Revista Linguí∫tica

Volume 22, número 1, jan./abr. de 2026

Organizadores da edição:
Christina Abreu Gomes (UFRJ),
Marcelo Alexandre Silva Lopes de Melo (UFRJ),
Maria da Conceição de Paiva (UFRJ)

Contribuições da Sociolinguística no Século XXI: teoria, interfaces e aplicações – uma homenagem a Anthony Julius Naro

Motivado pelo 4º Seminário PEUL: uma homenagem a Anthony J. Naro e aos 45 anos do PEUL, ocorrido em 2025, este número reúne trabalhos da área da Sociolinguística voltados para questões atuais dos estudos de variação e da mudança linguística, com base em diferentes abordagens teóricas e metodologias. Traz um dossiê com trabalhos apresentados no evento e uma seção com outros artigos da área da Sociolinguística.

ISSN: 2238-975X

UFRJ

Comissão Editorial

Editor-chefe

Ana Regina Vaz Calindro, Universidade Federal do Rio de Janeiro, Brasil

Editores

Aniela Improta França, Universidade Federal do Rio de Janeiro, Brasil
Isabella Lopes Pederneira, Universidade Federal do Rio de Janeiro, Brasil
Marcus Maia, Universidade Federal do Rio de Janeiro, Brasil
Marije Soto, Universidade Estadual do Rio de Janeiro, Brasil

Editores Honorários

Maria Luiza Braga, Universidade Federal do Rio de Janeiro, Brasil

Editor Fundador

Lilian Ferrari, Universidade Federal do Rio de Janeiro, Brasil

Conselho Editorial

Anthony Naro, Universidade Federal do Rio de Janeiro, Brasil (*in memoriam*)
Fernanda Ferreira Spoladore, Michigan State University, EUA
Gabriela Matos, Universidade de Lisboa, Portugal
Kees Hegenveld, Universidade de Amsterdam, Holanda
Leticia Sicuro Correa, Departamento de Letras - PUC/RJ
Leo Wetzels, Universidade Livre de Amsterdam, Holanda
Luiz Amaral, University of Massachusetts, USA
Maria Armanda Costa, Universidade de Lisboa, Portugal
Maria Cecília Mollica, Universidade Federal do Rio de Janeiro, Brasil
Roberto Almeida, Concordia University, Canada
Ruth Lopes, Universidade Estadual de Campinas, Brasil

Organização da Edição

Christina Abreu Gomes (UFRJ)
Marcelo Alexandre Silva Lopes de Melo (UFRJ)
Maria da Conceição de Paiva (UFRJ)

Redação e Assinaturas

Programa de Pós-Graduação em Linguística - Faculdade de Letras da UFRJ
Av. Horácio de Macedo, 2151 - Sala F. 321. Ilha do Fundão Cidade Universitária
CEP 21941-917 - Rio de Janeiro - RJ. E-mail: ppglinguistica@letras.ufrj.br

Editor Operacional e Editoração Eletrônica

Patricia Mabel Kelly Ramos, pattydesign/Brasil

Revista Linguí∫tica

Rio de Janeiro | volume 22 | número 1 | jan. - abr. de 2026

Contribuições da Sociolinguística no Século XXI: teoria, interfaces e aplicações – uma homenagem a Anthony Julius Naro |

Contributions of Sociolinguistics in the 21st Century: Theory, Interfaces, and Applications – A Tribute to Anthony Julius Naro

UFRJ

Sumário

Apresentação

7 **Contribuições recentes da pesquisa Sociolinguística sobre variação e mudança linguística na produção, percepção, avaliação e atitudes**
Christina Abreu Gomes, Marcelo Alexandre Silva Lopes de Melo e Maria da Conceição de Paiva

Homenagem

13 **No túnel do tempo com Anthony J. Naro: lembranças de Marta Scherre**
Maria Marta Pereira Scherre

Entrevista

24 **Entrevista com Ana Maria Carvalho (AC) por Xoán Carlos Lagares (XL)**
Ana Maria Carvalho e Xoán Carlos Lagares

Artigos: Variação e mudança em dados de produção

31 **As gramáticas dos gramáticos: a sintaxe dos clíticos pronominais em grammaticas históricas do português**
Saulo Rogério Pacheco Rocha e Marco Antonio Rocha Martins

54 **Variação no emprego de artigo definido diante de pronome possessivo na norma de Maputo**
Danielle Kely Gomes

74 **Construções conclusivas no português escrito: variação e gênero textual**
Mayra França Floret

96 **Variação dos pronomes de 2ª pessoa do singular com função de sujeito nas músicas de Poesia Acústica: análise sociofuncionalista sobre a influência da variável gênero**
Dennis Castanheira e Júlia Saadi

118 **Objetos acusativos de 3ª pessoa na variedade carioca do português brasileiro**
Adriana Martins Simões

154 **A expressão do sujeito pronominal em peças de teatro capixabas: uma análise em tempo real da 2ª pessoa do discurso**
Iesa Venturin Mulinari de Rodran e Leila Maria Tesch

188 **A alternância tu/você na escrita piauiense novecentista: os condicionamentos linguísticos e sociais**
Luiz Egito de Souza Barros e Márcia Cristina de Brito Rumeu

218 **Tu, você, nós e a gente: uma análise sociolinguística das variantes pronominais no português falado em São Luís – MA**
Elimária Oliveira Lima e Letícia Gantzias Abreu

239 **Sistematização de estudos sociolinguísticos sobre a representação da primeira pessoa do plural na fala alagoana: um olhar para a variável comunidade**
Elyne Giselle de Santana Lima Aguiar Vitório

Artigos: Percepção, atitudes/avaliação

260 **Crenças e atitudes da elite letrada frente à variação nas estratégias de relativização**
Dante Lucchesi

288 **O sotaque das capitais da região sudeste do Brasil: um estudo de avaliação**
Leila Maria Tesch

314 **Do papel à consciência: a escrita de cartas como instrumento de reflexão sobre a diversidade linguística**
Susana Menezes Araujo e Ormezinda Maria Ribeiro

339 **"Se você acha que capixaba não tem sotaque, fala arroz aí rapidinho": a ditongação na variedade capixaba**
Ana Carolina Carnieli Duarte e Raquel Gomes Chaves

359 **Gênero social e mudança linguística: reflexões teórico-metodológicas para a Sociolinguística**
Camila Bordonal Clempi e Angélica Rodrigues

382 **Amigos gaysssss: efeitos da duração de /-s/ final em percepções de gênero/sexualidade em vozes masculinas**
Marcus Garcia de Sene e Ronald Beline Mendes

406 **Pajubá como laboratório sociolinguístico no século XXI**
Renato Régis Barroso e Yonara Cristina de Souza dos Santos

Dossier: Seminário Peul

430 **Entre ø e forma explícita como marcador relacional em orações relativas: reorganização do sistema no PB**
Cecilia Mollica

448 **Análise sociolinguística da lenição de /d/ em gerúndios na região metropolitana do Rio de Janeiro**
Matheus Victor Alves Pereira e Eliete Figueira Batista da Silveira

http://dx.doi.org/10.31513/linguistica.2026.v21n2a75478

Contribuições recentes da pesquisa Sociolinguística sobre variação e mudança linguística na produção, percepção, avaliação e atitudes

Recent contributions of Sociolinguistic research on language variation and change based on production, perception, evaluation and attitudes

Christina Abreu Gomes[1]

Marcelo Alexandre S. Lopes de Melo[2]

Maria da Conceição de Paiva[3]

O tema do volume 22, número 1 da Revista Linguí∫tica, é voltado para questões atuais sobre a variação linguística como parte do conhecimento internalizado do falante. Foi inspirado no *4º Seminário do PEUL "Contribuições da Sociolinguística no século XXI – teoria, interfaces e aplicações"*. Esta edição do Seminário foi comemorativa dos 45 anos do grupo de pesquisa Programa de Estudos sobre o Uso da Língua (PEUL), sediado na UFRJ, e uma homenagem ao Professor Emérito Anthony Julius Naro (1942-2024), seu fundador. As edições do Seminário PEUL têm por objetivo estabelecer um espaço de diálogo entre pesquisadores da área sobre questões recentes no estudo da variação e da mudança. O 4º Seminário, realizado na Faculdade de Letras da UFRJ, em 03 e 04 de junho de 2025, contou com 39 apresentações orais de trabalho e 29 apresentações de posters de pesquisadores de diferentes instituições brasileiras (UERJ, UFF, UFES, UnB, UFSC, Seduc/UFT, IFTO, UFT, UFPA, UFSCAR, SME-RJ, Colégio Pedro II – RJ, UFS, UECE, além da UFRJ). Integraram a programação do Seminário duas palestras dos professores convidados Paul O´Neill (LMU, Munique) e Ana Maria Carvalho (University of Arizona). Como parte da programação, foi exibido o documentário "Vozes das Margens", de Ana Maria Carvalho, seguido de debate mediado pelo Professor Xóan Lagares da UFF. Todas as sessões do Seminário foram abertas irrestritamente para ouvintes. A sessão de exibição do documentário contou com a presença de mais de 100 estudantes de graduação da Faculdade de Letras da UFRJ.

A partir do estudo seminal de Labov (1963), em Martha's Vineyard, publicado na revista Word, intitulado *The social motivation of a sound change*, o desenvolvimento subsequente dos estudos sobre variação e mudança tem contribuído para questões atinentes à natureza da linguagem humana com evidências sobre o caráter heterogêneo do conhecimento linguístico internalizado e a não

[1] Universidade Federal do Rio de Janeiro (UFRJ), christinagomes@letras.ufrj.br, https://orcid.org/0000-0003-0358-2029.

[2] Universidade Federal do Rio de Janeiro (UFRJ), malmelo@letras.ufrj.br, https://orcid.org/0000-0002-8025-0530.

[3] Universidade Federal do Rio de Janeiro (UFRJ), paiva@club-internt.fr, https://orcid.org/0000-0002-8261-6575.

autonomia da língua em relação a usos e usuários, ao tratar da organização cognitiva da variação, da relação entre língua e sociedade, seja com base em dados de produção ou de percepção. Este número da Revista Linguí∫tica traz 18 artigos, selecionados com base no critério de avaliação duplo-cega por pares, organizados em duas seções: uma composta por artigos de trabalhos não relacionados a apresentações no *4º Seminário PEUL* e outra composta de artigos relacionados a apresentações no referido Seminário.

O presente número começa com uma homenagem de Maria Marta Pereira Scherre (UnB) ao queridíssimo Professor Anthony J. Naro, abordando diferentes aspectos de sua contribuição à universidade brasileira como pesquisador e professor, incluindo aspectos de sua cativante personalidade. Ainda, antecedendo os 18 artigos, há a entrevista da Professora Ana Maria Carvalho (University of Arizona) pelo Professor Xóan Lagares (UFF), tomando como base a instigante conversa, estabelecida entre os dois professores, sobre contato linguístico, português uruguaio e política e identidade linguística, na seção de debate, mediada pelo Professor Xóan Lagares, que se seguiu à exibição do documentário "*Vozes das Margens*", na manhã do dia 4 de julho, ambas como parte da programação do 4º Seminário PEUL. A entrevista é antecedida de uma apresentação da entrevistada pelo entrevistador, Xóan Lagares.

A seguir são apresentados os artigos que integram este número na ordem em que aparecem no sumário. Os artigos espelham a diversidade de temas relacionados à variação e à mudança linguística com dados do português brasileiro e de outras línguas, desenvolvidos com base em diferentes abordagens teóricas e metodologias. Os 9 primeiros artigos abordam a variação e a mudança em dados de produção. Os 7 artigos subsequentes são voltados para a percepção, avaliação e atitudes linguísticas. Os dois artigos do Dossiê 4º Seminário PEUL trazem estudos sobre variáveis linguísticas do PB em dados de produção coletados em amostras de fala espontânea.

O artigo de Saulo Rogério Pacheco Rocha e Marco Antônio Rocha Martins, fundamentado no modelo de Competição de Gramáticas, traz um estudo diacrônico sobre a mudança na colocação dos pronomes clíticos no português europeu, a partir de dados coletados em um corpus inédito de 6 gramáticas históricas do período compreendido entre os séculos XV e XIX. O artigo *As Gramáticas dos Gramáticos: a sintaxe dos clíticos pronominais em grammaticas históricas do português* mapeia a mudança na sintaxe de clíticos pronominais do Português Europeu de sistema majoritariamente proclítico para um sistema de predominância da ênclise.

O segundo artigo, de autoria de Danielle Kely Gomes, intitulado *Variação no emprego de artigo definido diante de pronome possessivo na norma de Maputo*, aborda a questão dos condicionamentos ou restrições linguísticas e sociais que atuam na variação de realização do artigo definido diante de possessivo na norma urbana do português falado em Moçambique. Essa variedade do português se caracteriza pelo contato entre o português e línguas autóctones do grupo Bantu. O estudo identifica condicionamentos de ordem estrutural e social em dados de uma amostra com falantes que se avaliam com diferentes graus de contato com o português ou como falantes do português como L1.

O artigo *Construções Conclusivas no português escrito: variação e gênero textual* de Mayra França Floret, tomando como referencial teórico os Modelos Baseados no Uso e da Gramática de Construções Baseada no Uso, investiga a variação no uso de construções conclusivas do português conectadas pelos esquemas [Preposição Demonstrativo Substantivo], [Preposição Demonstrativo] e [Preposição Substantivo Preposição Demonstrativo], com dados de amostra composta por textos de três gêneros distintos: redações do ENEM, textos de blog e artigos científicos. O artigo apresenta evidência de que os gêneros textuais atuam como um condicionamento relevante da variação em questão.

O artigo de Dennis Castanheira e Julia Saadi *Variação dos pronomes de 2ª pessoa do singular com função de sujeito nas músicas de poesia acústica: análise sociofuncionalista sobre influência da variável gênero*, observa a variação pronominal de 2ª pessoa do singular entre os pronomes "tu", "você" e "cê" com função de sujeito nas músicas do projeto *Poesia Acústica*, na perspectiva sociofuncionalista que conjuga a Sociolinguística Variacionista e o Funcionalismo Norte-Americano. O estudo examina o papel da variável gênero/sexo como condicionamento da variação.

O artigo *Objetos acusativos de 3ª pessoa na variedade carioca do português brasileiro*, de Adriana Martins Simões, analisa a realização variável do objeto acusativo anafórico de 3ª pessoa na variedade carioca, em uma abordagem que conjuga a teoria gerativista e a sociolinguística, que aponta para a competição de gramáticas que se manifesta sob a forma de variação linguística.

Iesa Venturin Mulinari de Rodran e Leila Maria Tesch, no artigo *A expressão do sujeito pronominal em peças de teatro capixabas: uma análise em tempo real da 2ª pessoa do discurso*, apresentam um estudo sobre a evolução da mudança na expressão pronominal da 2ª pessoa do discurso. O estudo foi desenvolvido com base em uma amostra composta por 12 peças teatrais capixabas, divididas em quatro períodos/sincronias, e publicadas entre os anos de 1889 e 2004. Foram observadas variáveis linguísticas e sociais. Os resultados indicam substituição de *tu* por *você* a partir da década de 1930, além do efeito da ambiguidade contextual e morfológica como grandes impulsionadoras da mudança, e o *locus* inicial da mudança em dados de personagens característicos de classes mais baixas e menos instruídas.

O artigo de Luiz Egito de Souza Barros e Márcia Cristina de Brito Rumeu, intitulado *A alternância tu/você na escrita piauiense novecentista: os condicionamentos linguísticos e sociais*, investiga a variação entre as formas de referência ao sujeito de segunda pessoa do singular na escrita de piauienses ao longo do século XX, à luz da Sociolinguística Histórica e da Teoria da Variação. Com base em cartas produzidas entre as décadas de 1920 e 1990, o estudo demonstra a predominância da forma "você", embora ainda em intensa alternância com "tu", identificando condicionamentos linguísticos e sociais que favorecem o emprego da forma inovadora no português brasileiro.

O artigo de Elimária Oliveira Lima e Letícia Gantzias Abreu, intitulado *Tu, você, nós e a gente: uma análise sociolinguística das variantes pronominais no português falado em São Luís*, analisa a variação no emprego dos pronomes de segunda e primeira pessoas no português ludovicense, tomando como referencial a Sociolinguística Variacionista. O estudo evidencia que as alternâncias "tu/você" e "nós/a gente" constituem um processo sistemático de variação, condicionado por fatores regionais,

linguísticos e pragmático-discursivos, revelando características particulares da variedade maranhense em relação ao quadro geral do português brasileiro.

Em *Sistematização de estudos sociolinguísticos sobre a representação da primeira pessoa do plural na fala alagoana: um olhar para a variável comunidade*, Elyne Giselle de Santana Lima Aguiar Vitório reúne resultados de diferentes pesquisas sobre a representação da primeira pessoa do plural na variedade alagoana, considerando o uso de *nós* e *a gente* nas posições de sujeito e de não sujeito, bem como os padrões de concordância verbal. A análise demonstra que o tipo de comunidade constitui um importante condicionamento da variação, evidenciando diferenças entre comunidades urbanas e rurais quanto ao comportamento das variantes pronominais e das estratégias de concordância.

Dante Lucchesi, no artigo *Crenças e atitudes da elite letrada frente à variação nas estratégias de relativização*, apresenta um estudo de percepção linguística sobre a avaliação das diferentes estratégias de relativização do português brasileiro por integrantes da elite letrada de Salvador e do Rio de Janeiro. Os resultados revelam a persistência da valorização da variante prescrita pela tradição gramatical, mesmo quando pouco utilizada na fala cotidiana, além de identificarem perfis distintos de avaliadores, refletindo a força da ideologia normativista e diferentes posicionamentos diante da variação linguística.

O artigo de Leila Maria Tesch, *O sotaque das capitais da Região Sudeste do Brasil: um estudo de avaliação*, investiga a percepção dos sotaques das capitais da Região Sudeste a partir dos pressupostos da Sociolinguística e dos estudos de percepção e avaliação linguística. Com base em respostas obtidas por meio de questionário on-line, o estudo demonstra que os sotaques de Rio de Janeiro, São Paulo, Belo Horizonte e Vitória recebem avaliações distintas quanto ao prestígio, à beleza, à correção e à agradabilidade, evidenciando o papel dos estereótipos linguísticos e sociais na construção dessas percepções.

No artigo intitulado *Do papel à consciência: a escrita de cartas como instrumento de reflexão sobre a diversidade linguística*, Susana Menezes Araujo e Ormezinda Maria Ribeiro apresentam uma proposta de intervenção pedagógica voltada à conscientização de estudantes do Ensino Médio sobre a legitimidade da variação linguística e os efeitos do preconceito linguístico. O estudo, fundamentado na Sociolinguística Educacional, foi desenvolvido a partir de oficinas em uma escola pública do Distrito Federal e mostra que a produção de cartas contribuiu para ampliar a consciência dos alunos acerca da diversidade linguística e para promover atitudes de respeito às diferentes variedades do português brasileiro.

O artigo de Ana Carolina Carnieli Duarte e Raquel Gomes Chaves, *"Se você acha que capixaba não tem sotaque, fala arroz aí rapidinho": a ditongação na variedade capixaba*, investiga o processo de ditongação com glide /j/ na fala da Região Metropolitana da Grande Vitória, à luz da Teoria da Variação e Mudança Linguística e da Fonética Acústica. Os resultados apontam que o fenômeno é condicionado por fatores linguísticos e sociais, como faixa etária, posição do acento, qualidade da vogal e estilo de fala, sugerindo que a ditongação constitui um processo sensível ao monitoramento estilístico e às características fonológicas dos itens lexicais.

Camila Bordonal Clempi e Angélica Rodrigues, em *Gênero social e mudança linguística: reflexões teórico-metodológicas para a Sociolinguística*, propõem uma releitura da variável gênero a partir das teorias da performatividade e da tecnologia de gênero, afastando-se de interpretações essencialistas. A partir da análise da alternância entre marcas de segunda e terceira pessoas em cartas publicadas na imprensa do início do século XX, o estudo demonstra que os contextos discursivos e as representações de gênero influenciam o comportamento das variantes, oferecendo novas perspectivas para a compreensão do significado social da mudança linguística.

Em *Amigos gaysssss: efeitos da duração de /-s/ final em percepções de gênero/sexualidade em vozes masculinas*, Marcus Garcia de Senne e Ronald Beline Mendes apresentam um experimento de percepção baseado na técnica matched-guise para investigar o papel da duração da fricativa /-s/ em final de palavra na construção de percepções de gênero e sexualidade. Os resultados mostram que o alongamento da fricativa favorece a percepção de vozes masculinas como menos masculinas ou mais associadas à identidade gay, evidenciando o papel indexical dessa variável fonética no português brasileiro.

O artigo de Renato Régis Barroso e Yonara Cristina de Souza dos Santos, intitulado *Pajubá como laboratório sociolinguístico no século XXI*, analisa o Pajubá como um repertório linguístico estruturado e socialmente regulado, articulando pressupostos da Sociolinguística Variacionista, da Linguística Queer e da Análise do Discurso Crítica. A pesquisa demonstra que o uso do Pajubá é sistemático e condicionado por fatores identitários, performativos e contextuais, reafirmando sua relevância para a compreensão das relações entre variação linguística, identidade e resistência social.

O artigo de Maria Cecília de Magalhães Mollica, *Entre Ø e forma explícita como marcador relacional em orações relativas: reorganização do sistema no PB*, investiga a alternância entre o apagamento e a realização explícita de preposições em orações relativas do português brasileiro, com base em dados de fala espontânea da amostra Recontato/2000 (Projeto PEUL). O estudo demonstra que a variação é condicionada principalmente por fatores internos ao sistema linguístico e interpreta os resultados como evidência de uma reorganização funcional do sistema relacional do português brasileiro, e não como um processo de simplificação gramatical.

Em *Análise sociolinguística da lenição de /d/ em gerúndios na Região Metropolitana do Rio de Janeiro*, Matheus Victor Alves Pereira e Eliete Figueira Batista da Silveira investigam o processo de lenição da consoante /d/ em formas de gerúndio na fala da Região Metropolitana do Rio de Janeiro, à luz da Sociolinguística Variacionista. As análises partem de dados do Corpus Concordância, do Projeto CORPORAPORT, composto por entrevistas com falantes das duas localidades, Nova Iguaçu e Copacabana o estudo identifica elevada frequência de redução do segmento. Os resultados demonstram que fatores como escolaridade, contexto fonético-fonológico, frequência lexical e informante condicionam a variação, sugerindo que o fenômeno representa uma possível mudança em curso no português brasileiro.

Esperamos que os leitores apreciem os artigos publicados nesse número da Revista Linguí∫tica e que se sintam estimulados a avançar o debate sobre os diferentes aspectos da relação entre língua e sociedade na abordagem da pesquisa sociolinguística.

http://dx.doi.org/10.31513/linguistica.2026.v21n2a75285

No túnel do tempo com Anthony J. Naro: lembranças de Marta Scherre

Through the time tunnel with Anthony J. Naro: recollections by Marta Scherre

Maria Marta Pereira Scherre[1]

Nessa minha singela homenagem ao eterno orientador, amigo e colega acadêmico Naro, nosso apelido carinhoso para Anthony Julius Naro, faço inicialmente três destaques.

O primeiro destaque vai para a sua grande habilidade com relação à linguagem humana, que já se evidencia na guinada de sua formação acadêmica nos Estados Unidos, da graduação em Matemática, de 1959 a 1963 (na Polytechnic University), para o Doutorado (PhD) em Linguística, de 1963 a 1968, sob a orientação de Paul Kiparsky (no MIT - Massachusetts Institute of Technology).

O segundo destaque é a capacidade aglutinadora de Naro, sem nunca querer ser líder, mas o fato é que sempre foi. Por meio de sua atuação **é** que a Sociolinguística Variacionista brasileira alcançou patamares internacionais. Seguramente, Naro foi o grande homem da internacionalização com relação aos estudos sociolinguísticos sobre o português brasileiro. Ele foi um pesquisador e divulgador além do seu tempo, quando ainda não se enfatizava a grande importância da internacionalização na Academia, na minha maneira simples de ver o mundo. Indiretamente, Naro foi o grande responsável pela criação do Programa de Estudos sobre o Uso da Língua (PEUL), o primeiro grupo de Sociolinguística no Brasil, organizado na Faculdade de Letras da Universidade Federal do Rio de Janeiro (UFRJ), desde a década de 1980, com sua primeira sede em uma das salas da Faculdade de Letras ainda na Avenida Chile, no centro do Rio de Janeiro, e depois na sala F-318 do prédio da Faculdade de Letras, na ilha do Fundão.

O terceiro destaque é a imensa generosidade de Naro com relação a todos os membros do PEUL e, em especial, no trato com seus orientandos e colegas da Academia ao longo de mais 45 anos de atividades no Brasil: de 1974 a 2022. No meu caso, a nossa frutífera parceria em eventos nacionais e internacionais se deu de 1982 a 2017, ao longo de 35 anos, com cerca de 54 apresentações fora do Brasil, que ocorreram, em especial, no evento New Ways of Analyzing Variation (NWAV), de 1990 a 2017, por 27 anos, em cerca de 22 NWAVs, com 26 apresentações, 22 da dupla Naro e Scherre e Scherre e Naro, em quantidade equilibrada e 04 com outros colegas da Universidade de Brasília (UnB) e da Universidade Federal de Espírito Santo (UFES). Repito: Naro era naturalmente o homem da internacionalização.

[1] Universidade Federal do Espírito Santo (UFES), Centro de Ciências Humanas e Naturais, mscherre@gmail.com, https://orcid.org/0000-0003-2977-0431.

"Naro ainda está aqui!" Ele vive nas lembranças de frutíferas parcerias acadêmicas, cercadas pela floresta e pelas árvores da amizade e da saudade. A certeza de que "Naro ainda está aqui" se evidencia com o seu último texto, escrito em 2022 e publicado em 2024 no livro *Uma Estrada da Sociolinguística Brasileira*: homenagem à trajetória acadêmica de Marta Scherre, sob a organização de Cíntia Pacheco e Carolina Andrade, ambas da Universidade de Brasília, minhas companheiras de estrada acadêmica há cerca de 20 anos. A seguir, passo a palavra ao homenageado Naro por meio desse último texto. Certamente ele está me ouvindo feliz, uma homenagem de sua homenageada.

Naro, me permite negritar algumas de suas palavras? A gente fazia muito isso em nossas inúmeras escritas eletrônicas, também com coloridos. Eis, então, as palavras do texto Memorial de Anthony Julius Naro (2024, p. 29).

A dupla Scherre e Naro (e Naro e Scherre) começou em meados da década dos anos 70 do século passado. No início, Marta era aluna de mestrado; depois, colaboradora no projeto Competências Básicas; logo, aluna de doutorado; depois, colega de departamento na UFRJ; chefe do departamento; colega de pesquisa, e por aí vai... Preenchemos uma série de pré-requisitos necessários ao sucesso acadêmico, a saber: uma base teórico-metodológica aceita pelos colaboradores e uma base de dados compartilhada. Mas temos abordagens complementares nessas áreas. **Marta prefere trabalhar até os mínimos detalhes desde a codificação inicial até um número sem fim de rodadas com diferenças mínimas de combinações de variáveis, para, através desse caminho, chegar às conclusões e subsequentes publicações; eu prefiro uma visão mais distante, abstraída dos detalhes, na tentativa de alcançar tendências gerais. A abordagem da Marta vai da árvore até a floresta; a minha vai da floresta até a árvore. Os dois percorremos o mesmo terreno, mas usando mapas diferentes, até que chegamos ao mesmo lugar. Para Marta, o trabalho científico se funda na realidade sincrônica da comunidade atual de fala; eu prefiro incluir e até me fundamentar na deriva diacrônica que levou à situação atual. As duas metodologias complementares do trabalho científico resultam em caminhos analíticos complementares que chegam a conclusões em comum. A parceria, apesar de transitar por vias diferentes, tem funcionado bem, porque cada um de nós sabe se colocar na perspectiva do outro**. (Naro, 2024, p. 29).

Nessa homenagem, convido o leitor a transitar por nosso último texto em parceria, um texto de Marta Scherre, Anthony Naro e Lilian Yacovenco publicado em 2018, em um número especial da revista *Diadorim* do Rio de Janeiro, em homenagem à ilustre colega Dinah Callou.

Minhas lembranças no início do túnel do tempo remontam aos anos de 1974 a 1978, quando eu era aluna do Mestrado em Letras, na Pontifícia Universidade Católica do Rio de Janeiro – PUC/Rio, com Naro orientador de Mestrado, um orientador exemplar. Relato aqui alguns momentos pitorescos com Naro, por meio de diálogos que me vêm memória.

A primeira disciplina no Mestrado, no primeiro semestre de 1974, foi de sintaxe gerativa avançada. Naro fez o seguinte comentário escrito no meu trabalho final dessa disciplina, em que havia representações sintáticas por meio de árvores: "Suas árvores parecem de mato selvagem". Mesmo

assim, fui aprovada na disciplina, mas não com nota máxima. Depois fiz mais duas disciplinas de sintaxe gerativa para iniciantes, também com Naro, em semestres consecutivos. E ele me perguntou no primeiro dia de aula: "Marta, por que você está fazendo sintaxe gerativa para bebês, se você já fez sintaxe gerativa avançada?" E eu, calmamente, respondi que era para entender o que eu não havia entendido na disciplina de sintaxe gerativa avançada. E ele se limitou a sorrir. Na segunda disciplina de sintaxe gerativa para iniciantes, em que me matriculei como aluna efetiva, obtive a nota máxima, acrescida de um comentário elogioso. Em um certo dia, em encontro festivo na casa dele no final do ano, no Dia de Ação de Graças americano, quando ele reunia colegas, alunos e orientandos na sua casa em Copacabana, ele veio todo sorridente abrir e porta e falou para meu esposo (Anecir Scherre, de apelido Lessa): "Lessa, você sabia que a Marta recebeu nota máxima (A+) na disciplina que fez comigo nesse segundo semestre?" Foi um momento muito estimulante, para uma mineira do interior que estava cursando o Mestrado na PUC-RIO.

Voltando novamente no tempo: após o final da disciplina em sintaxe avançada, Naro me fez uma pergunta, verdadeiramente inusitada para mim, à época: "Marta, você gostaria de nos auxiliar nas entrevistas de uma pesquisa com os falantes que cursam o Movimento Brasileiro de Alfabetização (MOBRAL)?" Eu disse: "Claro que sim Naro, mas a questão é que eu não sei fazer pesquisa. Não tive essa oportunidade no meu curso de graduação no interior de Minas Gerais, de 1970 a 1973, em Coronel Fabriciano, onde havia um campus da PUC-Minas. Resposta tranquila e serena do Naro: "Não precisa saber fazer pesquisa não. A gente ensina tudo. É só ter boa vontade e gostar de aprender". E foi a minha sorte grande na vida de pesquisadora, iniciada pelas mãos do Naro.

Assim, de 1974 a 1977, fui auxiliar de pesquisa do *Projeto Competências Básicas do Português*, que Naro desenvolveu com a Professora Miriam Lemle, a convite da Professora Miriam, o primeiro projeto de pesquisa sob os princípios teóricos e metodológicos da Sociolinguìstica Variacionista Laboviana realizado com dados do português brasileiro falado pelos alunos que frequentavam o MOBRAL no Rio de Janeiro. O relatório final desse projeto de pesquisa seminal para os estudos variacionistas brasileiros foi publicado por Lemle e Naro (1977). O contato de Naro com William Labov se iniciou exatamente à época da realização desse projeto, porque as instituições patrocinadoras Fundação MOBRAL e Fundação Ford especificaram que haveria necessidade de tratamento estatístico dos dados de fala que seriam coletados. Naro solicitou a Labov a indicação de uma pessoa que pudesse vir ao Brasil prestar assessoria nas análises estatísticas das pesquisas sociolinguísticas a partir da base de dados organizada. Labov indicou Gregory Guy, que à época era aluno de Doutorado sob orientação de Labov, na Universidade da Pensilvânia, nos Estados Unidos. Segundo depoimento do Naro, que consta da nota 1 do texto de Marta Scherre e Cláudia Roncarati (2008, p. 49), "seu primeiro contato com a Teoria da Variação se deu quando, em 1974, ele foi convidado pela professora Miriam Lemle a participar do projeto de pesquisa 'Competências Básicas do Português', financiado pelo MOBRAL e pela Fundação Ford."

Outro momento pitoresco com Naro ocorreu em um certo momento da etapa de análise estatística de dados de 10 falantes da cidade do Rio de Janeiro para a minha Dissertação de Mestrado. Depois de múltiplas *rodadas* feitas por mim, Naro me pergunta: "Marta, quando você vai começar a escrever sua Dissertação?" Respondi: "Uai, quando eu terminar a análise estatística dos dados". Ele retrucou: "Para o Mestrado, você já poderia ter começado a escrever há seis meses. Seus resultados já estão ótimos. Não precisa fazer nenhuma rodada mais. Você já tem resultados para uma Tese de Doutorado nos padrões americanos: dados inéditos, resultados inéditos e análise coerente." Lá fui eu escrever minha Dissertação de Mestrado, ainda com vontade de fazer mais rodadas, de olhar mais atentamente o caminho das árvores para tentar chegar à floresta. Tudo deu certo. Defendi minha Dissertação de Mestrado (*A regra de concordância nominal em português*) em 1978.

É fundamental salientar que o primeiro semestre de 1975 foi o marco histórico da ampliação dos estudos sociolinguísticos brasileiros, como se pode ver, a seguir, na longa citação de Scherre e Roncarati (2008, p. 37), necessária nessa homenagem à grandiosidade de Naro no Brasil.

> Em [1975], o Professor Anthony Julius Naro daria uma disciplina sobre línguas em contato no curso de mestrado da PUC-RJ. Antes de iniciar a primeira aula, as alunas perguntaram se o referido professor não poderia ministrar um curso de Teoria da Variação Linguística de linha laboviana. Com a simplicidade científica que lhe é peculiar, o professor Naro respondeu que, naquele dia, não poderia iniciar um curso sobre Teoria da Variação, porque não havia preparado a aula naquele sentido. Portanto, os alunos teriam de ouvi-lo falar sobre línguas em contato, mas, na semana seguinte, ele poderia recomeçar a disciplina falando sobre Teoria da Variação. E assim aconteceu. Ali, naquele momento, estava plantada a semente das pesquisas variacionistas sobre o Português falado no Brasil, que germinou através de inúmeras dissertações de mestrado e [de teses] de doutorado sob orientação do professor Naro; floresceu plenamente com a formação de um grupo de pesquisa, [o PEUL, o *Grupo do Rio*], sob sua coordenação, com base inicialmente, nos princípios teóricos e metodológicos as Sociolinguística Laboviana, denominada também de Teoria da Variação Linguística (Laboviana) ou de Sociolinguística Quantitativa, e se ampliou no amplo território brasileiro, com a formação de novos grupos assessorados pelo *Grupo do Rio*, entre os quais se destacaram o VARSUL (*Variação Linguística Urbana da Região Sul*) e o VALPB (*Variação Linguística na Paraíba*).

Para evidenciar que, realmente, "Naro ainda está aqui!", volto ao túnel do tempo no ano de 1979, quando ocorreu o *New Ways of Analyzing Variation in English* (NWAVE) *VIII*, que contou com a participação presencial de Naro e Sebastião Votre, o primeiro orientando de Doutorado de Naro, que defendeu sua Tese de Doutorado na PUC-Rio, em 1978, sob o título *Aspectos da variação fonológica na fala dos analfabetos do Rio de Janeiro*, e participante do *Grupo do Rio* à época de sua criação, em 1979. O NWAVE VIII, título que ainda terminava com *in English*, foi realizado também em 1979 na Universidade de Quebec em Montreal, no Canadá. Só em 1984, no NWAV XIII, pela primeira na Universidade da Pensilvânia, na Filadélfia, é que eliminaram *in English*, tendo em vista o fato de que participantes de diversas outras línguas passaram a participar desse evento ímpar para as pesquisas variacionistas.

Só Naro e Sebastião Votre foram ao NWAVE VIII, em 1979, evento internacional histórico para o *Grupo do Rio* e para o Brasil. Por informações de Sebastião Votre e de um e–mail do Naro para mim em 2014, hoje sei que Naro foi convidado por um dos organizadores desse evento, David Sankoff, para falar das pesquisas variacionistas que estavam sendo realizadas no Rio de Janeiro pelos membros do *Grupo do Rio*. Nesse evento, Naro e Sebastião Votre apresentaram trabalhos autorais. Além disso, Naro sintetizou e citou cinco pesquisas de três orientandas suas, com Dissertações de Mestrado defendidas em 1997 e 1978, e duas outras pesquisas individuais, de participantes da disciplina histórica em 1975 na PUC-Rio e integrantes do *Grupo do Rio*, que recebeu o nome de PEUL: Programa de Estudos sobre o Uso da Língua a partir de 1987, um grupo que sempre se renovou (*cf*. Scherre; Roncarati, 2008, p. 37-38, 39, 40-43). O Grupo do Rio, atual PEUL, sempre foi sediado na Faculdade de Letras da Universidade Federal do Rio de Janeiro, como consta das referências bibliográficas deste texto.

A seguir, seguem detalhes das cinco pesquisas sintetizadas e/ou citadas por Naro no NWAV VIII, em 1979:

1) Helena Gryner e Alzira Tavares de Macedo, sobre a pronúncia do S pós-vocálico (Pesquisa individual: Alzira Macedo entrou para o *Grupo do Rio* em 1982 e Helena Gryner, em 1992);

2) Giselle Machline de Oliveira e Silva, sobre pronomes de segunda pessoa, com foco na forma *você* (Pesquisa individual: integrante da equipe de criação do *Grupo do Rio*, em 1979);

3) Maria Cecília de Magalhães Mollica, sobre o uso do pronome cópia nas construções relativas (Mestrado em 1977 – PUC-RJ, sob orientação de Naro, participante do *Grupo do Rio* a partir de 1980);

4) Maria Marta Pereira Scherre, sobre a concordância de número entre os elementos dos sintagmas nominais (Mestrado em 1978 – PUC-RJ, sob orientação de Naro, participante do *Grupo do Rio* a partir de 1980); e

5) Nelize Pires de Omena, sobre os pronomes pessoais de terceira pessoa em função de objeto direto (Mestrado em 1978 – PUC-RJ, sob orientação de Naro, participante do *Grupo do Rio* a partir de 1980).

Os organizadores do NWAV VIII abriram espaço para o envio dos textos dos sete trabalhos brasileiros para o livro *Variation Omnibus,* publicado em 1981. No prefácio desse livro, David Sankoff e Henrietta Cerdergren registraram e batizaram o grupo liderado por Naro:

> We are happy to have welcomed an unusually large number of innovative and well-executed contributions by overseas variationists, and particularly to have introduced the work of **the Rio de Janeiro group** to NWAVE." (Sankoff; Cedergren, 1981, p. ix, negritos nossos).

David Sankoff e Henrietta Cerdergren (1981, p. ix) também relatam no *Preface* do *Variation Omnibus* que "[t]he scope and the depth of the papers presented, almost all which are printed here, are clear measures of the dynamism of this field."

O *Variation Omnibus* apresenta 61 textos organizados em quatro grandes temas, em caixa alta, sem os detalhes entre parêntesis: MULTI-VARIABLE STUDIES (três textos); PLURAL MARKING IN ROMANCE LANGUAGES & RELATED SUBJECTS (18 textos); SYNTACTIC VARIATION (24 textos); CONTEXT, DOMAIN & REGISTER (16 textos).

Os sete textos sobre o português brasileiro se distribuem da seguinte forma. Em *PLURAL MARKING IN ROMANCE LANGUAGES* estão: Sebastião J. Votre (Universidade Gama Filho e Universidade Federal do Rio de Janeiro) Phonological and syntactic aspects of denasalization in spoken Brazilian Portuguese; Maria Marta Pereira Scherre (Pontifícia Universidade Católica e Universidade Federal, Rio de Janeiro) Variation de la règle d'áccord du nombre dans le syntagme nominal en Portugaise; e Helena Gryner e Alzira Tavares de Macedo (Université Fedérale, Rio de Janeiro) La prononciation du S post-vocalique: deux processus de changement linguistic en portugais. Em *SYNTACTIC VARIATION* estão: Maria Cecília M. Mollica (Pontifícia Universidade Católica & Universidade Federal, Rio de Janeiro) La copie dans les constructions relatives em Portugais; Nelize Omena (Pontifícia Universidade Católica, Rio de Janeiro) Pronom personnel de la troisieme personne: ses formes variantes en fonction accusative; e Anthony J. Naro (Universidade Federal & Pontifícia Universidade Católica, Rio de Janeiro) Morphological constraints on subject deletion. Em *CONTEXT DOMAIN & REGI*STER está Giselle Machline de Oliveira e Silva (Universidade Federal Rural, Rio de Janeiro) Perspective sociolinguistique de la forme você à Rio de Janeiro.

Os cinco trabalhos dos participantes do *Grupo do Rio* que não estavam lá foram publicados em francês. A tradução de alguns deles foi feita por Giselle Machline, francesa de nascimento e brasileira de vivência e de coração, colega também generosa, característica de todos os participantes do PEUL. Naro não fez a tradução dos textos, porque ele se dizia incapaz de fazer traduções de textos corridos para o inglês. Nessas situações, a intuição do inglês não lhe vinha facilmente à mente, pela interferência do português brasileiro, que já falava e escrevia com a maestria de um nativo.

Aliás, ele me dizia, com ar de brincadeira, que ele já era "um sem língua''. Para textos nossos publicados em inglês, quando da dupla Naro e Scherre, o líder da redação era ele, que escrevia diretamente em inglês, e eu participava lendo, sugerindo alterações e/ou acréscimos sempre em tópicos, em português, a pedido dele. No caso da dupla Scherre e Naro, eu era a líder das ideias, mas escrevia tudo em grandes tópicos em português, para, novamente, não interferir na intuição do inglês dele. Depois eu lia e relia os textos corridos em inglês, fazendo anotações pertinentes em português. A gente escrevia e reescrevia inúmeras vezes nossas ideias e usávamos coloridos para nossas observações. Jamais usávamos inserções eletrônicas. Ele se dizia incapaz de dominá-las. Eu fazia essa tarefa quando necessário fosse. De vez em quando, Naro repetia uma frase que ainda permeia minha mente quando estou escrevendo, como agora. "Marta, já passou da hora de a gente fechar esse texto. Ou a gente acaba com o texto ou texto acaba com a gente".

Depois que terminei minha Tese de Doutorado, uma verticalização de análise da concordância nominal, com dados de 64 falantes coletados pelo PEUL na década da 1980, a Amostra Censo 1980, resolvi analisar os dados da concordância verbal nessa mesma base de dados, porque Naro estava trabalhando com Sebastião Votre, na linha funcionalista. Comecei a apresentar resultados sobre concordância verbal em reuniões do PEUL para exposição de resultados de pesquisas dos participantes. Certo dia, em uma de nossas múltiplas conversas, Naro me perguntou se podia participar da pesquisa sobre concordância verbal que eu estava dando continuidade. Respondi a ele com lágrimas de alegria nos olhos: "Naro, trabalhar com você, como colega, é a maior honra da minha vida acadêmica. Obviamente que a minha resposta é sim!!!" Aí, Naro me disse: "Marta, só vou te dizer um detalhe muito importante: é mais fácil ser minha orientanda do que minha colega de trabalho. Como colega, eu não sou muito consistente. Mudo muito de ideia. Enfim, estou te avisando para você não ficar chateada comigo e ter bastante paciência". E aí nossa conversa fluiu por uns bons minutos e selamos a nossa nova parceria.

O primeiro grande teste da nossa parceria como colegas de trabalho ocorreu em 1989, após a minha defesa de Tese de Doutorado, que ocorreu em 1988, com o título de *Reanálise da concordância nominal em português*, já na UFRJ, onde eu já trabalhava desde 1993.

Eu sempre gostei de viajar. Era um sonho desde criança, pouco realizado. Fui morar no Rio de Janeiro em 1974, quando Lessa e eu nos casamos. Lessa é o apelido antigo do Anecir, que vem dos tempos de futebol, na década de 1960, arte que ele ainda pratica até hoje. Meu sonho era, inicialmente, conhecer o nosso Brasil, de dimensões continentais e, depois, conhecer outros países. Lessa me dizia que devíamos *fazer o pé-de-meia* primeiro, para só depois viajar. Após 15 de casados, um dia, em uma conversa amorosamente acirrada, eu falei com Lessa: "meu amor, parece que nosso *pé-de-meia* tá furado, porque nunca chega o momento de a gente viajar para, pelo menos, conhecer o Brasil." Conversa vai, conversa vem, logo entendi que, na verdade, Lessa não gostava de viajar e eu levei 15 anos de casada para entender isso. Acho que ele mesmo não sabia direito. Essa constatação fez eu dizer a ele: "Lessa, vou viajar e conhecer o mundo pelo meu trabalho, que vai me permitir galgar os degraus da minha carreira acadêmica e realizar parte dos meus sonhos. Hoje mesmo vou fazer dois resumos e, amanhã, vou à casa do Naro e pedir para ele traduzir e submeter ao 19º New Ways of Analyzing Variation (NWAV) que ia se realizar em 1990 na Universidade da Pensilvânia, na Filadélfia, Estados Unidos. E assim aconteceu.

Organizei dois resumos durante a noite e no dia seguinte fui à casa do Naro e disse pra ele. "Naro, fiz dois resumos, um Scherre e Naro e outro Naro e Scherre para você ler e ver se são resumos que têm chance de serem aceitos para a apresentação no NWAV XIX. Naro me disse que gostava de viajar, mas não gostava muito de ir a congressos. Mesmo assim aceitou minha ousadia e leu os dois resumos. Na sequência, ele me disse que os dois resumos tinham mais o meu jeito de ser do que o dele. Assim, ele fez a versão do Naro e Scherre sem grandes alterações, mas refez completamente o de Naro e Scherre, de tal forma que, para o NWAV XX, a se realizar em Washington, DC, eu pude aproveitar o resumo que eu tinha feito para o que seria o Naro e Scherre. O fato é que tivemos os dois

abstracts feitos para a sessão de *papers* no padrão NWAV, que, à época e por muitos anos, só aceitava submissão de *abstracts* de pesquisas consolidadas, com tratamento estatístico adequado, realizado em sua maioria pelos programas da série VARBRUL, que tomaram corpo a partir de década de 1970. Registro aqui a maestria do Naro em fazer *abstracts* no padrão NWAV, que segue as orientações da Linguistic Society of America (LSA).

Enfim, como era de se esperar, os *abstracts* foram aceitos e a dupla Scherre e Naro e Naro e Scherre iniciou sua longa caminhada por cerca de 26 NWAVs, de 1990 a 2017, com os dois primeiros trabalhos apresentados no *XIX NEW WAYS OF ANALYZING VARIATION (NWAVE 19) - 1990 UNIVERSIDADE DA PENSILVÂNIA*, Philadelphia**,** sob os seguintes títulos: SCHERRE, Maria Marta Pereira; NARO, Anthony Julius. *Marking in Discourse: Birds of a Feather*; e NARO, Anthony Julius; SCHERRE, Maria Marta Pereira. *Linguistic Change and Variation: Flows and Counterflows in the Speech Community*. Viajamos para Filadélfia, apresentamos nossos trabalhos e publicamos o texto de Scherre e Naro na *Language Variation and Change* na primeira edição de 1991. O de Naro e Scherre já havia sido enviado para publicação no *Cadernos de Estudos Linguísticos* da Unicamp, que saiu ainda em 1991.

Assim, 11 anos após o *Grupo do Rio* ser apresentado ao NWAV, a dupla Naro e Scherre e Scherre e Naro voltou a levar pesquisas variacionistas sobre o português brasileiro para o esse principal e importante evento internacional, que ocorre ora nos Estados Unidos ora no Canadá, e já está na sua versão 54. O NWAV 54 acontecerá de 3 a 4 de outubro de 2026 na Universidade de Montréal, no Canadá, país onde tudo começou para o Brasil pelas mãos do Naro, em companhia de Sebastião Votre, também um forte estimulador para que outros participantes do *Grupo do Rio* também se inserissem no processo de intercionalização. E assim aconteceu com muitos pesquisadores brasileiros, muito além do *Grupo do Rio,* uma outra longa e boa história ainda por contar.

No final da minha viagem pelo túnel do tempo, sempre recordando que "Naro ainda está aqui", faço um relato de uma longa viagem que a dupla Naro e Scherre e Scherre e Naro fez Canadá para participar do NWAV 45, em 2016, gentilmente acompanhados por Lessa e Wanderley. O NWAV 45 foi organizado pela Universidade Simon Fraser e pela Universidade de Vitória, em Vancouver, Canadá. O evento ocorreu no *SFU Harbour Centre campus*, o campus central da *Simon Fraser University* no centro de Vancouver.

Vamos então contar um pouco dessa história, com a última apresentação da dupla Naro e Scherre e Scherre e Naro, em que Naro evidencia mais uma vez sua imensa generosidade.

Inicialmente, só o Naro é que havia recebido o convite dos organizadores do NWAV 45 para fazer uma palestra de abertura do evento, uma grande honra. No convite ao Naro, mencionaram meu nome:

> Dear Tony, Simon Fraser University and University of Victoria have the honour of hosting NWAV 45 in 2016 (…) in Vancouver, BC, Canada. The dates that we have in mind are November 3-6 2016. (…) (…) We are therefore writing you in the sincere hope that you would be willing to participate as a plenary for NWAV 45. The topic we have in mind–though you would of course have freedom in shaping the talk as best suits you–concerning your work with Maria Marta Pereira Scherre and colleagues. (…) (…) (…) (…). Most sincerely,

> Panos Pappas, SFU Alexandra D´Arcy, UVic (Pappas, Panos; D´Arcy, Alexandra – e-mail enviado a Anthony J Naro, em 14 de novembro de 2014).

No mesmo dia, Naro me encaminhou o e-mail recebido e me escreveu o seguinte:

> Marta, veja este convite. Mencionaram seu nome explicitamente. Mandaram para você também? Dizem até que vão pagar as despesas, como o David fez no primeiro NWAV que eu fui. Se você topar, vou responder que aceitaMOS (e que podemos dividir o pagamento também, caso não tenham mandado para você). Um abraço plenário, Naro (Naro, Anthony J. – e-mail enviado a Marta Scherre, em 14 de novembro de 2014).

Aceitei o convite de Naro, com honra imensurável, mas com a seguinte principal ponderação: "Mencionaram sim meu nome mas eu não recebi o convite" (Scherre, Maria Marta Pereira, e-mail enviado ao Naro, em 14 de novembro de 2014). Mesmo assim, Naro enviou a seguinte mensagem aos organizadores do NWAV 45:

> Dear Panos & Alexandra: It will be a pleasure and an honor for me to participate in the 2016 NWAV in a plenary session. As you note in your kind invitation, my work on the topic of standardization and change within the community has been carried out in collaboration with my colleague Maria Marta Pereira Scherre, who usually takes part in NWAV together with me. For this reason, I have invited her to participate as co-author and she has accepted. (…). Of course, 2016 is a long way off so we have lots of time to work out details. Tony (Naro, Anthony J. – e-mail enviado a Panos e Alexandra, em 15 de novembro de 2014).

Os organizadores do NWAV 45 enviaram a seguinte resposta ao Naro, com cópia mim:

> Dear Tony, We are thrilled that you have accepted our invitation, and we are very happy to have Marta present with you. Together your work is formidable, and this is a great honour for us. (…) (…) For now, let us simply say Thank You, and we are very much looking forward to working with you both toward the success of NWAV 45. (…) (…) (…) Most sincerely, Alex and Panos (Pappas, Panos; D´Arcy, Alexandra – e-mail enviado a Naro, com cópia para mscherre@gmail.com, em 16 de novembro de 2014).

Na sequência, Panos nos enviou a seguinte mensagem:

> Hello Tony e Marta I Would also like do add how excited I am about your acceptance. We are looking forward to hosting you both In Vancouver in 2016 Panos (Pappas, Panos - e-mail enviado a Naro e Marta, em 16 de novembro de 2014).

No dia seguinte, eu enviei a seguinte mensagem aos organizadores do NWAV 45:

> Dear Panos and Alexandra, I would also like to say that I am very happy with the kind invitation: undoubtedly, this a real honor for me. I will do my best! Sincerely. Marta Scherre (Scherre, Maria Marta Pereira – e-mail enviado a Panos e Alexandra, com cópia para Naro, em 17 de novembro de 2014).

Mais tarde, soubemos que estávamos convidados para a primeira palestra de abertura do NWAV 46, honra verdadeiramente incomensurável. Naro sugeriu que, nessa palestra da dupla Naro e Scherre

(2016), apresentássemos uma visão geral dos princípios que governam a variação dos fenômenos de concordância verbal no português, sob o título: *General Principles Governing Variation in Brazilian Portuguese: the Scale of Prominence*.

Os organizadores do evento acrescentaram que Naro e eu estávamos também convidados para proferir uma palestra no evento *Simon Fraser University - Department of Linguistics Colloquium Series* de 2016. Para aproveitar a onda do NWAV 45, essa versão do Colóquio foi programada para dois de novembro de 2016, às vésperas do NWAV 45. No caso do Colóquio, Naro sugeriu que eu liderasse essa fala, configurando uma palestra da dupla Scherre e Naro (2016) e também sugeriu que fizéssemos uma revisão ampliada sobre as tendências sincrônicas da concordância de número na comunidade, na linha de Naro e Scherre (2013), sob o título *Revisiting Sociolinguistic Trends in the Brazilian Speech Community*.

Desta forma, com a forte sensação de que "Naro ainda está aqui" e ainda com muita história pra contar da dupla Naro e Scherre e Scherre e Naro, fecho a minha homenagem ao orientador, coautor e amigo Naro, neste evento PEUL 45 anos – 4º SEMINÁRIO DO PEUL – Em homenagem a Anthony Julius Naro e ao GRUPO PIONEIRO DA SOCIOLINGUÍSTICA NO BRASIL. Muito do que relatei aqui está além do Lattes. Os detalhes da intensa vida acadêmica do Naro ainda podem ser acessados em seu CV-Lattes.

Referências

GRYNER, Helena; MACEDO, Alzira Tavares de. La prononciation du S post-vocalique: deux processus de changement linguistic en portugais. *In*: SANKOFF, David; CEDERGREN, Henrietta. (org.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, Inc., 1981, pp. 135-140.

LEMLE, Miriam; NARO, Anthony J. *Competências básicas do português*. Rio de Janeiro, MOBRAL, Fundação FORD, 1977. Disponível em: https://cremeja.org/a7/wp-content/uploads/2019/09/3-1-competencias-basicas-do-portugues-1977.pdf. Acesso em: 09 out. 2025.

MOLLICA, Maria Cecília M. La copie dans les constructions relatives em Portugais. *In*: SANKOFF, David; CEDERGREN, Henrietta. (org.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, Inc., 1981, pp. 329-334.

NARO, Anthony J. Morphological constraints on subject deletion. *In*: SANKOFF, David; CEDERGREN, Henrietta. (org.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, Inc., 1981, pp. 351-357.

NARO, Anthony J. Memorial de Anthony Julius Naro. *In*: PACHECO, Cintia da Silva; ANDRADE, Carolina Queiroz. (org.). *Uma estrada da Sociolinguística Brasileira*: homenagem à trajetória acadêmica de Marta Scherre. Campinas, SP: Pontes Editores, 2024, p. 29.

NARO, Anthony Julius. *Currículo do Sistema de Currículos Lattes*. [*S. l.*], 2018. Disponível em: http://lattes.cnpq.br/4161994799982051. Acesso em: 22 jul. 2026.

NARO, Anthony J.; SCHERRE, Maria Marta Pereira. Variação e mudança linguística: fluxos e contrafluxos na comunidade de fala. *Cadernos de Estudos Linguísticos*, Campinas, v. 20, pp. 9-16, jan./jun. 1991. Disponível em:https://periodicos.sbu.unicamp.br/ojs/index.php/cel/article/view/8636853/4574 . Acesso em: 24 abr. 2026.

NARO, Anthony J.; SCHERRE, Maria Marta Pereira. Remodeling the age variable: Number concord in Brazilian Portuguese. *Language Variation and Change*, Cambridge University Press, v. 25, n. 1, pp. 1-15, Mar. 2013. DOI: https://doi.org/10.1017/S0954394512000269.

NARO, Anthony J.; SCHERRE, Maria Marta Pereira. *General principles governing variation in Brazilian Portuguese*: the scale of prominence. *In*: NEW WAYS OF ANALYZING VARIATION (NWAV), 45., 2016, Vancouver. Palestra plenária [...]. Vancouver: Simon Fraser University/University of Victoria, 2016. Disponível em: https://ling-webapps.fass.sfu.ca/nwav45/invited-speakers/. Acesso em: 24 abr. 2026.

OLIVEIRA e SILVA, Giselle Machline de. Perspective sociolinguistique de la forme você a Rio de Janeiro. *In*: SANKOFF, David; CEDERGREN, Henrietta. (org.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, Inc., 1981, pp. 481-487.

OMENA, Nelize. Pronom personnel de la troisieme personne: ses formes variantes en fonction accusative. *In*: SANKOFF, David; CEDERGREN, Henrietta. (org.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, Inc., 1981, pp. 343-350.

SANKOFF, David; CEDERGREN, Henrietta J. Preface. *In*: SANKOFF, David; CEDERGREN, Henrietta. (org.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, Inc., 1981, p. ix.

SCHERRE, Maria Marta Pereira. La variation de la règle d'áccord du nombre dans le syntagme nominal en Portugais. *In*: SANKOFF, David; CEDERGREN, Henrietta. (org.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, Inc., 1981, pp. 125-133.

SCHERRE, Maria Marta Pereira; NARO, Anthony J. Marking in discourse: Birds of a feather**.** *Language Variation and Change***,** Cambridge University Press, v. 3, n. 1, pp. 23-32, Mar. 1991. DOI: https://doi.org/10.1017/S0954394500000430.

SCHERRE, Maria Marta Pereira; NARO, Anthony J. *Revisiting Sociolinguistic Trends in the Brazilian Speech Community*. 2016. Palestra apresentada no Simon Fraser University Department of Linguistics Colloquium Series, Vancouver, Canadá. Disponível em: https://www.sfu.ca/linguistics/events/colloquia.html. Acesso em: 24 abr. 2026.

SCHERRE, Maria Marta Pereira; NARO, Anthony J.; YACOVENCO, Lilian Coutinho. Nós e a gente em quatro amostras do português brasileiro: revisitando a escala da saliência fônica. *Diadorim*, Rio de Janeiro, v. 20, n. esp., pp. 428-457, dez. 2018. Disponível em: https://revistas.ufrj.br/index.php/diadorim/article/view/23285. Acesso em: 3 mai. 2025.

SCHERRE, Maria Marta Pereira; RONCARATI, Cláudia. Programa de Estudos sobre o Uso da Língua (PEUL): origens e trajetórias. *In:* VOTRE, Sebastião; RONCARATI, Cláudia (org.) *Anthony Julius Naro e a Linguística no Brasil* – uma homenagem acadêmica. Rio de Janeiro: FAPERJ/7LETRAS, 2008, pp. 36-58.

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. Faculdade de Letras. PEUL: Programa de Estudos sobre o Uso da Língua. Rio de Janeiro, [1979]. Disponível em: https://peul.letras.ufrj.br. Acesso em: 5 maio 2026.

VOTRE, Sebastião J. Phonological and syntactic aspects of denasalization in spoken Brazilian Portuguese. *In*: SANKOFF, David; CEDERGREN, Henrietta. (org.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, Inc., 1981, pp. 97-103.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a75286

Entrevista com Ana Maria Carvalho (AC) por Xoán Carlos Lagares (XL)

Interview with Ana Maria Carvalho (AC) by Xoán Carlos Lagares (XL)

Ana Maria Carvalho[1]

Xoán Carlos Lagares[2]

Entrevistada

A professora Ana Maria Carvalho é uma destacada sociolinguística brasileira com um importante trabalho acadêmico sobre contato linguístico, especializada desde o seu doutoramento – realizado na Universidade de Califórnia, Berkeley – no português uruguaio, uma variedade linguística falada na região norte do Uruguai, ao longo da fronteira com o Brasil, nos departamentos de Rivera, Artigas, Cerro Largo, Tacuarembó e Salto.

Ela é professora do Departamento de Espanhol e Português da Universidade de Arizona e diretora do Programa de Aquisição e Ensino de Segunda Língua (SLAT), assim como membro do Centro de Estudos Latino-Americanos. Entre a sua vasta produção acadêmica destacamos a organização da coletânea *Português em contato* (Madri/Frankfurt: Ibero-americana/Vervuet, 2009), com capítulos que abordam situações de contato do português com outras línguas na África, em América, na Ásia e na Europa.

Nos seus estudos sobre o contato entre o português e o espanhol nas comunidades bilingues no norte do Uruguai, Ana M. Carvalho tem estudado questões relacionadas à variação e mudança linguística, a atitudes e ideologias linguísticas presentes na comunidade de falantes, explorando também a dimensão identitária dos usos, assim como políticas linguísticas implicadas em propostas de ensino bilíngue nesse território.

Junto com o editor Richie Machado, desenvolveu o projeto de elaborar um documentário sociolinguístico em que falantes de português uruguaio expusessem diretamente às câmeras a sua experiência com as línguas. Desse projeto surgiu o filme "Vozes das Margens" (2024), dirigido por ambos, um documentário em que o testemunho desses falantes se faz presente, entrelaçado com belas imagens do espaço natural e cultural onde vive e persiste o português do Uruguai.

Em junho de 2025, com motivo do IV Seminário do Programa de Estudos sobre o Uso da Língua (PEUL), que comemorava os seus 45 anos com uma homenagem a Anthony Julius Naro, o filme foi apresentado por Ana M. Carvalho na Faculdade de Letras da UFRJ. Naquela ocasião,

[1] Universidade do Arizona (UArizona), anac@arizona.edu, https://orcid.org/0000-0002-8577-5352.

[2] Universidade Federal Fluminense (UFF-CNPq), xlagares@id.uff.br, https://orcid.org/0000-0003-0974-2118.

tivemos oportunidade de manter uma conversa sobre os muitos aspectos sociolinguísticos levantados pelo documentário, que tentamos trazer agora para este formato de entrevista escrita. Aqui, a diretora de "Vozes das Margens" nos fala sobre as motivações e as características do gênero cinematográfico "documentário sociolinguístico", e os desafios e oportunidades que apresenta para os pesquisadores e comunidades de falantes, além de explicar aspectos fundamentais dessa situação concreta de contato linguístico: os efeitos da sobreposição do espanhol como língua do Estado, a questão da nomeação da variedade linguística, com suas implicações ideológicas, ou a relação com o português brasileiro e o conflito entre identidade local e ideal normativo.

XL: "Vozes das Margens" é um documentário sociolinguístico que se centra na experiência linguística dos falantes do português do Uruguai, contada por eles mesmos. O título do filme enfatiza essa perspectiva, que está presente na própria concepção do documentário, na sua realização e montagem. Como você chegou, a partir do seu trabalho de pesquisa na área, até essa concepção de documentário? E em decorrência dessa questão: na sua opinião, quais as características que deveria ter um filme que aborde questões de linguagem para merecer a denominação de "documentário sociolinguístico"?

AC: Em 1995 fui pela primeira vez ao norte do Uruguai, divisa com o Brasil, para coletar dados para a minha tese de doutorado. Nessa ocasião me hospedei com uma família local por 6 meses, fiz entrevistas sociolinguísticas com mais de 80 pessoas e observei horas e horas de interações espontâneas. Estes dados me possibilitaram não somente escrever minha tese e terminar meu doutorado, mas também material para análise por muitos anos subsequentes. Depois dessa primeira visita retornei várias vezes para coletar mais dados, e quanto mais o tempo passava, mais agradecida eu me sentia por ser recebida sempre de portas abertas pela comunidade, que generosamente compartilhava comigo seu *mate* e suas histórias. Estas experiências me ensinaram muito sobre sociolinguística, conhecimento que eu passava para meus estudantes, conferências e publicações.

Depois de 30 anos de coleta e análise de dados sobre o contato do português com o espanhol na região, me pareceu que já era tempo de seguir o princípio ético de reciprocidade com as comunidades que estudamos e desenvolver um projeto em conjunto com elas e para elas. A ideia de um documentário que retratasse a realidade sociolinguística segundo os próprios falantes de português uruguaio me pareceu uma oportunidade de conferir autoridade às vozes que muitas vezes são relegadas às margens. Estas vozes são periferizadas não somente pela estratificação sociolinguística e compartimentalização diglóssica que mantêm o português fora dos espaços públicos, mas também pelo fato de que nós, linguistas, muitas vezes excluímos as vozes e os corpos dos participantes dos nossos trabalhos científicos, ao transformá-los em dados despersonalizados, como mencionou Bill Labov (2020). *Vozes das Margens* é uma tentativa de subverter esta dinâmica.

Walt Wolfram define um documentário sociolinguístico como material que apresenta não apenas as línguas, mas também as experiências que os falantes têm com elas (Wolfram, 2010, p. 799). Ou seja, combina o registro das variedades minorizadas com narrativas sobre as dinâmicas sociais e experiências coletivas que atravessam as práticas linguísticas. Estas experiências, por sua vez, são também fundamentais para a difusão de conceitos sociolinguísticos entre o público em geral e importantíssimas para desmascarar as ideologias linguísticas prescritivas e puristas subjacentes ao preconceito linguístico presente no cotidiano desses falantes.

XL: O termo "Portunhol" é utilizado, frequentemente, para referir situações sociolinguísticas muito diferentes em que as línguas espanhola e portuguesa estão em contato. A situação retratada no filme não é apenas a de uma fronteira entre Estados, em que as pessoas de ambos os lados entram em interação frequente, dando lugar a formas híbridas das duas línguas nacionais. O processo político de sobreposição do espanhol, como língua de Estado, converteu o português falado no Uruguai numa língua em situação minoritária, em relação diglóssica com o idioma oficial. Vários depoimentos no filme relatam essa desigualdade no uso social das línguas; entre a casa e a escola, por exemplo. No seu parecer, há alguma estabilidade nesse bilinguismo diglóssico ou se trata de uma fase de um processo histórico de assimilação linguística?

AC: É verdade que a realidade no norte do Uruguai é bem diferente dos encontros entre o espanhol e o português que ocorrem nas outras fronteiras entre o Brasil e seus vizinhos, já que o uso do português na região era generalizado antes mesmo da chegada do espanhol (Barrios, 1996). E você explica muito bem as dinâmicas diglóssicas na região que estão por trás do uso do espanhol, idioma nacional, público e prestigioso, e do português uruguaio, idioma usado em casa, em família, em contextos físicos e sociais que implicam interações íntimas, solidárias. E é sem dúvida graças a essas dinâmicas que o português se manteve até hoje nessas comunidades desde a imposição do espanhol nos finais do século XIX, já que os contextos de uso de cada língua foram mantidos bem separados. Ao longo das últimas décadas, no entanto, sobretudo graças à urbanização, o espanhol foi gradualmente entrando nos lares, competindo com o português, que aos poucos vai saindo. Esta dinâmica irremediavelmente leva à substituição das variedades locais pelas variedades supralocais, um processo já completado em muitas famílias de classe média nas cidades da fronteira. Típico deste processo é o uso do português pelos pais em casa e o uso do espanhol pelos seus filhos, mecanismo que leva à substituição linguística em duas ou três gerações. Quando isto ocorre, o bilinguismo se desestabiliza e corre o risco de ser eliminado, levando ao monolinguismo em espanhol. No entanto, ainda hoje, nas comunidades rurais e nas classes trabalhadoras nas cidades, o português uruguaio se mantém graças ao forte valor identitário regional que ele representa para a região. Ambas as dinâmicas, de preservação do português e de mudança em direção ao espanhol, são representadas no documentário. Como estes cenários são fortemente dependentes de complexos fatores sociais, econômicos e geográficos, é muito difícil prever o futuro.

XL: Uma das questões que mais chamam a atenção nos depoimentos dos falantes entrevistados no filme é a quantidade de denominações empregadas para referir a fala local: entreverado, fronteiriço, portunhol... Também são usados diferentes glotônimos para denominar as línguas de Estado entre as quais eles se identificam: uruguaio, espanhol, espanhol puro, por um lado, e brasileiro, português, bayano, por outro. Quais identidades se representam nesses nomes e qual responde melhor à identidade local? Em que medida essa identidade se reconhece no termo portunhol?

AC: Nas comunidades, como retratado em *Vozes das Margens*, os falantes usam diversos termos para se referirem à variedade local. O mais intrigante é o termo 'bayano', que, segundo me contaram, vem de bageano, residente de Bagé, uma cidade no Rio Grande do Sul. No filme, notamos que, além de 'bayano', 'português', 'brasileiro' e 'fronterizo', 'entreverado' e 'portunhol' são também termos populares e remetem à percepção que os falantes têm de falarem um dialeto único e híbrido, uma mistura entre o português e o espanhol. Além disso, e como você notou, as comunidades se referem ao espanhol como 'uruguaio', o que remete às políticas linguísticas nacionalistas que foram impostas pelo governo central e até hoje são difundidas pelas escolas, que equivalem 'um idioma' a 'uma nação', e, portanto, 'ser uruguaio' a 'falar uruguaio' (e 'ser brasileiro' a falar 'brasileiro'). Interessantemente e fortalecendo esta hipótese, os falantes de português como primeira língua se referem ao espanhol na primeira pessoa, como "nossa língua" em situações em que claramente o espanhol foi aprendido na escola. O "nossa" é consequência da ideia de formação de nação fortemente construída através do idioma espanhol, como expliquei em 2014. O uso desses termos pela comunidade e suas relações com a identidade regional ainda carecem de uma análise sistemática.

Em termos de estudos científicos, a denominação das variedades do português faladas no norte do Uruguai tem sido material de controvérsia entre os linguistas. No início, Rona se referiu ao 'fronterizo' (Rona, 1965). Elizaincín mais tarde propôs *Dialectos de portugués de Uruguay*, ou *DPU* (Elizaincín 1978), termo este que foi posteriormente reproduzido em vários trabalhos. Hensey (1984) chegou a denominá-las *interlíngua* ou *interleto*. A noção de mistura de idiomas também é refletida no uso de *portunhol* (Sturza, 2022; Lipski, 2006). Da minha parte, as análises de meus dados e de outros colegas variacionistas me levaram a questionar a adequação desses termos e a sugerir *português uruguaio* (PU), designação hoje adotada pelo governo uruguaio.

Minha posição se baseia em trabalhos quantitativos que revelam padrões estruturais claramente portugueses, com a presença do espanhol limitada principalmente a empréstimos lexicais e alternância de códigos, processos comuns em comunidades bilíngues. Há pouquíssimas divergências sintáticas, fonológicas, ou morfológicas na estrutura do português uruguaio como consequência do contato com o espanhol, o que justificaria a ideia de uma terceira variedade híbrida, suficientemente diferente do que se fala em outras comunidades falantes de português. Portanto, considerando que o português uruguaio constitui um dialeto do português, ainda que com particularidades lexicais decorrentes do

contato linguístico, parece-me importante denominá-lo, nos trabalhos científicos, simplesmente como uma variedade regional do português. Desse modo, evitamos a difusão de estereótipos linguísticos associados à ideia de uma mistura aleatória de línguas, conotação que alguns dos termos anteriores carregam.

XL: Em realidades como a descrita no filme costuma haver uma tensão entre a preservação da identidade local e o ideal linguístico representado por uma língua padronizada, com a qual os falantes se identifiquem de alguma maneira. Quais são as posturas, dentro dessa comunidade, em relação ao papel que o português brasileiro poderia exercer numa possível política linguística, orientada, por exemplo, ao ensino?

AC: Os falantes de português uruguaio são extremamente inseguros e convictos de que falam uma variedade inferior àquelas faladas tanto por monolingues de espanhol no sul do Uruguai como por monolingues de português no Brasil. Esta atitude negativa em relação às variedades bilíngues é fruto de ideologias linguísticas que prezam variedades nacionais, prestigiosas, urbanas e 'puras'. A insegurança linguística pode até levar alguns grupos, como os jovens da classe média, e em alguns estilos (mais formais) a assimilarem algumas características linguísticas supralocais, usando o português brasileiro como modelo (Carvalho, 2004) e distanciando-se do dialeto local. No entanto, é importante mencionar que o português uruguaio se manteve como idioma local graças ao forte vínculo com a identidade local, do qual o português brasileiro é totalmente desprovido.

Em relação ao ensino, em 2002 fui enviada à fronteira por um projeto que pretendia implementar aulas de português na região. Na ocasião entrevistei pais e professores de escolas primárias em Artigas e Rivera sobre o que achavam da ideia. Constatei um consenso entre todos sobre a importância de ensinar português para que aprendessem "o português verdadeiro". Posteriormente, aulas de português como língua estrangeira foram incluídas no currículo de 74 escolas fronteiriças. Em 2018 tive a oportunidade de observar algumas dessas aulas e constatei que, infelizmente, o ensino de português não busca legitimizar a variedade local especificamente, mas sim substituí-la pelo dialeto padrão brasileiro. Esta agenda também é a nacional, como ilustrado nos dizeres da ministra de Educação e Cultura do Uruguai, em uma entrevista para uma revista, quando afirmou que "os uruguaios sabem portuñol, mas têm que aprender português"[3]. Na minha opinião, a ideia de substituir o português uruguaio pelo padrão brasileiro nas comunidades fronteiriças é extremamente prejudicial, pois legitimiza o português brasileiro em território uruguaio, reduzindo o valor de um importante patrimônio cultural fortemente ligado à terra, à ancestralidade e à herança cultural. Em vez de diminuir, exacerba a desigualdade linguística e a periferização do português local.

[3] María Julia Muñoz: "Los uruguayos hablan portuñol, queremos que hablen portugués". Montevidéu: Portal, 2017

XL: Em relação à pergunta anterior, existe alguma reivindicação política de reconhecimento oficial e, portanto, de "elaboração" de um "português uruguaio", como língua dessa comunidade? Do seu ponto de vista, o cultivo literário, por exemplo, poderia ser considerado como um passo nessa direção?

AC: A presença do português uruguaio nas mídias sociais, a meu ver, constitui uma forma de subversão e resistência diante da ausência do português nos espaços públicos, contribuindo para a construção de sentidos de localidade e de legitimidade linguística. O mesmo ocorre com o uso de português uruguaio na literatura, como vemos no documentário através da poesia de David Benevidez. Não sei se essas iniciativas contribuirão para a sua elaboração, e não sei se é necessária tal elaboração, mas acredito que contribuem para sua legitimização.

Em relação a uma política de reconhecimento oficial, a ANEP (Administración Nacional de Educación Pública) reconhece o português como língua regional [4], enquanto que o espanhol é a língua nacional. Em relação a políticas locais, um movimento chamado *Jodido Bushinshe* visou postular o idioma local como patrimônio cultural imaterial através da criação e difusão de arte em português uruguaio como tentativa de reconhecimento e proteção. Albertoni (2018) explica que essa iniciativa, organizada por um grupo de intelectuais e artistas de Rivera, foi marcada por contradições entre os objetivos de seus ativistas e as perspectivas dos habitantes que utilizam o português em seu cotidiano, bem como por tensões decorrentes do conflito entre ideologias de autenticidade e de mercantilização.

XL: Começamos esta entrevista falando do gênero "documentário sociolinguístico". Gostaria de encerrá-la perguntando sobre os aprendizados que você teve, como linguista, ao enfrentar o desafio de contar essa realidade linguística fora dos muros da academia, para um público não especialista. Muito se fala, atualmente, nas universidades brasileiras, de transferência de conhecimento e de popularização da linguística. Os conhecimentos sociolinguísticos foram essenciais para planejar e realizar um filme como esse, capaz de fazer um retrato fiel da situação, mas, por outro lado, podemos dizer que popularizar esse conhecimento significou também renunciar à tentação de oferecer explicações baseadas no saber acadêmico? Como foi lidar com esse aparente paradoxo?

AC: Sim. Um importante objetivo de *Vozes das Margens* era difundir conceitos da sociolinguística (como variação linguística, atitudes e ideologias, estratificação sociolinguística, preconceito, diglossia, transmissão e ruptura de línguas minorizadas, etc.). A minha instrução para os protagonistas, em *off,* era simplesmente que narrassem suas experiências com as línguas, e foi o que eles fizeram. As narrativas não me trouxeram nenhuma surpresa. Realmente incluíram experiências que refletiam as dinâmicas sociolinguísticas exatamente como já havíamos identificado nos estudos anteriores. Outro objetivo era fazer um documentário sem a presença de especialistas, para que as vozes fossem somente dos falantes, das margens. Isso foi desafiador porque a maioria das pessoas pode descrever sua realidade linguística, mas não tem o conhecimento para analisá-la, pois aprendemos sobre estas questões quando

[4] https://www.anep.edu.uy/codicen/politicas-linguisticas/portugues?utm_source=chatgpt.com

fazemos cursos de linguística. Isto implica que o trabalho de edição e montagem é muito importante. Walt Wolfram, um sociolinguista americano e autor de vários documentários sociolinguísticos, afirma que as decisões de edição são sumamente importantes para lidar com "os objetivos que competem entre si, entre a voz das comunidades e o conhecimento linguístico do especialista, para que estereótipos negativos e mitos linguísticos não sejam difundidos." (2010, p. 799). Portanto, as decisões relativas à seleção e à organização das narrativas exigiram cuidadosa reflexão.

Referências

ALBERTONI, P. Mercantilización y autenticidad en la frontera uruguayo-brasileña: el portuñol en el siglo XXI. *Trabalhos de Linguística Aplicada*, v. 60, n. 2, pp. 410-424, 2018.

BARRIOS, G. Planificación lingüística e integración regional: el Uruguay y la zona de frontera. *In*: TRINDADE, A. Menine; BEHARES, L. (org.). *Fronteras, educación, integración*. Santa María: Pallotti, 1996.

CARVALHO, A. M. "I speak like the guys on TV": Palatalization and the urbanization of Uruguayan Portuguese. *Language Variation and Change*, v. 16, n. 2, pp. 127-151, 2004.

CARVALHO, A. M. Sociolinguistic continuities in language contact situations: The case of Portuguese in contact with Spanish along the Uruguayan-Brazilian border. *In:* AMARAL, P.; CARVALHO, A. M. (eds.), *Portuguese/Spanish interfaces*. Amsterdam: John Benjamins, 2014, pp. 263-294.

ELIZAINCÍN, A. *Algunas precisiones sobre los dialectos portugueses en Uruguay*. Montevidéu: Universidad de la República, 1978.

HENSEY, F. Uruguayan Portuguese as an interlanguage and interlect. *In*: *Language in the Americas*. Nova York: Cornell University Press, 1984, pp. 148-160.

LABOV, W. *Justice as a linguistic matter* [Video]. ABRALIN Speakers Series. 19 de maio de 2020.

LIPSKI, J. Too close for comfort? The genesis of "portuñol/portunhol". *In*: FACE, T.; KLEE, C. (ed.). *Selected proceedings of the 8th Hispanic Linguistics Symposium*. Somerville: Cascadilla Press, 2006. p. 1-22.

RONA, P. *El dialecto fronterizo del norte del Uruguay*. Montevidéu: Adolfo Linardi, 1965.

STURZA, E. Portunhol: Língua, história e política. *Gragoatá*, v. 24, n. 48, pp. 95-116, 2022.

WOLFRAM, W. Collaborative issues in language variation documentaries. *Language and Linguistics Compass*, v. 4, n. 9, pp. 793-818, 2010.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a71952

As gramáticas dos gramáticos: a sintaxe dos clíticos pronominais em grammaticas históricas do português

The grammars of grammarians: the syntax of pronominal clitics in historical grammars of Portuguese

Saulo Rogério Pacheco Rocha[1]

Marco Antonio Rocha Martins[2]

RESUMO

Este artigo analisa a mudança na colocação dos pronomes clíticos no português europeu, em gramáticas históricas entre os séculos XV e XIX. Para tanto, foi constituído, transcrito e editado, um *corpus* inédito de seis gramáticas históricas, utilizando um modelo de OCR especializado, o Early Portuguese Printing (EPP), treinado na plataforma Transkribus. A análise quantitativa realizada, fundamentada no modelo da Competição de Gramáticas de Anthony Kroch (1989, 2021 [2001]), mapeia a mudança na sintaxe dos clíticos pronominais, de um sistema majoritariamente proclítico (indício de produção da gramática do Português Clássico, a gramática conservadora) para um de predominância enclítica (indício do Português Europeu, a gramática inovadora em implementação no *corpus*). Os resultados mostram que a produção sintática dos gramáticos acompanha a curva de mudança já descrita para essa mudança noutros gêneros, como o estudo de Galves *et al.* (2005), e que a produção de obras normativas, pelo menos até o século XIX para os europeus, era escrita com base em suas gramáticas adquiridas. Além disso, as relações sócio-históricas levantadas como hipóteses para a gênese da gramática do PE, no século XVIII, como a dissolução da esfera cultural da Língua Barroca (Paixão de Sousa, 2004), se relacionam de forma consistente com os contextos sócio-históricos e biográficos de cada autor. O artigo, dessa forma, valida o uso de textos normativos como objeto para a análise da sintaxe diacrônica e contribui com um novo *corpus* digital para a área.

PALAVRAS-CHAVE: Sintaxe Diacrônica. Competição de Gramáticas. Clíticos em português. Gramáticas Históricas. Humanidades Digitais.

ABSTRACT

This paper analyses the change in the placement of clitic pronouns in European Portuguese in historical grammars between the 15th and 19th centuries. For this purpose, an unpublished *corpus* of six historical grammars was compiled, transcribed, and edited using a specialized OCR model, Early Portuguese Printing (EPP), trained on the Transkribus platform. The quantitative analysis, based on Grammar Competition model (Kroch, 1989, 2021 [2001]), maps the change in the syntax of pronominal clitics from a predominantly proclitic system (indicative of Classical Portuguese grammar, the conservative grammar) to a predominantly enclitic one (indicative of European Portuguese, the innovative grammar implemented in the *corpus*). The results reveal that the syntactic production of grammarians follows the curve of change already described for this shift in other genres, such as the study by Galves *et al.* (2005), and that the production of normative works, at least until the 19th century for Europeans, was written based on their acquired grammars. In addition, the socio-historical relationships raised as hypotheses for the genesis of PE grammar in the 18th century, such as the dissolution of the cultural sphere of the Baroque Language (Paixão de Sousa, 2004), are consistently related to

[1] Universidade Federal de Santa Catarina (UFSC), saulo.r@posgrad.ufsc.br, https://orcid.org/0000-0003-3715-6706.

[2] Universidade Federal de Santa Catarina (UFSC)/CNPq, marcomartins.ufsc@gmail.com, https://orcid.org/0000-0002-3999-3893.

Recebido em 26/12/2025 - Aceito em 10/06/2026

the socio-historical and biographical contexts of each author. The paper validates the use of normative texts as objects for analyzing diachronic syntax and contributes a new digital *corpus* to the field.
KEYWORDS: Diachronic Syntax. Competition of Grammars. Clitic. Historical Grammars. Digital Humanities.

1. Introdução

O estudo da mudança linguística, especialmente no domínio da sintaxe, tem se beneficiado imensamente da construção de *corpora* digitais que permitem a análise quantitativa de fenômenos ao longo do tempo. A trajetória da colocação pronominal clítica na história do português, um dos tópicos mais estudados da sintaxe diacrônica, ilustra uma complexa transição de gramáticas ao longo do tempo. Embora a pesquisa na área tenha se consolidado a partir de *corpora* de gêneros textuais variados, como textos literários, cartas e peças de teatro[3], o gênero gramáticas e obras normativas, talvez por sua natureza prescritiva e monitorada, permanece como um campo de investigação menos explorado.

Este artigo parte da hipótese de que, apesar de seu caráter normativo, gramáticas e textos normativos escritos por gramáticos não são imunes à variação e à mudança que operam nas línguas. Pelo contrário, a prosa desses autores constitui um campo de observação privilegiado para observar a "competição de gramáticas" (aqui empregado nos termos de Kroch, 1989), permitindo investigar a complexa relação entre a norma prescrita, a gramática internalizada do autor e o uso linguístico de sua época. Nesse sentido, este artigo busca responder as segu questões como: a sintaxe de gramáticos que escreveram gramáticas históricas reflete os padrões e a mudança em relação à sintaxe dos clíticos pronominais na história do português? De que forma o contexto sócio-histórico de cada autor se manifesta em seus padrões sintáticos?

Dessa forma, portanto, o objetivo central aqui é apresentar, com base em Pacheco Rocha (2025), a curva de mudança na colocação dos clíticos em um *corpus* inédito e especializado de seis gramáticas portuguesas, publicadas entre 1536 e 1864, e comparar esse quadro àquele descrito por outros autores para os padrões do português escrito nesse período. A metodologia combina a criação de um *corpus* digital de alta fidelidade, por meio da plataforma Transkribus e de um modelo de OCR próprio, com a análise quantitativa dos dados à luz do referencial teórico da Competição de Gramáticas (Kroch, 1989, 2021 [2001]). O artigo está estruturado em seções que apresentam o embasamento teórico, a composição do *corpus*, o perfil de cada autor, e, finalmente, a análise detalhada dos dados sintáticos.

2. A teoria e o *Corpus*

Nosso conceito de gramática vem da teoria gerativa chomskyana. Gramática é um conjunto de regras que definem o conhecimento linguístico de um indivíduo que sabe e fala uma língua natural, conforme Chomsky (1986), ou de uma geração de indivíduos, como define Mendivel-Giró (2015).

[3] Ver, por exemplo, os corpora do Corpus Histórico do Português Tycho Brahe https://www.tycho.iel.unicamp.br/corpus/ e do Projeto para a História do Português Brasileiro https://sites.google.com/site/corporaphpb/.

A análise que apresentamos se assenta em um arcabouço de pressupostos da teoria da gramática, que concebe a mudança não como um processo aleatório, mas como uma substituição de sistemas regrados, constrangidos, por princípios de uma Gramática Universal. Adota a proposta de *Gramática Histórica* de Mendívil-Giró (2015) como objeto de estudo; ou, dito de outro modo, um "retrato" da gramática de uma comunidade geracional – ou Língua-I de indivíduos de uma mesma geração – num tempo e espaço definidos. Nessa perspectiva, indivíduos e suas gramáticas são vistos como unidades fundamentais da mudança que representam os pontos que delineiam trajetórias da mudança diacrônica ao longo de extensos períodos históricos, e os dados observáveis e empiricamente verificáveis são interpretados como estágios dessas curvas históricas, quanto mais dados mais claramente se delineia a curva de tendência de mudança. Como consequência disso, faremos neste artigo a distinção entre a **gramática** dos gramáticos, que neste trabalho fará referência ao conhecimento linguístico geracional dos autores-gramáticos, as gramáticas históricas em competição; e a **grammatica** dos gramáticos, que se refere às obras por eles escritas, suas *grammaticas*.

O embasamento teórico deste artigo se ancora também no modelo da Competição de Gramáticas, proposto por Anthony Kroch (1989, 2021 [2001]). Os pressupostos teóricos adotados postulam que a mudança linguística ocorre quando uma gramática inovadora se desenha/muda numa comunidade de falantes e passa a coexistir e competir com uma gramática conservadora, a estabelecida. Com o tempo, a nova gramática é gradualmente adotada pelas novas gerações, e a frequência de uso de fenômenos superficiais que a instanciam pode ser medida estatisticamente, geralmente descrevendo uma curva em formato de "S" (*f-curve*). A aplicação deste modelo à história do português permite analisar, por exemplo, a variação diacrônica na colocação dos pronomes clíticos, que é uma hipótese já bastante investigada. A competição entre uma gramática mais antiga, o Português Clássico (PCl), que favorecia a próclise em contextos de variação diacrônica, e uma nova gramática, o Português Europeu (PE), que passou a preferir a ênclise, já que a natureza dos híbrida dos clíticos, isto é, não tão morfologicamente restrita e nem tão sintaticamente livre, indica transformações substanciais na *gramática* histórica reproduzida na população em que os autores dos vestígios históricos estão inseridos. Para um período em específico, o avanço da forma inovadora é, então, observado na frequência crescente das ênclises nos textos de cada período.

2.1. Sobre gramáticos, gramáticas e grammaticas – o *corpus*

Pacheco Rocha (2025) organizou um *corpus* inédito de "grammaticas" do português: obras de cunho metalinguístico redigidas por autoridades da língua portuguesa em seu tempo[4]. A escolha por este gênero textual, que diverge de abordagens convencionais focadas em dados menos monitorados, parte da hipótese de que mesmo textos prescritivos não são imunes a processos de mudança linguística. Pelo contrário, servem como um campo de observação privilegiado para a "competição de gramáticas"

[4] O corpus está integralmente disponível em: https://doi.org/10.5281/zenodo.18354169 para informações metodológicas a respeito de sua constituição, veja Pacheco Rocha (2026).

que opera em todo o sistema linguístico, ainda que de forma velada. Além disso, esse recorte permite investigar a complexa relação entre a norma prescrita e o uso linguístico atestado em *corpora* mais amplos e variados (como cartas, peças de teatro etc.). A análise das grammáticas em contraste com estudos já realizados em *corpora* de outras naturezas permite formular questões centrais para a sintaxe diacrônica: em que medida a produção normativa de uma época reflete as transformações em curso na língua? Os gramáticos atuavam como agentes de conservação, de inovação, ou simplesmente registravam tardiamente as mudanças já consolidadas pela comunidade de falantes? De que forma a gramática do PE substituiu a do PCl nas obras normativas e metalinguísticas?

O *corpus* desta pesquisa é composto por um conjunto de seis obras de cunho metalinguístico, aqui chamadas de *grammaticas*, publicadas em Portugal entre os séculos XVI e XIX. As cópias digitais foram obtidas do acervo da Biblioteca Nacional de Portugal e seu processo de transcrição e edição foi realizado na plataforma Transkribus, com um modelo de OCR treinado especificamente para este estudo, o *Early Portuguese Printing*[5] (EPP). Os autores e respectivos anos de nascimento e as obras estão listados no que segue:

1. **João de Barros** (1496-1570): *Grammatica da Lingua Portuguesa* (1540).
2. **Fernão de Oliveira** (1507-1581): *Grammatica da Lingoagem Portuguesa* (1536).
3. **Duarte Nunes de Leão** (c.1530-1608): *Origem da Lingoa Portuguesa* (1606).
4. **Luís Caetano de Lima** (1671-1757): *Orthographia da Lingua Portugueza* (1736).
5. **Bernardo de Lima e Melo Bacellar** (1736-?): *Grammatica Philosophica e Orthographia Racional da Lingua Portugueza* (1783).
6. **Francisco Júlio Caldas Aulete** (1823-1878): *Grammatica Nacional* (1864).

O fenômeno linguístico selecionado para a análise quantitativa neste *corpus* é, como dito, a colocação dos pronomes clíticos em contextos de variação diacrônica e que serão especificados na seção 3.2 que segue[6]. E, partindo do pressuposto de que os dados sintáticos só podem ser plenamente interpretados à luz do contexto sócio-histórico de seus produtores, a seção seguinte apresenta um perfil biográfico de cada gramático, de forma a demonstrar e correlacionar a forma como suas trajetórias individuais e dinâmicas sociopolíticas e históricas que moldaram suas obras e o padrão gramatical da língua portuguesa.

---

[5] O EPP foi tornado público em outubro de 2024 e está disponível para todos os usuários na plataforma Transkribus no link: transkribus.org/model/early-portuguese-printing.

[6] Neste artigo, o termo "contextos de variação" se refere aos ambientes sintáticos nos quais se observa historicamente no português a alternância e a competição entre a próclise e a ênclise. Conforme será detalhado na seção metodológica, adota-se aqui uma síntese entre os Contextos de Variação I e II descritos por Galves, Britto e Paixão de Sousa (2005) e a classe de sentenças "não dependentes" de A. M. Martins (1994), abarcando principalmente sentenças matrizes com elementos pré-verbais (sujeitos, advérbios ou sintagmas preposicionais) e segundas coordenadas.

João de Barros, o Nobre Suprassumo da Aristocracia

João de Barros (1496-1570), nascido em Vila Verde, foi filho, provavelmente bastardo, de Lopo de Barros, um nobre da província de Gestosa/Portugal (Panegassi, 2013). Criado no paço real, recebeu uma sólida formação em latim, grego e humanidades, o que garantiu a ele proximidade com o futuro rei D. João III. Na vida adulta, ocupou cargos importantes na administração imperial, como por exemplo, o de tesoureiro da Casa da Índia, a partir de 1533. Embora mais conhecido por sua obra historiográfica, *As Décadas da Ásia*, Barros publicou trabalhos linguísticos fundamentais, como a Cartinha (1539), a grammatica da Lingua Portuguesa (1540) e o Diálogo em louvor da nossa linguagem (1540). A produção de Barros está intrinsecamente ligada à intensificação da expansão ultramarina portuguesa no século XVI. Conforme descrito por Russell-Wood (2014), a crescente necessidade de comunicação e catequização nos novos territórios estimulou a produção de materiais didáticos impressos. Um exemplo precoce dessa demanda são as "cartinhas" de instrução religiosa enviadas ao Congo já em 1504 (Mattos e Silva, 2001), um tipo de instrumento pedagógico que seria aprimorado por gramáticos como Barros nas décadas seguintes. Nesse contexto, sua grammatica reflete o projeto humanista de dignificar o português, o transformando em uma língua de cultura e em um instrumento para a afirmação e administração do império.

Fernão de Oliveira, o Frade e Capitão do Mar

Diferentemente de João de Barros, as informações biográficas de Fernão de Oliveira (c. 1507-1582) são fragmentadas, reconstituídas a partir de fontes como processos inquisitoriais e cartas, muitas vezes com imprecisões (Santos, 2022). Nascido na Gestosa, mesma província de Barros, e filho de um "juiz de órfãos", Oliveira se muda para Évora ainda durante sua juventude, período sobre o qual relata, em sua grammatica, um episódio raro de consciência sociolinguística, no qual descreve o preconceito que sofria em Évora por sua pronúncia beirã da terminação verbal, que ele próprio associava à de João de Barros, tido como modelo de prestígio. Em Évora, Oliveira ingressou na Ordem Dominicana, mas em determinado momento abandonou a vida em convento, por volta de 1532, e passou a viver na Espanha, onde aprofundou seus estudos humanistas e provavelemnte começou a escrever sua grammatica (Buescu, 1978, p. 49-50). De volta a Lisboa, publicou sua *grammatica da Lingoagem Portugueza* em 1536, provavelmente escrita durante seu tempo na Espanha (*cf.* Buescu, 1983). Sua vida posterior foi marcada pela instabilidade: viajou, comandou uma nau francesa e foi preso duas vezes pela Inquisição portuguesa por supostas ligações com o protestantismo. Apesar das perseguições, foi nomeado professor na Universidade de Coimbra entre 1555 e 1557 e legou obras importantes como a *Arte da Guerra no Mar* e a primeira *História de Portugal*. A trajetória intelectual de Oliveira foi moldada pelo contexto conturbado do século XVI português. A sua obra emerge em um período de consolidação do império, mas também de forte rivalidade cultural com a Espanha, que fomentava um sentimento de "unidade linguística ibérica" sob a hegemonia do castelhano (*cf.* Paixão de Souza, 2004). Nesse cenário, a gramática de Oliveira pode ser lida não apenas como um

trabalho erudito, mas como um ato político, uma tentativa de dignificar a língua portuguesa e afirmar sua autonomia frente ao latim e ao crescente prestígio do castelhano.

Nunes de Leão, o Historiador Hispânico

Duarte Nunes de Leão (c. 1530-1608) foi um proeminente intelectual e jurista formado pela Universidade de Coimbra. De ascendência judaica, sendo filho de um médico "hebreu" (Buescu, 1983), sua produção intelectual foi vasta, mas também alvo de notória vigilância editorial pela Igreja, que exigia autorizações das mesas censórias para suas publicações. Suas principais obras linguísticas são a *Orthographia da Lingoa Portuguesa* (1576) e, escrita no final de sua vida, a *Origem da Lingoa Portuguesa* (1606), um dos primeiros estudos sobre a história da língua como fenômeno histórico. Sua trajetória se desenrola durante a ascensão do poder filipino e a subsequente União Ibérica (1580-1640), à qual ele era publicamente favorável. Seu posicionamento pode ser interpretado à luz de sua condição de "cristão-novo" e de uma possível origem familiar espanhola. As dedicatórias de suas obras são evidências eloquentes: a Orthographia (1576), publicada às vésperas da união, é dedicada a "Lourenço da Sylva" um nobre português, enquanto a Origem (1606), já sob domínio espanhol, é oferecida diretamente ao rei Filipe II. Nesse contexto, é notável que, no auge da união política, Leão se dedique a celebrar a origem da língua portuguesa, tratando os feitos "portugueses" e "espanhóis" como parte de uma Hispânia unificada. Conforme expõe Paixão de Sousa (2004), a influência do castelhano nesse período se aprofundou, resultando numa "relação assimétrica" que afetou os padrões sintáticos do português, especialmente em autores politicamente alinhados e, principalmente, bilíngues luso-espanhol como Leão.

Luís Caetano de Lima, o Diplomata e aqui o Último dos Nobres

Luís Caetano de Lima (1671-1757) foi uma figura central para compreender o século XVIII português, e sua biografia é bem documentada por fontes contemporâneas. Nascido em Lisboa em uma família nobre, recebeu uma formação erudita que incluía seis idiomas e o preparou para uma proeminente carreira como diplomata de D. Pedro II (de Portugal) e D. João V em capitais como Paris, Londres e Roma (*cf.* Assunção *et al.*, 2022). A trajetória de Lima se confunde com a ascensão de Portugal como um novo poder econômico global, independente da Espanha desde 1640 e impulsionado pelo ouro do Brasil, encontrado nesse período histórico. Lima atuou como um agente cultural deste novo estado, sendo enviado a Paris para adquirir livros para a biblioteca real e ocupando cargos como Mestre dos Infantes. Sua produção intelectual reflete essa orientação cosmopolita, com a publicação de gramáticas de francês e italiano, que objetivavam o aprendizado dessas línguas por portugueses (Assunção *et al.*, 2022, p. 32). Sua principal obra para a língua portuguesa, a *Orthographia da Lingua Portugueza* (1736), é um marco deste período de intensa atividade intelectual. A obra de Lima surge em um momento de reconfiguração cultural. Após a independência de 1640, a forte influência espanhola do universo cultural "barroco" sobre as elites letradas portuguesas começou a declinar, dando lugar à hegemonia cultural francesa (Paixão de Sousa, 2004; Galves, 2012). Essas duas influências parecem

ter tido efeitos distintos: enquanto a proximidade com o castelhano favoreceu uma convergência linguística pan-ibérica, a ascensão do prestígio francês parece ter estimulado um movimento de diferenciação, isolando o português. A Orthographia de 1736 é, portanto, um reflexo dessa transição, preparando o terreno para as novas tradições gramaticais e as ideias iluministas do século XVIII.

Bernardo Bacellar, o Racionalista e Primeiro dos Gramáticos Modernos

Bernardo de Lima e Melo Bacellar (1736-1831), nascido em Chaves, iniciou sua trajetória na Ordem Franciscana sob o nome de "Frei Bernardo de Jesus Maria". Sua carreira foi profundamente marcada pela tutela de Frei Cenáculo, seu superior e uma das figuras centrais na reforma intelectual de Portugal. Foi por meio dessa ligação que Bacellar teve acesso às ideias iluministas que, acredita-se (*cf*. Melo, 2020), motivaram sua drástica decisão de abandonar a vida clerical, inclusive seu nome, e, posteriormente, se mudar para Paris. Na capital francesa, sua imersão nos círculos eruditos se consolidou com seu trabalho na Biblioteca Real, de onde manteve uma colaboração intelectual contínua com seu protetor, enviando a ele cópias traduzidas de manuscritos que tinha acesso (Torres, 1996, p. 461). A grammatica Philosophica (1783) é a culminação desse percurso intelectual. A publicação da obra, viabilizada pela posição de Frei Cenáculo na Real Mesa Censória, marca a introdução da tradição racionalista de Port-Royal na gramaticografia portuguesa. O acesso privilegiado de Bacellar à vanguarda do pensamento europeu é ilustrado pela citação, em seu livro de 1783, da *Encyclopédie méthodique* publicada apenas um ano antes, em 1782 (Torres, 1996, p. 461). Esse desenvolvimento individual só foi possível dentro do contexto mais amplo das reformas pombalinas, que embora perseguissem as ideias perigosas dos estrangeirados (Saraiva, 2012, p. 67), permitiam àquelas úteis ao regime. Nesse cenário, a gramática de Bacellar representa uma ruptura com a tradição de herança ibérica, refletindo a nova orientação cultural de Portugal no século XVIII e inaugurando a competição de normas que viria a redefinir o debate sobre a língua portuguesa.

Júlio Caldas Aulete, o Professor de Duas Nações

Francisco Júlio Caldas Aulete (1823-1878), nascido em Lisboa, foi filho de um funcionário público e se tornou uma figura proeminente como professor e autor de diversos livros educacionais. Embora sua figura seja frequentemente ofuscada pelo sucesso póstumo de seu dicionário, Aulete atuou em importantes instituições portuguesas e, conforme os obituários da época, dedicou sua vida ao ensino da língua portuguesa, vindo a falecer de tifo em maio de 1878 (Polachini, 2022). Embora seja difícil mensurar se o próprio Aulete tinha a dimensão do impacto transnacional de seu projeto, as evidências de sua vasta circulação são inequívocas. A *grammatica Nacional* não apenas foi adotada no ensino público português, mas também no prestigioso Colégio Pedro II, no Rio de Janeiro. A obra era comercializada na capital do Império do Brasil, por vezes com o título alterado para "grammatica Portugueza", uma adaptação que gerou inclusive um comentário no jornal "O Lobisomem", da época, que ironizava o caráter nacional da gramática oficial portuguesa adotada no Brasil. O impacto da produção de Aulete no Brasil foi tão significativo que sobreviveu ao autor.

Conforme descreve Verdelho (2002, p. 43), o Dicionário Aulete, publicado postumamente, encontrou no Brasil um público consumidor fundamental, de suas 8 edições, as 5 maiores e mais recentes foram feitas no Brasil. Essa apropriação pelo mercado editorial brasileiro se consolidou no século XX, após 1958, quando os direitos da obra foram adquiridos pela editora Delta (*cf.* Nunes, 2002, p. 111), e perdura até hoje, com o dicionário sendo atualizado pela também brasileira Lexikon Editora Digital[7]. A atuação de Aulete deve ser compreendida no contexto do século XIX português, um período de intensa instabilidade, mas também de um esforço contínuo de modernização do Estado e da educação após as crises das Guerras Napoleônicas e da independência do Brasil. As sucessivas reformas educacionais buscavam criar um cidadão letrado e unificado. É nesse cenário que o "intento civilizatório" de Aulete ganha sentido: suas obras, especialmente a *Grammatica Nacional* (Aulete, 1864), não eram apenas ferramentas de ensino, mas instrumentos pedagogicamente muito poderosos, diferentemente das outras *grammaticas* tratadas neste trabalho, a grammatica de Aulete era de fato voltada à atuação em sala de aula, e essa qualidade teórica e prática foram méritos que que ajudaram a cristalizar uma norma linguística moderna, padronizada e europeia, fundamental para o projeto de nação do liberalismo português e do conservadorismo brasileiro.

3. As Metodologias

3.1. A Transcrição das grammaticas

Optamos pela transcrição direta dos fac-símiles das primeiras edições de cada gramática, em vez de basear-nos em edições digitais pré-existentes. Essa decisão foi motivada, primeiramente, pela necessidade de garantir controle total sobre o processo de codificação e normalização dos dados, de forma a assegurar uma estrutura uniforme em todo o *corpus*, o que é fundamental para a confiabilidade da análise computacional. Em segundo lugar, a decisão foi uma necessidade prática, visto que poucas transcrições confiáveis estavam disponíveis online, e as existentes se encontravam em formatos, como PDF de imagem ou, quando muito, com transcrições OCR inadequadas para o processamento linguístico almejado.

Para as transcrições, portanto, empregamos a plataforma Transkribus, administrada pela cooperativa Read-Coop. A escolha se deu por sua capacidade de criar modelos de reconhecimento de caracteres especializados (*fine-tuning* dos modelos disponíveis*)* em diferentes qualidades gráficas de escritas históricas. A necessidade de um modelo próprio, o *Early Portuguese Printing* (EPP), foi confirmada após testes demonstrarem que os modelos disponíveis eram insuficientes para a especificidade do *corpus*. A estratégia de treinamento adotada foi a de processar as obras em ordem cronológica, das mais arcaicas (século XVI) às mais modernas (século XIX). A hipótese subjacente a essa forma de "aprendizado anti-curricular" (Bengio, 2009) é a de que, ao despender mais tempo e trabalho no início do treinamento, em fornecer *ground truths* de qualidade para o modelo a aprender primeiro a maior variabilidade dos impressos antigos, a IA desenvolveria uma base de reconhecimento

[7] Disponível em: aulete.com.br.

mais robusta, que seria então refinada com os textos mais padronizados. Enfim, para garantir o rigor filológico da transcrição, especialmente em passagens ambíguas das obras de Barros (1540) e Oliveira (1536), foram consultadas como referência as edições críticas de Buescu (1975), Machado (1957) e, alternadamente, a obra de Oliveira transcrita no *Corpus Histórico do Português Tycho Brahe* (Veronesi; Galves, 2024).

Os modelos pré-treinados para o português, ao que tudo indica, também possuíam treinamento em idiomas próximos, como o castelhano, e por isso se mostraram inviáveis devido a particularidades da língua portuguesa, cuja variabilidade histórica no uso de diacríticos (timbre, tonicidade, crase, til, etc.) ou grafemas (como a cedilha subscrita não apenas no 'c', mas também no 'ę') constituiu a necessidade duma distribuição de dados única. Era, portanto, essencial um modelo que aprendesse essas características sem as assimilar a padrões de outras línguas, como transcrever 'ão' em 'ano', por exemplo. O modelo desenvolvido, denominado "Early Portuguese Printing" (EPP), alcançou em sua versão final (VIIb) uma Taxa de Erro de Caracteres (CER) de 2,58%, com um training set de 142.606 palavras. Embora o CER de um modelo para impressos idealmente se situe abaixo de 2%, o resultado do EPP é notável, considerando a complexidade de um *corpus* diacrônico que abrange três séculos de intensa variação tipográfica. Tornado público no Transkribus em outubro de 2024[8], o EPP é capaz de reconhecer com precisão diacríticos, letras gregas e outros elementos gráficos específicos do português histórico.

3.2. A Coleta dos dados

Para a análise da colocação de clíticos pronominais, este estudo adota uma síntese das abordagens de Martins (1994) e de Galves, Britto e Paixão de Sousa (2005) (doravante GBPS, 2005). A escolha pelo enquadramento de GBPS (2005) é estratégica, pois permite a comparação direta dos resultados obtidos no *corpus* de gramáticas com os de estudos quantitativos de larga escala, como o realizado pelas autoras no *corpus Tycho Brahe*. GBPS dividem os ambientes de variação em dois contextos principais, denominados Contexto I e Contexto II, divididos por diferentes "padrões de variação empírica (GBPS, 2005, p. 3).

O Contexto de Variação I, conforme originalmente proposto por GBPS (2005), inclui sentenças matrizes finitas afirmativas em que o complexo verbo-clítico é precedido por sujeitos referenciais não focalizados, advérbios pré-verbais que não são advérbios de VP fronteados, e sintagmas preposicionais pré-verbais não fronteados. Para este artigo, esse contexto foi ampliado para incluir todas as sentenças "não-dependentes", conforme descrito por A. M. Martins (1994, 2011, 2016), o que engloba sentenças simples, principais matrizes e coordenadas não-disjuntivas. A. M. Martins (1994, p. 15) afirma que, com exceção das coordenadas disjuntivas, não há variação na colocação dos pronomes dependente da natureza de simples/principais ou coordenadas e que dados de seu *corpus* notarial (1994) confirmam a irrelevância dessa distinção para a colocação dos clíticos. A qualidade

[8] Disponível em: transkribus.org/model/early-portuguese-printing. Os detalhes mais específicos do treinamento, como a lista de caracteres, as escolhas filológicas adotadas e uma descrição mais extensa do processo de treinamentos estão descritos em Pacheco Rocha (2025), disponível em: repositorio.ufsc.br/handle/123456789/265186.

das sentenças não-dependentes como contexto de variação para a colocação dos pronomes clíticos é reconhecida, inclusive, por GBPS (2005, p. 1). Dessa forma, buscamos uma análise de um contexto de variação diacrônica mais amplo, mantendo a minúcia qualitativa das análises de GBPS (2005), mas não circunscrita apenas às matrizes, como o estudo de 2005. Exemplos de cada subclasse extraídos do *corpus* incluem sentenças com sujeito pré-verbal como em (1), com advérbios pré-verbais como em (2) e com sintagmas preposicionais pré-verbais, como em (3):

(1) **A nóssa linguágem declinas<u>se</u>** em outras duas a hũa podemos chamár, uogál, (Barros, 1540, l. 28)

(2) **& graçiosa & cõpridamẽte <u>me</u>** conserua nella poys (Oliveira, 1536, l. 9)

(3) **Com isto <u>se</u> encontraõ** os Portugueses em tudo (Leão, 1606, p. 126)

O Contexto de Variação II como aqui o definimos, que segue GBPS (2005), compreende sentenças segundas coordenadas verbo-iniciais e sentenças em que o verbo é precedido por uma oração dependente. Neste artigo, preservamos a concepção original das autoras para esses casos. No entanto, a classe de "matrizes com subordinadas fronteadas" foi desconsiderada na análise quantitativa devido às suas particularidades analíticas, como extensão da sentença fronteada e o impacto estatístico dessa extensão na colocação dos clíticos, elementos que não coincidem com o objetivo deste trabalho. A análise do Contexto II, portanto, se resume às segundas coordenadas verbo-iniciais, dada a percepção de que o gênero textual das gramáticas privilegia tais construções coordenadas, conforme dados em (4) e (5) no que segue.

(4) os gregos dádo que tenham estas diferenças de gẹnero, **conheçẽ<u>nô</u>** per artigos. (Barros, 1540, l. 22)

(5) Substantivo Léva, fallando de gente de guerra e. o verbo léva, lévo. **fecha-<u>se</u>** porem no nome Trevo, e nos Verbos Devo. (Lima, 1736, p. 19)

As sentenças foram extraídas dos arquivos transcritos utilizando expressões regulares (*regex*) e a subsequente classificação sintática dos dados seguiu um método de anotação manual em planilha eletrônica.

4. Os resultados e a discussão: a competição de gramáticas nas grammaticas, instanciadas na sintaxe dos clíticos pronominais

Antes de apresentar, de fato, os resultados e a discussão dos dados, sumarizam-se abaixo, na tabela 1, os totais transcritos e coletados neste trabalho. O objetivo em sistematizar os dados desse modo é descrever a extensão de cada obra no *corpus* (total de palavras) e a distribuição absoluta dos dados coletados, evidenciando o peso quantitativo dos subcontextos (sujeito-verbo e coordenadas

verbo-iniciais), os quais serão objeto de escrutínio analítico isolado nas seções subsequentes. Nossa justificativa para apresentar os resultados individuais em subcontextos encontra respaldo em estudos dessas estruturas na história do português, que mostram a relevância da presença do sujeito em posição pré-verbal no condicionamento da ênclise ou da próclise e em primeiras coordenadas com o verbo em posição inicial não absoluta (GBPS, 2005; Carneiro, 2005; Martins, 2009; Cavalcante; Galves; Paixão de Sousa, 2015; Galves; Paixão de Sousa, 2017)

Tabela 1: Total de dados

Autor / Soma	Nº total de palavras	Nº de Sentenças com clíticos coletadas	Nº de Sentenças com clíticos em contexto de Variação 1	Nº de Clíticos em contexto Sujeito-Verbo	Nº de Clíticos em segundas coordenadas verbo-iniciais
Barros	23.873	511	33	18	34
Oliveira	22.974	423	23	9	28
Leão	21.263	425	34	13	33
Lima	26.997	720	62	24	36
Bacellar	25.219	341	31	23	14
Aulete	21.353	558	65	44	23
Totais	141.679	2.978	248	131	168

Fonte: Pacheco Rocha, 2025

Apresentamos no que segue os resultados e a discussão dos dados relevantes coletados das obras. A apresentação está dividida em três subseções: a primeira, 4.1, aborda o Contexto de Variação I; a segunda trata das sentenças com sujeitos pré-verbais, analisadas isoladamente; e a terceira, as segundas coordenadas verbo-iniciais.

4.1. Contexto de variação I

Originalmente descrito por GBPS (2005), esse ambiente sintático compreende as sentenças matrizes com sujeitos pré-verbais, com advérbios pré-verbais que não são advérbios de VP fronteados, e com sintagmas preposicionais pré-verbais não fronteados. Conforme explanamos na metodologia, o escopo desse contexto se expande aqui para abranger todas as "sentenças não-dependentes" (A. M. Martins, 1994), o que inclui, além das matrizes, as orações simples e as coordenadas não-disjuntivas.

Tabela 2: Totais de Ênclises e Próclises em Contextos de Variação

Autor / Soma	Nº/% de Ênclises		Nº/% de Próclises	
Barros	5	15,2%	28	84,8%
Oliveira	2	8,7%	21	91,3%
Leão	3	9,7%	28	90,3%
Lima	11	17,7%	51	82,3%
Bacellar	13	41,9%	18	58,1%
Aulete	62	93,9%	4	6,1%
Total	94	38,7%	149	61,3%

Fonte: Pacheco Rocha, 2025

Como dito, este contexto é composto por sentenças sujeito-verbo como (6), advérbio-verbo como em (7) e sintagma preposicional-verbo como em (8).

(6) **Na syntaxica reconhecem-se** os periodos, as orações, e a relação que as palavras tem entre si. (Aulete, 1864, p. 88)

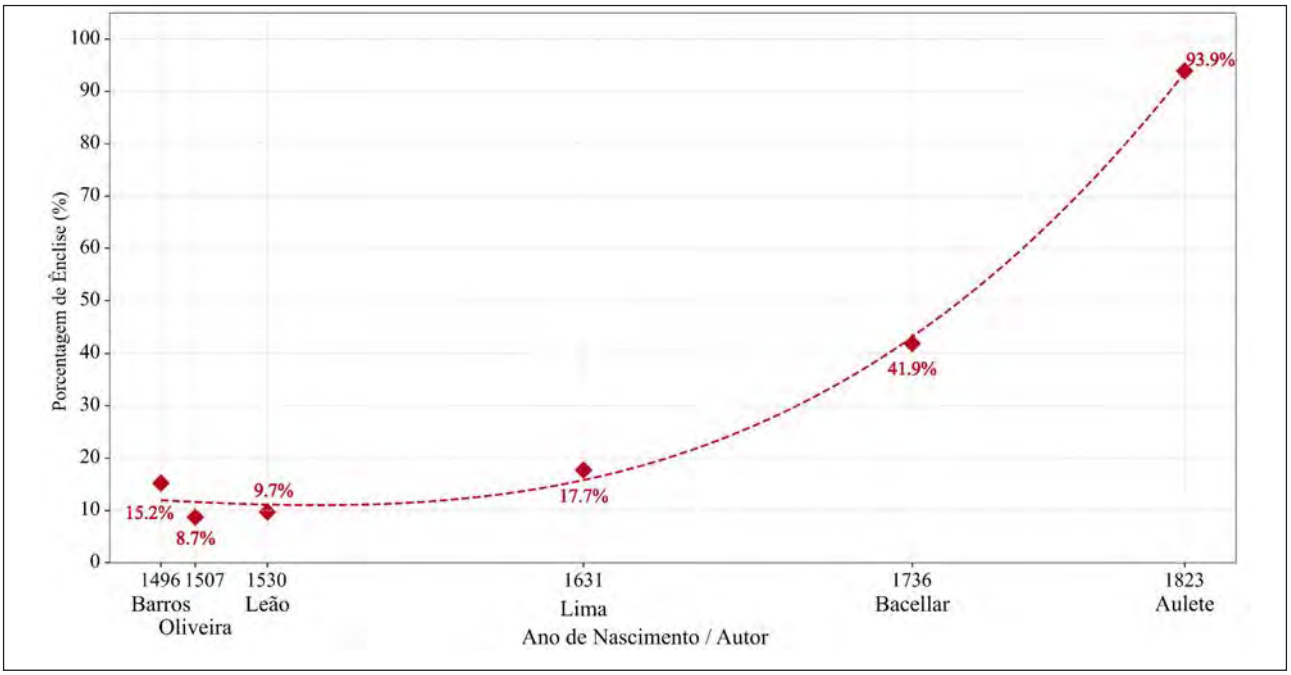
(7) **Daqui infereme** que pronunciamos Joaom, Japaom, e reputão-nos grosseiros. (Bacellar, 1783, p. 122-123)

(8) Sendo as letras SCH seguidas de O, como succede em Schola, Scholium, **accrescentase-lhe** em Portuguez hum E, (Lima, 1736, p. 210)

A progressão do uso da ênclise nesse contexto, desde o padrão proclítico do século XV até o domínio enclítico do século XIX, pode ser visualizada, nas gramáticas aqui analisadas, no gráfico a seguir:[9]

[9] Para evidenciar o padrão evolutivo teorizado e filtrar flutuações amostrais de cada autor, optamos por uma curva de tendência estimada via regressão polinomial de terceiro grau (cúbica) neste e nos demais gráficos deste trabalho. Diferente de abordagens lineares ou logísticas, o polinômio de grau três captura melhor as oscilações não lineares inerentes ao percurso diacrônico, essa escolha permite isolar a tendência teórica, de forma a ilustrar a dinâmica histórica dos dados para além das variações pontuais da amostra.

Gráfico 1: Evolução da Ênclise em Contextos de Variação I



Fonte: Pacheco Rocha, 2025

A evolução das ênclises no contexto de variação I dos gramáticos revela uma curva de mudança que se alinha o já descrito a respeito dessa mudança (*cf.* A. M. Martins 1994; GBPS, 2005); além disso, é perceptível como a curva também segue o descrito a respeito da sócio-histórica da língua, com a ascensão da gramática do PE após as gerações de 1640. O autor mais antigo, João de Barros (1496), nascido ainda no século XV, portanto, apresenta taxas de ênclise de 15,2%, o que claramente indica uma gramática de base proclítica típica do PCl, mas com uma taxa relativamente alta de ênclises quando comparadas com Oliveira e Leão, o que pode ser interpretado como um reflexo da sobrevivência de traços do PA, um sistema mais antigo e enclítico. Essa característica se conecta à biografia de Barros, na medida em que, embora tenha tido uma origem comum a Oliveira, ambos nasceram na Província da Beira e pertencentes a uma geração ainda próxima, ao contrário de Oliveira, Barros, em particular, cresce inserido num ambiente presumivelmente conservador de uma posição subalterna da Corte Real.

Oliveira (1507) e Duarte Nunes de Leão (1530) representam, na curva deste contexto, o auge do que se esperaria da produção do PCl no *corpus*, com as taxas de ênclise em ambiente de variação mais baixas dentre todos os autores, 8,7 e 9,7%, respectivamente. Esse padrão fortemente proclítico, como também era de se esperar, se conecta às suas biografias neste caso. Oliveira escrevera sua gramática, pelo menos em parte, na Espanha ou recém chegado dela, e a sua produção é, pelo menos ideologicamente, inspirada na gramática de António de Nebrija (Buescu, 1978, p. 49-50). Enquanto Leão era um historiador e intelectual alinhado ao governo filipino durante a União Ibérica, e publicou obras também em castelhano, num período de máxima influência do castelhano. Nesse contexto, a língua castelhana já havia generalizado a próclise em contextos finitos (Paixão de Sousa, 2004).

Esses dados, neste contexto, fornecem, assim, um testemunho linguístico do que se pode denominar "Português Hispânico" (Galves, 2012).

A virada geracional e política se manifesta na obra de Luís Caetano de Lima (1671). Nascido em uma Portugal já independente, sua produção mostra um retorno da ênclise ao patamar de 20% similar ao do primeiro autor, Barros. Este fato, é verdade, até poderia ser interpretado como um início de produções identificáveis como do padrão do PE, mas como os dados são poucos, a produção não é tão distinta da de Barros, e a curva média não demonstra, de fato, um desvio, a produção mais enclítica de Lima nesse contexto é mais confiavelmente explicada pela diminuição da influência do castelhana em sua produção, já que nada em sua biografia indica que Lima tivesse conhecimento ou pelo menos um uso constante o bastante de castelhano para desviar o seu uso de um patamar já estabelecido na sua própria gramática, a do PCl.

A competição de gramáticas se torna inequívoca, de fato, com a produção de Melo Bacellar (1736). Sua grammatica apresenta uma taxa de ênclise de 41% nesse contexto, um padrão intermediário que não teria como ser explicado senão entre uma competição com uma forma inovadora, conflito entre a gramática do PCl e a emergente gramática do PE. A biografia de Bacellar, como exposto, mostra como o autor, nascido já no século XVIII, está inserido num universo cultural completamente diferente de seus antecessores nessa diacronia, o autor fora um "estrangeirado", um português que abandonou sua terra natal e viveu imerso nas ideias do iluminismo francês e do racionalismo, sua produção, tanto de ênclises quanto de ideias, para o caso da *grammatica philosophica*, o posiciona como a ponte entre a velha esfera cultural ibérica e a nova, de base francófila e mais autocentrada em Portugal, que permitirá o surgimento dos traços rastreáveis do PE na produção dos autores neste período. Finalmente, a obra de Francisco Júlio Caldas Aulete (1823) conclui essa trajetória com uma taxa de ênclise de 89%, marcando um passo ainda mais avançado na consolidação e implementação da gramática do PE no gênero normativo do século XIX.

4.2. Sentenças com sujeitos pré-verbais

Com a análise geral do Contexto de Variação I estabelecida, esta seção se aprofunda em um de seus componentes mais reveladores: as sentenças não-dependentes finitas com sujeitos pré-verbais. A opção metodológica por isolar este subcontexto se alinha ao procedimento de Galves, Britto e Paixão de Sousa (2005)[10]. As autoras demonstram que a curva de ascensão da ênclise se torna

[10] Ademais, conforme Galves, Britto e Paixão de Sousa (2005) apontam, isolar o contexto de sujeito pré-verbal permite um controle analítico mais fino sobre o comportamento do clítico de terceira pessoa, historicamente associado a construções de matriz passiva ou indeterminadora com o *se*. No entanto, o presente trabalho optou por realizar uma análise total, sem segmentar quantitativamente o clítico *se* das demais pessoas pronominais, por duas razões: primeiramente, pela própria natureza do gênero textual investigado (as *grammaticas*), cuja prosa metalinguística naturalmente impõe um condicionamento discursivo direcionado a formulações impessoais e de apassivamento; em segundo lugar, devido ao volume absoluto de dados coletados por autor neste subcontexto específico (*cf.* tabela 3). Uma fragmentação estatística ulterior reduziria a amostragem a índices marginais por cada categoria, o que comprometeria a representatividade e a confiabilidade das generalizações percentuais necessárias para o mapeamento da curva diacrônica a que o trabalho se propõe. Ver, ainda, para uma argumentação em relação à separação dos subcontextos SV, Carneiro (2005), Martins (2009), Cavalcante, Galves e Paixão de Sousa (2015) e Galves e Paixão de Sousa (2017).

metodologicamente mais clara e pronunciada quando se observa estritamente o contexto Sujeito-Verbo, separando-o de outros elementos pré-verbais. Este contexto isola, portanto, as estruturas em que um sujeito referencial e não focalizado precede o verbo, como em (9) e (10):

(9) **Os Italianos <u>a</u> representão** acrescẽtando hũ g. (Leão, 1606, p. 131)

(10) **O som consoante com a vogal çhama-<u>se</u>** syllaba. (Bacellar, 1783, p. 22)

O total de dados deste contexto está sistematizado na tabela abaixo:

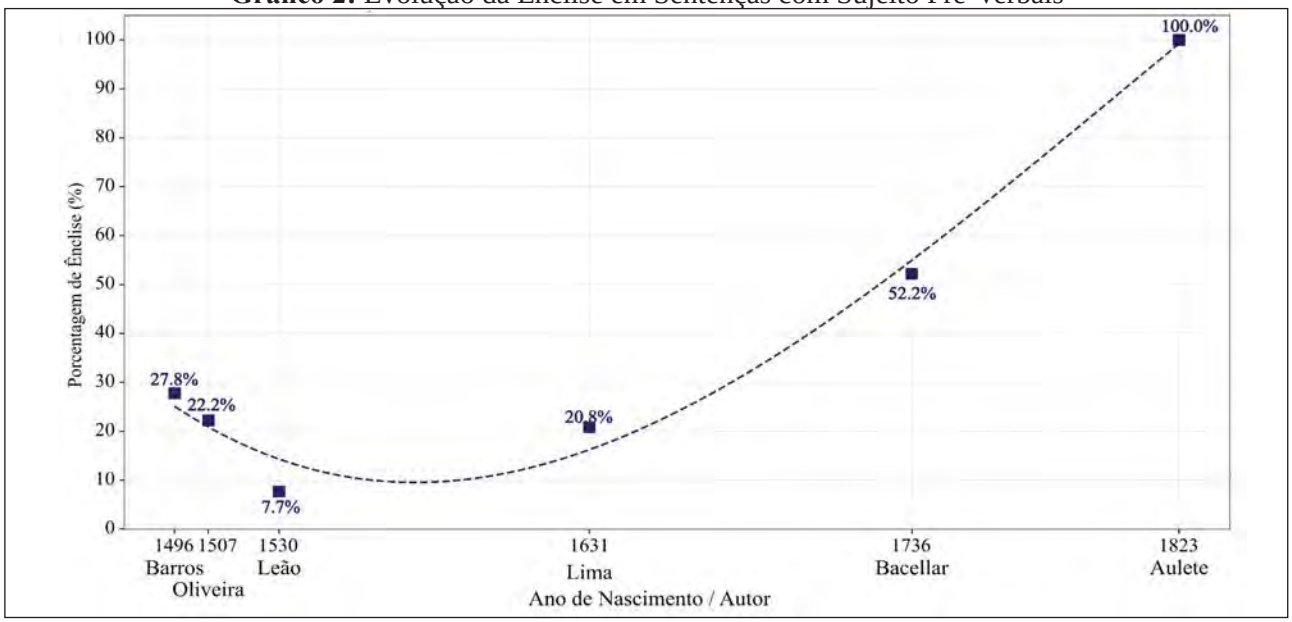
Tabela 3: Totais de Dados em Sentenças com Sujeitos Pré-Verbais

Autor / Soma	Nº/% de Ênclises		Nº/% de Próclises	
Barros	5	27,8%	13	72,2%
Oliveira	2	22,2%	7	77,8%
Leão	1	7,7%	12	92,3%
Lima	5	20,8%	19	79,2%
Bacellar	12	52,2%	11	47,8%
Aulete	44	100,0%	0	0,0%
Total	69	52,7%	62	47,3%

Fonte: Pacheco Rocha, 2025

A evolução deste padrão sintático no *corpus* de gramáticas, desde o século XVI até o século XIX, demonstra uma transformação sistemática na gramática dos autores, conforme visualizado no gráfico a seguir:

Gráfico 2: Evolução da Ênclise em Sentenças com Sujeito Pré-Verbais



Fonte: Pacheco Rocha, 2025

Quando tomado isoladamente, esse contexto revela uma proximidade muito maior entre a produção de João de Barros (1496) e Fernão de Oliveira (1507), que exibem taxas de ênclise de 27,78% e 22,22%, respectivamente, embora o padrão seja majoritariamente proclítico, característico do PCl. O curioso é que, nesse caso, a seleção das sentenças apenas com sujeitos pré-verbais separa com mais expressividade a produção de Oliveira (1507) da de Leão (1530), que produziu apenas 7,7% de ênclises nesse contexto. A explicação para essa constatação está, para a análise deste trabalho, justamente na relação mais intrínseca com a cultura e língua castelhana que a biografia de Leão apresenta, já que, além de desempenhar funções públicas no governo filipino e dedicar suas obras ao rei espanhol Felipe II, que unificou Portugal e Espanha numa invasão armada, Leão é o único dos autores nesta diacronia que escreveu e publicou obras em castelhano. Esse uso acentuadamente proclítico, portanto, isola Leão dos demais e se alinha não apenas com a noção de um "Português Hispânico" (Galves, 2012), como até com as descrições mais desmedidas de A. M. Martins (2016, p. 418) de que a produção proclítica desse período se trataria de uma "variedade socioletal [entre a elite política, social e cultural] que tenderia para a convergência com o espanhol relativamente à generalização das próclises em frases finitas; ao mesmo tempo".

Outro indício de que a produção de Leão de fato está fora da curva do esperado para a produtividade de ênclises do PCl neste contexto é a estabilidade que o autor nascido em 1671, Luís Caetano de Lima, mostra, com uma taxa de ênclises de 20,83%, um retorno aos níveis mais claramente compatíveis com o PCl de Barros, e Oliveira. Este movimento de retomada da curva após Leão, portanto, é mais facilmente explicável por uma produtividade "normal" e média do PCl neste gênero de produções linguístico-normativas. A transição para o PE só se torna explícita com Bacellar (1736). Sua produção, com 52,17% de ênclises, é ainda mais alta e representa, ao mesmo tempo, um ponto de equilíbrio na competição de gramáticas, um período em que os padrões clássico e moderno coexistiram de forma paritária, e uma curva de implementação ainda mais íngreme que o Contexto de Variação I. Indício que corrobora essa interpretação é a produção de Aulete ((nascido em 1823), cuja gramática apresenta 100% de ênclises neste contexto. Este uso categórico demonstra a implementação definitiva desse padrão sintático representativo da gramática do PE no discurso normativo português.

A obra de Aulete apresenta uma particularidade metodológica que a distingue das demais no *corpus*. Por ser a primeira grammatica explicitamente voltada ao ensino, é repleta de excertos de autores clássicos (Camões, Antônio Vieira, Almeida Garrett, Sá de Menezes etc.), cujos textos reproduzem padrões sintáticos do PCl. Para isolar a grammatica do próprio Aulete, foi realizada uma depuração dos dados, expurgando todas as ocorrências de próclise encontradas nessas citações. Essa filtragem revela que, embora Aulete usasse os clássicos como autoridade cultural, a prosa de Aulete é, ainda assim, regida por sua gramática adquirida, a do português europeu moderno do século XIX.

A segunda particularidade da obra de Aulete é sua influência decisiva na consolidação da norma culta no Brasil durante a segunda metade do século XIX. Sua *grammatica Nacional* representa, na diacronia aqui estabelecida, a implementação de um padrão sintático do PE sobre um antigo padrão,

ligado à gramática do PCl. O ponto crucial é que foi precisamente esta obra que a elite letrada brasileira adotou como referência. Conforme exposto, a gramática foi não apenas implementada no Colégio Pedro II, mas também amplamente comercializada e influente no Brasil, consolidando um padrão linguístico em conformidade com o padrão europeu moderno num momento-chave da formação da identidade nacional brasileira.

4.3. Segundas coordenadas com verbo-iniciais

Esta seção se dedica ao último contexto de variação selecionado neste trabalho: o das segundas coordenadas com verbo-iniciais. Este contexto faz parte do que GBPS (2005) denominam do Contexto de Variação II por apresentar uma dinâmica de mudança diferente da do Contexto I. As segundas coordenadas verbo-iniciais são como (11) e (12), em que, dentro de uma sequência de sentenças em coordenação, a segunda se inicia com o complexo clítico-verbo sem um sujeito ou outro elemento anteposto.

(11) E daquellas gẽtes, com q̃ os Hespanhoes assitiuhaõ comercio & vezinhança, tomarão hũ as lingoas & **as confundirão** cõ a sua, (Leão, 1606, p. 16)

(12) Se a ellas se seguir a Vogual I, lança-se sóra o S, e **pronunciase** o CH como simplez C. (Lima, 1736, p. 210)

A sistematização do total de dados deste contexto no *corpus* de grammaticas segue abaixo, bem como a soma de seus totais:

Tabela 4: Total de Dados em sentenças Segundas Coordenadas Verbo-Iniciais

Autor / Soma	Nº/% de Ênclises		Nº/% de Próclises	
Barros	22	64,7%	12	35,3%
Oliveira	17	60,7%	11	39,3%
Leão	9	27,3%	24	72,7%
Lima	19	52,8%	17	47,2%
Bacellar	8	57,1%	6	42,9%
Aulete	19	82,6%	4	17,4%
Total	94	56,0%	74	44,0%

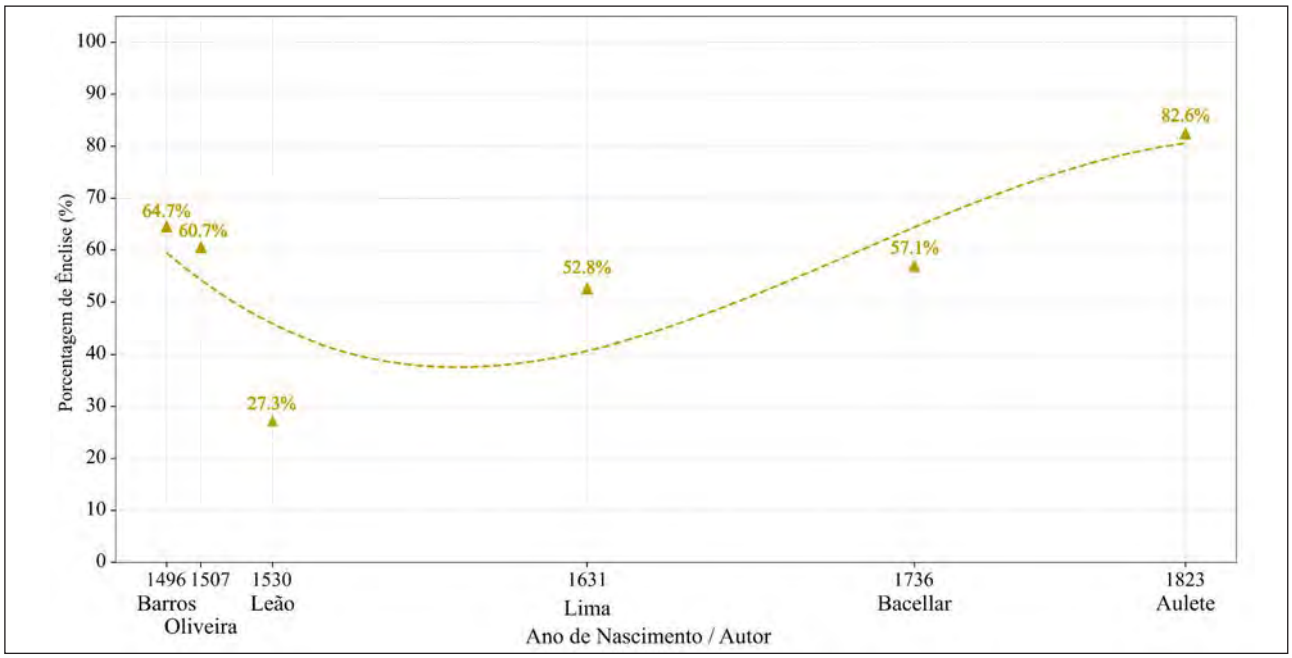
Fonte: Pacheco Rocha, 2025

A relevância deste contexto para a análise diacrônica é particular, já que, diferentemente duma simples dinâmica de substituição, GBPS (2005) mostram que este é um ambiente em que, desde o início do período analisado no século XVI, já se observavam altas e muito variáveis taxas de

ênclise entre os diferentes autores presentes nos corpora do *Tycho Brahe*. A mudança observável, portanto, não é um recuo sistemático da próclise, mas sim um período de "variação idiossincrática", talvez sensíveis a fatores prosódicos (GBPS, 2005, p. 11), que abre espaço para a generalização e estabilização da ênclise a partir de 1700. A análise desta construção, na qual a segunda oração se inicia pelo complexo verbo-clítico, é especialmente produtiva para este *corpus*, dada sua alta frequência na prosa dos gramáticos, sobretudo nas obras mais antigas.

A evolução da colocação clítica neste contexto específico nas grammaticas, como descrito por GBPS (2005) de fato revela um padrão diferente daquele observado anteriormente para os Contextos I. A trajetória da ênclise ao longo dos séculos, conforme os dados das *grammaticas*, é visualizada no gráfico a seguir.

Gráfico 3: Evolução da Ênclise em Sentenças Segundas Coordenadas Verbo-Iniciais



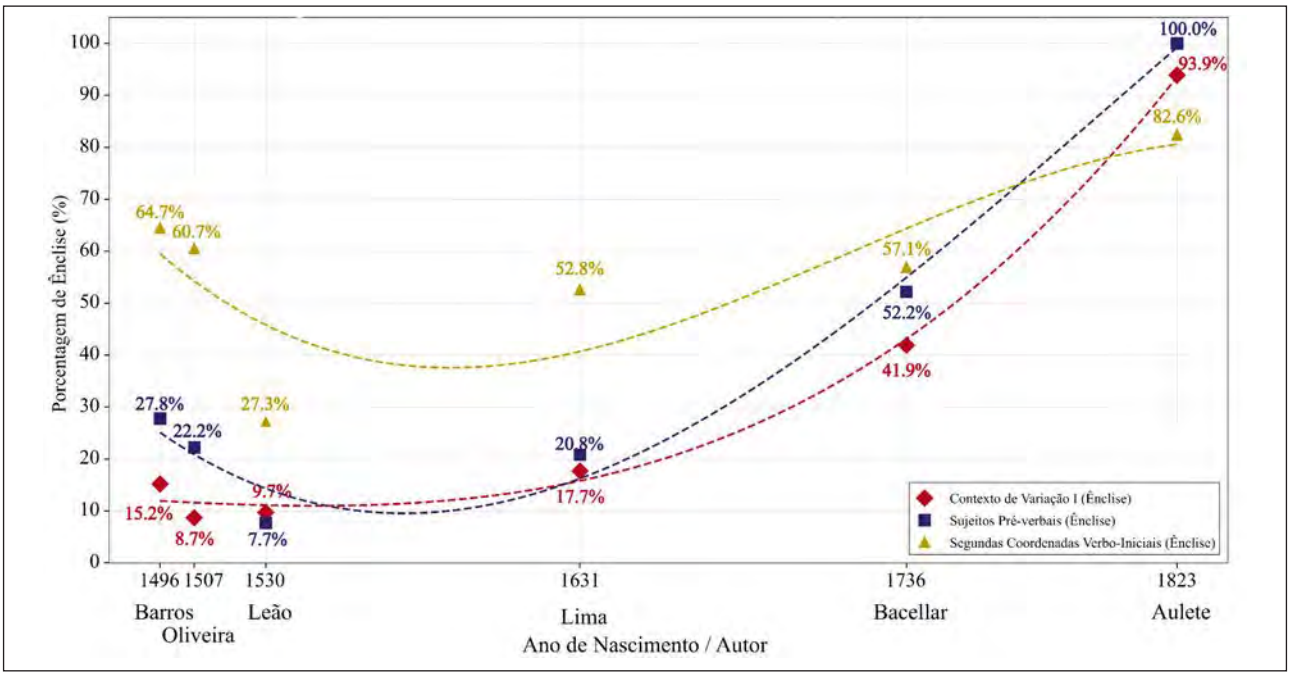
Fonte: Pacheco Rocha, 2025

A mudança pode ser caracterizada por dois pontos principais. O primeiro é a aparente estabilidade na produção de quatro dos seis autores. Barros (1540), Oliveira (1536), Lima (1736) e Bacellar (1783) apresentam taxas de ênclise que se mantêm em um equilíbrio notável, variando entre 52% e 65%. Este padrão sugere que a gramática do PCl licenciava uma variação estável entre próclise e ênclise neste contexto, uma espécie de indiferença que permitia ambas colocações, mesmo durante a fase de transição para o PE representada por Bacellar. O segundo ponto é a forma como Leão (1606) e Aulete (1864) rompem essa estabilidade em direções opostas. A produção de Leão, com apenas 27% de ênclises, se desvia drasticamente dos dados dos demais autores da mesma geração e se afasta da curva de tendência que ligaria Oliveira a Lima. Leão monstra, também nesse contexto, uma preferência clara pela próclise que não se comporta como a dos demais autores. Este dado,

não é surpresa, novamente reforça a hipótese de uma forte influência da gramática do castelhano, que no período já havia restringido a ênclise em contextos finitos de forma mais rigorosa. No outro extremo, Aulete, com 82% de ênclises, marca o fim do equilíbrio anterior e a consolidação do padrão enclítico do PE neste contexto, seguindo a mesma tendência de sua produção em sentenças com sujeito pré-verbal.

A comparação com os dados de GBPS (2005) para o mesmo contexto revela um cenário distinto. Enquanto o nosso *corpus* de grammaticas mostra uma curva relativamente estável, os dados do corpus *Tycho Brahe* são muito mais dispersos, especialmente nos dados referentes ao PCl, antes do século XVII, o que levou as autoras a classificarem, a princípio, a variação neste contexto como "idiossincrática", com taxas de ênclise que podiam variar de 20% a 80% em autores do mesmo período.

Gráfico 4: Evolução da Ênclise em Grammáticas Históricas Portuguesas (Séculos XV-XIX)



Fonte: Pacheco Rocha, 2025

Para explicar a aparente estabilidade encontrada no *corpus* de grammaticas em contraste com a dispersão no corpus de GBPS, a hipótese mais provável é de que a estabilidade seja uma coincidência resultante da pequena amostragem de obras. Com mais autores, a variação poderia se mostrar igualmente idiossincrática e mais pontos fora da curva como Leão puxariam a curva para baixo. Ainda assim, como exposto anteriormente para os outros contextos, a produção de Leão é justamente a que foge ao padrão dos autores clássicos e produz ainda mais próclises também nesse contexto é muito reveladora da natureza dessa variação, a permeabilidade da sintaxe do castelhano nos autores bilíngues anteriores ao século XVII pode, aqui, ser também uma chave explicativa para essa variação diacrônica.

5. Considerações finais

A análise quantitativa do *corpus* de seis gramáticas históricas apresentada neste artigo permitiu um novo mapeamento da trajetória da colocação dos clíticos pronominais no português europeu entre os séculos XV e XIX. Confirmamos que não houve, por conta do gênero textual, desvio da já conhecida curva de transição de um sistema de base proclítica, característico do PCl, para um padrão majoritariamente enclítico, indício de uso da gramática do PE, seguindo uma curva de mudança consistente com as descrições do fenômeno em GBPS (2005) e A. M. Martins, (1994). A análise mostrou, ainda, uma forte correlação entre os padrões sintáticos de cada autor e seu respectivo contexto sócio-histórico, com destaque para a influência do castelhano em Duarte Nunes de Leão (1606) e a ascensão do padrão do PE durante um período cultural marcado pela francofila das camadas letradas portuguesas no século XVIII, representada por Bernardo Bacellar (1736).

Do ponto de vista metodológico, este trabalho valida a utilização de textos de natureza prescritiva como fonte para a investigação da sintaxe diacrônica, mostrando que a prosa das *grammaticas* é, também ela, um palco que revela variação e mudança linguística, o que se torna particularmente evidente no caso de Caldas Aulete e a norma brasileira da virada dos séculos XIX e XX. Sua *Grammatica Nacional* (Aulete, 1864) não apenas mostra a consolidação da gramática inovadora do PE sobre a conservadora do PCl, mas também atua como um veículo de exportação dessa mesma norma para o Brasil. A adoção de sua obra pela elite letrada brasileira na segunda metade do século XIX, inclusive no Colégio Pedro II, ilustra de forma contundente como textos normativos europeus foram agentes ativo na disseminação de um padrão linguístico estrangeiro em um momento-chave de formação identitária brasileira, deixando não apenas vestígios estatísticos no uso linguístico daquele período, como também cicatrizes no atual ensino de língua portuguesa brasileiro.

Por fim, o modelo Early Portuguese Printing (EPP) e o *corpus* aqui apresentado constituem não apenas os instrumentos desta pesquisa, mas também um ativo para a comunidade acadêmica. A aplicação de ferramentas das Humanidades Digitais para a criação de *corpora* de alta fidelidade a partir de documentos históricos não é um fim em si mesma, mas um meio para viabilizar novas e mais aprofundadas investigações sobre a história das línguas, e das ideias. O avanço sobre os vestígios escritos do passado depende, cada vez mais, da construção dessa infraestrutura digital de acesso livre.

Referências

ASSUNÇÃO, Carlos; KEMMLER, Rolf; FERNANDES, Gonçalo; COELHO, Sónia; FONTES, Susana; MOURA, Teresa. *A orthographia da lingua portugueza (1736) de Luís Caetano de Lima*: estudo introdutório e edição. Vila Real: Centro de Estudos em Letras / Universidade de Trás-os-Montes e Alto Douro, 2022. (Ortógrafos Portugueses, 7).

AULETE, Francisco Júlio Caldas (1823-1878). Grammatica nacional. Lisboa: Typographia da Sociedade Typographica Franco-Portugueza, 1864. 96 p.; 20 cm. Cota do exemplar digitalizado: L-4532-V. Disponível em: https://purl.pt/17381. Acesso em: 22 jul. 2026.

BACELLAR, Bernardo de Lima e Melo (1736-c.). *Grammatica philosophica e orthographia racional da lingua portugueza*. Lisboa: Officina de Simão Thaddeo Ferreira, 1783. 196, [1] p.;8º (15 cm). Cota do exemplar digitalizado: RES-5748-P. Disponível em: https://purl.pt/29093. Acesso em: 22 jul. 2026.

BARROS, João de (1496-1570). *Grammatica da lingua portuguesa*. Olyssipone: apud Lodouicum Rotorigiu[m], Typographum, 1540. 60 f.; 4º (20 cm). Cota do exemplar digitalizado: RES. 5658 1 P. Disponível em: https://purl.pt/12148. Acesso em: 22 jul. 2026.

BENGIO, Yoshua. Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, Hanover, v. 2, n. 1, p. 1-127, 2009. DOI: https://doi.org/10.1561/2200000006.

BUESCU, Maria Leonor Carvalhão. *A Gramática da Linguagem Portuguesa (Fernão de Oliveira, 1536)*: Introdução, leitura actualizada e notas. Lisboa: Imprensa Nacional-Casa da Moeda, 1975.

BUESCU, Maria Leonor Carvalhão. *Gramáticos portugueses do século XVI*. Venda Nova: Instituto de Cultura e Língua Portuguesa, MEC – Secretaria de Estado da Cultura, 1978.

BUESCU, Maria Leonor Carvalhão. *Babel ou a ruptura do signo*. Lisboa: Imprensa Nacional/Casa da Moeda, 1983.

CARNEIRO, Zenaide de Oliveira Novais. *Cartas brasileiras (1809-1904)*: um estudo linguístico-filológico. 2005. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2005.

CAVALCANTE, Silvia Regina de Oliveira; GALVES, Charlotte; PAIXÃO DE SOUSA, Maria Clara. Topics, subjects and grammatical change: from Classical to Modern European Portuguese. *LaborHistórico*, Rio de Janeiro, v. 1, n. 2, pp. 97-107, 2015. Disponível em: https://revistas.ufrj.br/index.php/lh/article/view/4796/3504. Acesso em: 09 jun. 2026. DOI: https://doi.org/10.24206/lh.v1i2.4796.

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

GALVES, Charlotte. Periodização e competição de gramáticas: o caso do português médio. *In*: LOBO, Tânia; CARNEIRO, Zenaide; SOLEDADE, Juliana; ALMEIDA, Ariadne; RIBEIRO, Silvana (org.). *Rosae*: linguística histórica, história das línguas e outras histórias [*online*]. Salvador: EDUFBA, 2012. pp. 65-74.

GALVES, Charlotte; BRITTO, Helena; PAIXÃO DE SOUSA, Maria Clara. The change in clitic-placement from Classical to Modern European Portuguese: results from the Tycho Brahe Corpus. *Journal of Portuguese Linguistics*, v. 4, pp. 39-67, 2005.

GALVES, Charlotte; PAIXÃO DE SOUSA, Maria Clara. The change in the position of the verb in the history of Portuguese: subject realization, clitic placement, and prosody. *Journal of Historical Syntax*, [*S. l.*], 2017. Disponível em: https://historicalsyntax.org/hs/index.php/hs/en/article/view/28. Acesso em: 09 jun. 2026. DOI: https://doi.org/10.18148/hs/2017.v0i0.28.

KROCH, Anthony. Reflexes of grammar in patterns of language change. *Language Variation and Change*, v. 1, n. 3, pp. 199-244, 1989.

KROCH, Anthony. Syntactic change. *In*: BALTIN, Mark; COLLINS, Chris (org.). *The handbook of contemporary syntactic theory*. Oxford: Blackwell, 2001. pp. 699-729.

KROCH, Anthony. Mudança sintática. Tradução de Silvia Regina de Oliveira Cavalcante. *Working Papers em Linguística*, Florianópolis, v. 22, n. 2, e86882, 2021. Disponível em: https://periodicos.ufsc.br/index.php/workingpapers/article/view/86882. Acesso em: 08 jun. 2026. DOI: https://doi.org/10.5007/1984-8420.2021.e86882.

LEÃO, Duarte Nunes de (fl. 1530-1608). *Origem da Lingoa portuguesa*. Lisboa: Pedro Crasbeeck, 1606. [6], 150 p.; 4º (18 cm). Cota do exemplar digitalizado: RES. 277//1 V. Disponível em: https://purl.pt/50. Acesso em: 22 jul. 2026.

LIMA, Luís Caetano de, C.R. (1671-1757). *Orthographia da Lingua Portugueza*. Lisboa: Officina de António Isidoro da Fonseca, 1736. [24], 217, [3 br.] p.; 8º (14 cm). Cota do exemplar digitalizado: L-658-P. Disponível em: https://purl.pt/8. Acesso em: 22 jul. 2026.

MACHADO, José Pinheiro (org.). *João de Barros*: Gramática da Língua Portuguesa. 3. ed. Lisboa: Sociedade Astória, 1957.

MARTINS, Ana Maria. *Clíticos na história do português*. 1994. 628 p. Tese (Doutorado) – Universidade de Lisboa, Lisboa, 1994.

MARTINS, Ana Maria. Clíticos na história do português à luz do teatro vicentino. *Estudos de Lingüística Galega*, Santiago de Compostela, n. 3, pp. 83-109, 2011.

MARTINS, Ana Maria; CARRILHO, Ernestina. *Manual de linguística portuguesa*. Berlin: Walter de Gruyter GmbH, 2016.

MARTINS, Marco Antonio. *Competição de gramáticas do português na escrita catarinense dos séculos 19 e 20*. 2009. 326 f. Tese (Doutorado em Linguística) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2009.

MELO, Lívia de. A gramática filosófica de Bernardo de Lima e Melo Bacellar (1783). *Confluência*, Rio de Janeiro, n. 59, p. 166-176, jul.-dez. 2020.

MENDÍVIL-GIRÓ, José Luis. *El cambio lingüístico*: sus causas, mecanismos y consecuencias. Madrid: Síntesis, 2015.

NUNES, José Horta. Dicionarização no Brasil: condições e processos. *In*: NUNES, José Horta; PETTER, Margarida (org.). *História do saber lexical e constituição de um léxico brasileiro*. São Paulo: Humanitas/FFLCH/USP; Pontes, 2002. p. 99-122.

OLIVEIRA, Fernão de (c. 1507-1581). *Grammatica da lingoagem portuguesa*. Lisboa: Germão Galharde, 1536. [38] f.; 4º (20 cm). Cota do exemplar digitalizado: RES-274-V. Disponível em: https://purl.pt/120. Acesso em: 22 jul. 2026.

PACHECO ROCHA, Saulo Rogério. *O Português das "Obras Proueitosas" e dos "Ilustres & Muito Reuerendos Senhores"*: uma análise da mudança nos padrões de colocação pronominal clítica em gramáticas portuguesas do século 15 ao 19 e seus autores. 2025. 202 p. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2025.

PACHECO ROCHA, Saulo Rogério. A *Corpus* of Portuguese Historical and Metalinguistic Grammars (1536-1864). *Journal of Open Humanities Data*, v. 12, n. 65, pp. 1-8, 2026. DOI: https://doi.org/10.5334/johd.542.

PAIXÃO DE SOUSA, Maria Clara. *Língua barroca*: sintaxe e história do português nos 1600. 2004. 450 f. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 2004.

PANEGASSI, Rubens Leonardo. *O pasto dos brutos*: contexto de João de Barros, "horizonte histórico" e política nas Décadas da Ásia. 2013. Tese (Doutorado em História) – Universidade de São Paulo, São Paulo, 2013.

POLACHINI, Bruna Soares. A trajetória da Gramática (trans)Nacional de Caldas Aulete: livros em viagem. *Cadernos do Arquivo Municipal* [*online*], n. 18, 2022. Disponível em: https://cadernosarquivo.cm-lisboa.pt/index.php/am/article/view/33. Acesso em: 05 mar. 2025. DOI: https://doi.org/10.48751/CAM-2022-1833.

RUSSELL-WOOD, John. *Histórias do Atlântico português*. São Paulo: UNESP, 2014.

SANTOS, Eliéte Oliveira. *História de Portugal de Fernão de Oliveira*: edição semidiplomática e análise da retórica nos topônimos de Portugal. 2022. Tese (Doutorado em Letras) – Universidade Federal da Bahia, Salvador, 2022.

SARAIVA, José Hermano. *Portugal*: a companion history. 2. ed. Manchester: Carcanet Press, 2012.

TORRES, Amadeu. Melo Bacellar, filósofo da grammatica, na lexicografia de setecentos. *Revista Portuguesa de Filosofia*, Braga, v. 52, n. 1/4, pp. 935-943, 1996.

VERDELHO, Telmo. Dicionários portugueses, breve história. *In*: NUNES, José Horta; PETTER, Margarida (org.). *História do saber lexical e constituição de um léxico brasileiro*. São Paulo: Humanitas/ FFLCH/USP; Pontes, 2002. pp. 15-64

VERONESI, L. H. L.; GALVES, Charlotte. *The Tycho Brahe Platform*. 2024.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a71961

Variação no emprego de artigo definido diante de pronome possessivo na norma de Maputo

Variation in the use of the definite article before possessive pronouns in the Maputo norm

Danielle Kely Gomes[1]

RESUMO

O artigo descreve as restrições que atuam na variação no emprego do artigo definido diante de possessivo na norma urbana do português em Moçambique, uma variedade marcada por questões de contato entre o português e línguas autóctones do grupo bantu. Os dados foram recolhidos em 18 inquéritos da amostra MOÇAMBIQUE-Port (Vieira; Pissurno, 2016). O instrumental teórico e metodológico da Teoria da Variação e Mudança (Weinreich; Labov; Herzog, 1968; Labov, 2008) auxilia na formatação da hipótese de investigação e na delimitação dos procedimentos metodológicos. Controlaram-se restrições que poderiam afetar o emprego de artigo definido diante de possessivo na comunidade de fala de Maputo. A análise estatística foi operacionalizada com apoio das ferramentas disponíveis no Rbrul. Os resultados indicam que restrições relacionadas à escolaridade, ao tipo de sintagma, à caracterização semântica da posse e à faixa etária do informante são os condicionamentos que atuam para desfavorecer o emprego de definidos diante de possessivos na comunidade. As variáveis que controlam o estatuto de aquisição do português e as línguas de comunicação dominadas pelos informantes não foram selecionadas como relevantes para inibir a ocorrência do artigo definido, mas revelam indícios de que o contato afeta a gramática da comunidade de fala.

PALAVRAS-CHAVE: Artigo definido. Pronome possessivo. Variação. Português em Moçambique. Contato linguístico.

ABSTRACT

The article describes the constraints that influence variation in the use of the definite article before a possessive in the urban norm of Portuguese in Mozambique, a variety marked by issues of contact between Portuguese and Bantu Indigenous languages. The data were collected from 18 surveys in the MOÇAMBIQUE-Port sample (Vieira; Pissurno, 2016). The theoretical and methodological framework of the Theory of Variation and Change (Weinreich; Labov; Herzog, 1968; Labov, 2008) aids in shaping the research hypothesis and defining the methodological procedures. Constraints that could affect the use of the definite article before a possessive in the speech community of Maputo were controlled. Statistical analysis was carried out with the support of the tools available in Rbrul. The results indicate that constraints related to education, the type of phrase, the semantic characterization of possession, and the individual's age group are the factors that act to inhibit the use of definite articles before possessives in the community. The variables controlling the status of Portuguese acquisition and the languages of communication spoken by the informants were not selected as relevant for inhibiting the occurrence of the definite article, but they reveal signs that contact influences the grammar of the speech community.

KEYWORDS: Definite article. Possessive pronoun. Variation. Portuguese in Mozambique. Language contact.

[1] Universidade Federal do Rio de Janeiro (UFRJ), daniellekgomes@letras.ufrj.br, https://orcid.org/0000-0001-6335-1674.

Recebido em 27/12/2025 - Aceito em 10/06/2026

1. Introdução

O emprego de artigo definido diante de pronomes possessivos é uma regra variável descrita em diversos estudos sociolinguísticos que se debruçam sobre as variedades do português brasileiro (Oliveira e Silva, 1982, 1996; Callou; Oliveira e Silva, 1997; Baxter; Lopes, 2009; Menon, 2016; entre tantos outros). No português europeu, as investigações disponíveis revelam que não há variação no uso de artigos diante de possessivos: a variedade se define pelo uso categórico do determinante definido diante de possessivos. A ausência de artigo nesse contexto só é possível em construções cristalizadas (*Nossa Senhora*, *Vossa Senhoria*) e em estruturas com função vocativa (Não, *minha senhora*!), "sendo a não ocorrência do artigo definido nas restantes frases sentidas como anômala" (Inverno; Swolkien; 2003, p. 182).

No âmbito das variedades africanas do português, o panorama sobre a regra variável de emprego de artigo definido diante de possessivo ainda está em construção. Os trabalhos disponíveis se concentram na descrição da regra variável de artigo diante de possessivos em crioulos de base lexical portuguesa (Lucchesi, 1993), em comunidades isoladas caracterizadas por reestruturação parcial do português (Figueiredo, 2019), na variedade urbana do português de São Tomé (Gomes; Cordeiro, 2021), na escrita de estudantes da província de Nampula – província da região norte de Moçambique (Atanásio, 2002), na norma urbana falada em Luanda (Azevedo, 2025).

De forma a ampliar o espectro de análises variacionistas em dados de variedades do português em África, este artigo se dedica a descrever o emprego do artigo definido frente a pronomes possessivos em dados da modalidade oral do português falado em Maputo, capital de Moçambique. Como ilustram os exemplos de [01] a [08], a seguir, o emprego do artigo definido diante de possessivo é uma regra é variável na comunidade:

(01) temos ***a nossa língua nacional que é o português***... (PMO-C-2-m)

(02) fala-se… hum… ***nossa língua tradicional*** (PMO-B-1-m)

(03) Eu ia criar ***meus filhos*** como ***minha mãe*** fez (PMO-A-1-f)

(04) Isso me criou formas de conseguir alcançar ***o meu objetivo*** (PMO-C-2-m)

(05) tu e ***teu marido*** trabalham? (PMO-C-3-f)

(06) Você agora é licenciado em ciência de educação, então ***o teu lugar*** não é aqui (PMO-3-C-m)

(07) N***a minha zona*** muita gente fuma (PMO-A-1-f)

(08) Eu gosto muito de dar ***minha opinião*** (PMO-A-2-m)

Para o atendimento aos objetivos deste artigo, apresentam-se um panorama sobre a variação no emprego de artigo definido diante de pronomes possessivos em português, uma breve reflexão sobre codificação gramatical das noções de especificidade e definitude no contato entre o português e línguas autóctones do continente africano e a caracterização sócio-histórica e linguística da norma urbana do português em Moçambique. Posteriormente, descrevem-se os procedimentos metodológicos e discutem-se os resultados da análise variável. Por fim, tecem-se as considerações finais.

2. O artigo definido diante de possessivo em português

O artigo definido do português provém das formas acusativas do pronome demonstrativo *ille* (*illu(m)/illa(m)*; *illos/illas*). No latim vulgar, *ille* e *ipse* se especializam para estabelecer referência a nomes anteriormente mencionados no discurso. Castro (1991) destaca que a função anafórica dos demonstrativos começa a se esvaziar quando os pronomes passam a se manifestar diante de substantivos mesmo que as formas nominais não tenham sido mencionadas previamente. Com o esvaziamento da função anafórica dos demonstrativos, as formas se especializam e se modificam foneticamente, dando origem ao artigo definido.

O emprego de artigos definidos diante de possessivos é um processo historicamente documentado em português. Rinke (2010), em uma análise quali-quantitativa de dados do português europeu, afirma que

> enquanto o artigo definido acompanha geralmente os sintagmas nominais com pronome possessivo no português actual, o mesmo é sistematicamente omitido no português antigo. A partir do século XIV, existe uma fase de variação, na qual o emprego do artigo é opcional. No século XIX, o uso do artigo em combinação com os pronomes possessivos generaliza-se, com excepção dos termos de parentesco (Rinke, 2010, p. 136)

No português brasileiro, Callou e Oliveira e Silva (1997) atestam que "[N]o Brasil, o percentual de uso de artigo definido diante de possessivos e nomes próprios é quase idêntico, enquanto, em Portugal, o percentual de uso de artigos diante de possessivos é mais alto que diante de nomes próprios." (1997, p. 15). As autoras afirmam que a taxa de emprego de artigo definido diante de possessivos no português brasileiro contemporâneo corresponde à frequência de uso dos definidos na norma europeia entre os séculos XVI e XVIII.

Com base nos dados das capitais mapeadas pelo projeto Norma Urbana Culta, Callou e Oliveira e Silva (1997) identificam um processo variável que separa, por um lado, as duas capitais do Nordeste – Recife e Salvador – e, de outro, Porto Alegre. É possível identificar um *continnum* de aumento no emprego de definidos diante de possessivos na medida em que se percorre o eixo nordeste-sul, conforme ilustra a tabela 1, a seguir.

Tabela 1: *Continuum* de uso de artigos definidos diante de pronomes possessivos nas capitais do Projeto NURC/BR

Localidades	Frequência
Recife	60%
Salvador	66%
Rio de Janeiro	70%
São Paulo	70%
Porto Alegre	79%

Fonte: (Callou; Oliveira e Silva, 1997, p. 21)

Schei (2009) e Magalhães (2011) trazem contribuições importantes para a compreensão da regra variável de emprego de artigos diante de possessivos no português europeu. Schei (2009), apesar de concentrar a investigação no comportamento de artigo definido diante de possessivos em sete romances brasileiros, realizou uma análise do fenômeno em romances portugueses dos séculos XIX e XX. A autora verificou que o emprego de artigo definido no português europeu caracteriza-se em uma regra categórica, tendência notoriamente distinta da que se observa no âmbito do português do Brasil.

Magalhães (2011), em um mapeamento diacrônico da generalização do emprego de artigo definido diante de possessivos na variedade europeia, com base na análise de textos escritos entre os séculos XVI e XIX, observa duas tendências distintas para a evolução do fenômeno: i) em sintagmas nominais, a realização variável de artigo é atestada já no século XVI. No entanto, a variação se torna rarefeita em meados do século XVII, tendo seu desaparecimento total no século XVIII. No século XIX, há uma generalização significativa do emprego de artigo definido, de modo a evidenciar seu uso categórico; ii) em sintagmas preposicionais, a ausência de artigo diante de possessivo no século XVI era categórica. No século XVII os dados atestam o início da variação nesse contexto, tendência que perdurou até o século XIX.

Os trabalhos brevemente apresentados até aqui descrevem a regra variável em foco em *corpora* escritos e/ou normas urbanas das variedades brasileira e europeia. É importante considerar, contudo, estudos que descrevem a variação no emprego do artigo em variedades reestruturadas do português e normas do português ainda em processo de formação. Destacam-se, nesse âmbito, Baxter e Lopes (2009), sobre o falar de Helvécia; Figueiredo (2019), sobre o português de Almoxarife, Atanásio (2002), sobre a escrita de crianças do segundo ano do ensino primário de escolas na cidade de Nampula (província da região norte de Moçambique), Gomes e Cordeiro (2021), sobre a norma urbana são tomense e Azevedo (2025), sobre a norma urbana de Luanda.

Baxter e Lopes (2009) descrevem o comportamento de artigo definido em sintagmas nominais de variadas configurações no dialeto de Helvécia, uma comunidade linguística afro-brasileira, localizada no sul do estado da Bahia. Nos dados da localidade, os autores observaram que o emprego de definidos está correlacionado à estrutura morfossintática do sintagma nominal: sintagmas nominais

modificados por oração relativa não inibiram a ocorrência do artigo definido. Já a presença de pronomes possessivos no sintagma nominal foi uma condição que atuou para desfavorecer o uso do artigo definido. A diferença no comportamento desses dois condicionamentos que atuam na delimitação da referência pode ser atribuída ao fato de que os possessivos atuariam de forma mais consistente na atribuição da referência, sendo redundante a ocorrência de artigo definido nesse contexto.

No português de Almoxarife, variedade do português falada em uma comunidade de São Tomé, Figueiredo (2019) verificou que os homens tendiam a empregar mais o artigo diante de possessivos do que as mulheres. Uma provável justificativa para o comportamento dos informantes do sexo masculino se relaciona ao fato de que os homens têm mais contato com o mundo exterior à comunidade – por mais frequentemente migarem para os centros urbanos, por prestarem o serviço militar obrigatório e por terem um acesso maior à educação formal. As mulheres, ao contrário, apresentam maior variação no emprego de artigo definido, uma vez que tendem a se concentrar mais na comunidade.

No que se refere à atuação da variável social idade do informante, um indício para verificação do estatuto da mudança, Figueiredo (2019) comparou a variedade do português de Almoxarife com os trabalhos sobre as variedades do português dos Tongas, Roça Monte Café, São Tomé (Baxter; Lopes, 2005, 2006) e do português rural de Helvécia, Bahia (Baxter; Lopes, 2009). Na comparação, o autor percebeu no português de Almoxarife há um processo estável de variação, condicionado pelo isolamento ao qual a comunidade esteve sujeita, enquanto em Helvécia e no português dos Tongas há uma indicação de mudança em progresso no sentido de aquisição/emprego mais consistente de artigo definido pelos indivíduos mais jovens da comunidade, ou seja, "aquisição em direção à nivelação com a variante padrão". (Figueiredo, 2019, p. 373)

Atanásio (2002), ao investigar a escrita de crianças do ensino primário na província de Nampula (região localizada ao norte de Moçambique) constata que a ausência de artigos em sintagmas nominais no português de Moçambique é regra nos dados por ele analisados, independentemente da configuração do SN. O autor reforça que a ausência de artigo "não é um fenómeno verificável apenas na camada estudantil, mas abrange toda a sociedade moçambicana, incluindo jornalistas e académicos" (2002, p. 133). Para o autor, os moçambicanos não utilizam o artigo definido por conta de um processo de interferência das línguas bantu no processo de aquisição do português. As línguas bantu de Moçambique, no geral, não possuem um sistema de artigos, sendo a referência marcada por prefixos de classe associados ao nome. Desse modo, "admite-se que a interferência das línguas Bantu sobre o Português seja muito forte, para além da actuação das regras do PE" (Atanásio, 2002, p. 133).

Gomes e Cordeiro (2021), em uma descrição sobre o emprego de artigos diante de pronomes possessivos na norma urbana de São Tomé, capital de São Tomé e Príncipe, identificam uma regra variável de emprego de definidos, uma tendência distinta da observada no português europeu – sua norma de referência. A variação do artigo definido na variedade urbana do português em São Tomé apresenta-se como uma tendência mais próxima da verificada nas variedades do português brasileiro, em que o uso do artigo definido diante de possessivo configura-se como uma regra variável.

Tendo em vista a hipótese de que na norma são tomense a variável ausência de definidos seria a mais produtiva, já que o Forro – crioulo de base portuguesa que coexiste com o português na comunidade – não possui um sistema de artigos, os resultados da análise variacionista frustraram as expectativas. Em um universo de 517 dados, observou-se uma regra de fato variável, com ligeiro favorecimento para a presença de artigo (55%).

No que se refere às condições que inibem o emprego de artigos definidos na norma são-tomense, as autoras identificam que (i) indivíduos pouco escolarizados tendem a inibir o emprego de artigos definidos, enquanto os indivíduos com escolarização completa tendem a espelhar a norma de referência, com favorecimento para o emprego de definidos; (ii) a ausência de definidos diante de possessivos é mais saliente em sintagmas nominais, enquanto os sintagmas preposicionais tendem a se configurar como o contexto morfossintático que mais favorece o emprego de definidos diante de possessivos; (iii) sintagmas em função de predicativo, adjunto adverbiais e oblíquos tendem a ocorrer sem a marcação de artigos, enquanto as demais funções favorecem seu emprego; (iv) posses com valor semântico mais referenciais – relativas ao corpo, às relações pessoais mais próximas e às relações de parentesco tendem a inibir o emprego de definidos diante de possessivos. O contato entre o português e o crioulo forro não foi uma condição relevante estatisticamente para a ausência de definidos na norma urbana são-tomense, por conta de limitações na configuração da amostra.

Azevedo (2025), em um estudo sobre o emprego de artigo definido diante de possessivos na norma urbana de Luanda, Angola, constatou que há na comunidade uma regra variável, com favorecimento para o emprego de artigos definidos no contexto morfossintático investigado. Os resultados vão na contramão da hipótese postulada pelo autor, que partiu do princípio de que se verificariam índices baixos de empregos de definidos diante de possessivos na comunidade, também em função de um processo de interferência das línguas bantu que convivem com o português em Angola – quimbundo, umbundo e kikongo. Apesar de essas línguas possuírem uma estrutura análoga ao que seria o artigo definido, restrições de natureza morfológica, sintática e semântica muito distintas das do português atuam no emprego dos artigos, o que dificultaria, para um falante efetivamente bilingue, o estabelecimento de analogias entre o português e as línguas bantu de Angola.

No âmbito das restrições que atuam para inibir o emprego de artigo definido na norma de Luanda, Azevedo (2025) identificou que (i) se houvesse a possibilidade de contração entre artigo e preposição, a ocorrência de definidos era favorecida ; (ii) nomes no masculino inibem mais a ocorrência de artigos definidos do que nomes no feminino; (iii) nomes no singular tendem a inibir a ocorrência de artigos definidos; (iv) nomes concretos favorecem a ausência do definido, enquanto nomes mais abstratos tendem a figurar marcados pelo artigo; (v) sintagmas em função de predicativo, sujeito e adjunto adverbial tendem a inibir a ocorrência de artigos definidos; (vi) indivíduos que se identificam como falantes nativos de línguas Bantu, com o português como L2, tendem a inibir o emprego de artigos e (vii) os falantes jovens e os indivíduos adultos tendem a não empregar o artigo definido diante de possessivos, em um processo de afastamento à norma de referência, o português europeu, norma em que o emprego de definidos diante de possessivos é uma regra categórica.

Os trabalhos resenhados nesta seção atestam a complexidade no uso variável de artigo definido diante de possessivos. De uma forma geral, atesta-se uma grande variabilidade no emprego dos definidos, variação essa correlacionada a aspectos estruturais e sociais. Contudo, para melhor delimitar a hipótese de investigação sobre a norma urbana de Maputo, é importante entender como a marcação de definitude se configura em contextos de contato entre línguas europeias, especificamente o português, e línguas autóctones do continente africano. Tais reflexões serão apresentadas a seguir.

3. Considerações sobre a noção de definitude em contextos de contato entre línguas: o português e as línguas autóctones africanas

No âmbito da fundamentação da hipótese de investigação, é importante retomar reflexões sobre como o contato entre o português e as línguas autóctones do continente africano afeta a codificação de definitude na gramática das variedades do português faladas no continente**.** Retomam-se aqui as considerações de Janson (1984), sobre o papel das línguas lexificadoras da Europa Ocidental na formação de crioulos de base lexical românica, e a descrição de Lucchesi (1993) sobre os crioulos de base lexical portuguesa falados em São Tomé e em Cabo Verde.

Apesar de a variedade do português falada em Moçambique não ser derivada de processos de crioulização e nem ter atestado em sua história processos crioulizantes, a compreensão de como a gramática da definitude se articula em situações de contato abrupto pode fornecer pistas sobre como o mesmo processo se constitui em variedades constituídas a partir de processos de reestruturação parcial de línguas europeias em contextos de colonização (Holm, 2003), como o caso das variedades brasileira e das normas em Angola e em Moçambique.

Janson (1984) tece reflexões sobre o processo de formação do sistema de artigos em crioulos de base lexical portuguesa, francesa e inglesa, com base no pressuposto de que os crioulos não possuem comportamentos convergentes no que se refere ao estabelecimento dos traços de definitude e especificidade. Assim, é possível situar, de um lado, os crioulos de base lexical portuguesa e francesa e, de outro, os crioulos de base lexical inglesa. As características semelhantes entre os crioulos de base portuguesa e francesa são as mesmas encontradas em todas as línguas da Europa Ocidental, o que leva o autor a defender que a formação do sistema de artigo nos crioulos está ligada à mudança histórica da língua lexificadora.

No que se refere especificamente aos crioulos de base lexical portuguesa descritos pelo autor – o Papiamento (falado em ilhas do Caribe) e os crioulos falados nas ilhas do Golfo da Guiné (São Tomé, Príncipe e Ano-Bom) – o trabalho destaca que o uso do artigo indefinido nos referidos crioulos é similar, sendo seu emprego obrigatório para a primeira menção de pessoas e objetos no discurso. O emprego do artigo definido nesses crioulos também é semelhante: nopapiamento, há um artigo definido opcional (*e*) (com variação morfológica *es*) e sintagmas nominais sem artigo possuem referência definida. Nos crioulos da ilha do Golfo da Guiné, Janson (1984), a partir da descrição apresentada por Ferraz (1983), afirma que a partícula *se* desempenha a função de pronome demonstrativo e também pode exercer a função de um artigo.

O autor afirma que todos os crioulos analisados têm artigo indefinido no singular e no plural, assim como todos os crioulos por ele descritos apresentam uma partícula que pode funcionar como artigo definido. Contudo, Janson (1984) defende que os artigos não são características universais das línguas, pois

> o padrão com artigo definido no singular e no plural e um artigo indefinido no singular, derivado do numeral *um*, é típico das línguas da Europa Ocidental, mas não é frequente em outros lugares. Na minha opinião, o fato de ser este o padrão dominante nos crioulos é uma evidência muito forte contra qualquer hipótese inatista ou de tendências universais. Na verdade, parece um pouco presunçoso sugerir que um item encontrado principalmente na sintaxe europeia esteja embutido nos genes da humanidade. (Janson, 1984, p. 315)[2]

Lucchesi (1993) observa o sistema de artigos nos crioulos de base lexical portuguesa de São Tomé e Cabo Verde. Em sua análise, o autor destaca que há diferenças entre o sistema de referência das línguas crioulas e a língua lexificadora (o português), em função de fatores específicos que se articularam para a formação desses crioulos e a implicações gramaticais inerentes aos processos de contato linguístico.

O autor pôde notar que o emprego do artigo nos crioulos de Cabo Verde e São Tomé consiste em ser limitado e irregular, uma vez que o sistema de referência é pautado em uma única forma: o artigo indefinido. Quando um SN é introduzido no discurso por meio de um artigo indefinido, sua referência é desconhecida pelo ouvinte, mas o SN já possui uma característica 'específica', dado que esse item foi retirado de um conjunto maior. Com efeito, não há a necessidade do emprego do artigo em ocorrências subsequentes. Lucchesi (1993) constata que o emprego de uma estrutura à semelhança do artigo definido nessas línguas é opcional, uma vez que seu emprego desempenha um papel discursivo/pragmático, indicando ênfase no conteúdo referencial do SN ou apenas reforçando a noção de 'definitude'.

A erosão do sistema de artigos definidos do português nos crioulos pode estar ligada às propriedades morfossintáticas/fonológicas do artigo definido na língua lexificadora. Isso porque, ao contrário das demais línguas românicas – em que o artigo definido possui a estrutura silábica CV (como, por exemplo, no espanhol: *lo/la*; no francês: *le/la*) – , no português, a estrutura silábica do artigo definido é composta somente pelo núcleo vocálico (*o/a*), o que torna essa partícula suscetível à atuação de processos de sândi, com a fusão desse elemento a outras estruturas, principalmente preposições.. Em decorrência disso, o artigo definido tende a não suportar as mudanças que ocorrem durante o processo de desenvolvimento dos crioulos/pidgins.

Por associação, a configuração do sistema de marcação por definitude a partir de artigos também seria explicada por princípios análogos em situações de contato que levam à reestruturação parcial

---

[2] No original, "*Furthermore, the pattern with definite article in both the singular and the plural, and an indefinite article in the sigular, derived from the numeral one, is typical for Western European languages, but not very frequent elsewhere.In my opinion, the fact that this is the dominant pattern in the creoles is very strong evidence against any global hypothesis of innateness that an item found mainly in European syntax is embedded in the genes of mankind.*"

da gramática do português, como o que se verifica na formação das variedades do português faladas no Brasil, em Angola e em Moçambique. A variação no emprego dos artigos definidos nessas três variedades, na contramão do que acontece na variedade europeia - em que o uso de definidos é categórico nesse contexto -, pode ser justificada como um resquício do processo de reestruturação parcial da gramática da definitude do português europeu no processo de implementação / consolidação da língua no contexto colonial.

4. Aspectos sociolinguísticos do português em Moçambique

A República de Moçambique é um país litorâneo localizado no sudeste da África. Divide-se em onze províncias (Cabo Delgado, Gaza, Inhambane, Manica, Maputo, cidade de Maputo, Nampula, Niassa, Sofala, Tete e Zambézia) que se estendem por 801.537 km². No país a língua oficial é o português, que está em contato com pelo menos 20 línguas *bantu*, além de línguas estrangeiras como Inglês, Árabe, Hindi, Guajaradi e Urdu, e também línguas asiáticas (Firmino, 2002; Lopes, 1998 *apud* Chimbutane, 2018). Na tabela 2, a seguir, explicita-se – com base nos dados dos censos de 1980, 1997, 2007 e 2017 – a distribuição da população moçambicana consoante as línguas faladas pelos indivíduos.

Tabela 2: Avaliação dos Censos de 1980, 1997, 2007 e 2017

Ano	Falantes de língua Bantu como L1	Falantes de português	Falantes de português L1
1980	98,8%	24,4%	1,2%
1997	93%	39,8%	6,5%
2007	85,2%	50,4%	10,7%
2017	79,2%	58,1%	16,5%

Fonte: Chimbutane, 2018

Os dados censitários revelam que o número de falantes de português como L1 aumentou gradativamente e que o número de falantes de língua bantu vem diminuindo. Em 1980, 98,8% da população declarava ter uma língua bantu como língua materna, enquanto em 1977 esse número era de 93% e, em 2007, de 85,2%. Por outro lado, o número de falantes de português tem crescido ao longo dos anos. Em 1980, 24,4% da população afirmava falar português, número que subiu para 39,8% em 1997 e chegou a 50,4% em 2007, incluindo aqueles que tinham o português como primeira língua. O percentual de pessoas que têm o português como língua materna também está em ascensão. Em 1980, apenas 1,2% da população dizia ter o português como primeira língua; em 1997, essa porcentagem subiu para 6,5%, e em 2007, para 10,7%, evidenciando a tendência de crescimento dos falantes nativos de português. O Censo de 2017 mostra que 16,5% dos moçambicanos falam português como língua materna, enquanto os falantes de português como segunda língua representam 58,1% da população do país. (Chimbutane, 2018).

No que se refere especificamente a Maputo, capital de Moçambique e local de recolha dos inquéritos investigados neste trabalho, o mapeamento realizado pelo Instituto Nacional de Estatística (INE, 2017) revela que a população aproximada da capital é de 28 milhões de habitantes, o que a alça à cidade mais populosa do país. Maputo é considerada uma área urbana, na qual o português é a principal língua de interação.

Esse panorama não é uniforme pelas demais províncias do país. A difusão do português nas zonas mais rurais é fragmentada, já que nessas áreas as línguas da família bantu ainda constituem as línguas maternas das populações locais. Pissurno (2018) afirma que "a população da zona rural continua tendo o Português praticamente como uma língua estrangeira, de acesso restrito ao meio escolar" (2018, p. 82). As razões para a diferença entre a capital e as demais provinciais do país estão no processo histórico de formação de Moçambique.

Com a chegada de Vasco da Gama à costa do país em 1498, Moçambique é inserido no vasto império português. Apesar de a colonização portuguesa ser intensa na costa oeste do continente africano e na Ásia, os colonizadores à época não demonstraram grande interesse em ocupar o território, inserindo-o na estrutura imperial como um posto comercial, fonte de recursos naturais e pessoas escravizadas (Chimbutane, 2018). Como resultado da falta de uma política incisiva, o contato entre o português e as línguas locais foi bastante rarefeito nos momentos iniciais da colonização.

As relações entre Portugal e Moçambique se mantêm limitadas nos primeiros dois séculos após a chegada dos portugueses. O país ocupava uma "posição marginal" (Gonçalves, 2001, p. 977) no Império Português, principalmente quando em comparação às outras regiões do império. Esse panorama só começa a mudar no final do século XIX, com a delimitação das fronteiras e a escolha de Maputo (Lourenço Marques, à época) como capital. O processo de efetiva ocupação não foi pacífico e nem repentino, mas sim marcado por campanhas militares e conflitos entre portugueses e resistências locais.

A consolidação do processo de colonização, com imigração de colonos, e consequente fomento à educação, se dá a partir de 1918, com o fim das campanhas militares para ocupação efetiva do país, momento em que Portugal considerou o território, por fim, dominado. O número de imigrantes portugueses ainda era bastante baixo, mas Firmino (2010) destaca que a presença desses colonos acabou causando uma "competição desleal no trabalho" (p. 6) e que os moçambicanos tinham de provar uma qualificação que esbarrava justamente no conhecimento do português.

Desse modo, o português assume a função de mobilização social, função que ganha importância no momento em que a língua foi incorporada às chamadas "Escolas Indígenas", instituições com função de disseminar o português para os nativos. Concomitantemente, as línguas locais passaram a ser proibidas nessas instituições escolares, uma estratégia para enfraquecer as línguas autóctones, restringindo-as ao convívio familiar.

Os movimentos pela independência começam em 1962, com a fundação da Frente de Libertação de Moçambique (FRELIMO). Curiosamente, o português foi escolhido como língua franca para

mobilização rumo à independência. Uma escolha, à primeira vista, contraditória, mas que se justifica – na medida em que a opção por uma das línguas locais para a representação do ideal de nacionalidade poderia colocar uma região/povo em relação de proeminência aos demais, o que seria um entrave na união rumo à independência. (Firmino, 2010, p. 9).

A independência se consolida em 1975, depois de mais de uma década de guerra, e o português é reconhecido como língua oficial de Moçambique. O discurso ideológico/institucional se afirma em direção à escolha da língua do colonizador como língua oficial, já que o português "quebrava as barreiras das línguas maternas" (Firmino, 2010). Consequentemente, ampliam-se políticas linguísticas que promoviam o uso do português, por meio de propagandas e discursos do governo, e, principalmente, através de campanhas de educação em massa.

Só com a Constituição de 1990 se estabelece um movimento contrário à posição hegemônica do português. O documento constitucional define que "o Estado devia promover o desenvolvimento das línguas africanas na vida pública, inclusive na educação." (Chimbutane, 2018, p. 101). Essa posição política leva à legitimação da educação bilíngue, e ao reforço do contato entre as línguas, permitindo, cada vez mais, que a variedade do português em Moçambique fosse modelada pela interação entre o português e as línguas locais.

Contudo, as evidências históricas e sociolinguísticas disponíveis não permitem afirmar categoricamente que estamos diante de um "Português Moçambicano". Para Firmino (2010, p. 22),

> o português em Moçambique pode ser visto como um *continuum* que oscila desde as formas do mau português (pejorativamente chamado pretoguês) até às formas mais próximas do português europeu, cujo modelo europeu ainda é considerado como o padrão e, teoricamente, continua a ser difundido nas escolas.

Gonçalves (2021) oferece evidências adicionais para a interpretação do português em Moçambique como uma variedade ainda em construção, por destacar que "é prematuro, com base nas produções dos falantes, observar se os dados efetivamente refletem a emergência de uma variedade moçambicana". É possível identificar alguns traços característicos dessa variedade em construção – principalmente no que se refere a incorporações e adaptações de itens lexicais, à configuração morfossintática e ao sistema fonológico –, estando o português em Moçambique no estágio de "uma variedade típica de um contexto de língua europeia como L2".

5. Aspectos metodológicos

Os dados relativos à norma urbana de Maputo provêm da amostra MOÇAMBIQUE-Port[3] . A análise conta com 627 ocorrências de sintagmas nominais e preposicionais em que a marcação de posse é realizada por pronome possessivo. As ocorrências foram extraídas de uma amostra composta por 18 entrevistas sociolinguísticas, gravadas em Maputo em 2016, estratificadas de acordo com as variáveis sociolinguísticas sexo (masculino /feminino), escolaridade (níveis básico/intermediário

[3] Disponível em: https://corporaport.letras.ufrj.br/.

e superior) e faixa etária (entre 18 e 35 anos, entre 36 e 55 anos e mais de 56 anos). Para além do controle das variáveis sociais clássicas, é possível controlar, com base na autodeclaração dos informantes, o estatuto de aquisição do português e as línguas utilizadas na comunicação.

Por se tratar de uma análise de cunho variacionista (Weinreich, Labov, Herzog, 1968; Labov, 1972), postularam-se restrições estruturais e sociais que podem condicionar a variação. As variáveis linguísticas postuladas verificam questões relacionadas à natureza do sintagma, ao tipo de posse, à função sintática do sintagma, ao tipo de verbo com que se relaciona o sintagma possessivo, à pessoa do discurso e à presença de elementos entre o possessivo e o nome. No âmbito das variáveis sociais, controlam-se as condições que figuram na estratificação da amostra, além das variáveis estatuto de aquisição do português e trânsito pelas línguas da comunidade.

No quadro 1, a seguir, é possível observar os condicionamentos analisados para verificar as condições que levam à ausência do artigo definido diante de possessivos na norma urbana de Maputo.

Quadro 1: variáveis investigadas

linguísticas	**sociais**
Tipo de sintagma	Faixa etária
Tipo de verbo	Escolaridade
Pessoa do discurso	Sexo
Semântica da posse	Estatuto do português
Preposição	Línguas de comunicação
Função na estrutura de predicação	

Fonte: Elaboração da autora

Para a operacionalização das análises estatísticas, utiliza-se o pacote de programas disponível na plataforma *RStudio,* recorrendo-se ao modelo de regressão logística binomial na análise multivariada. O valor de aplicação considerado corresponde à ausência de artigos definidos diante de possessivos. Conforme discussão apresentada em seções anteriores, as línguas em contato com o Português no continente africano, de um modo geral, se caracterizam por um sistema muito particular de marcação da definitude. No caso das línguas bantu de Moçambique, conforme Atanásio (2002), não há uma estrutura análoga ao artigo definido no português, sendo as noções de especificidade/delimitação codificadas a partir de afixos que se adjungem ao nome. Assim sendo, espera-se que a norma urbana de Maputo reflita a tensão entre as gramáticas das línguas bantu e da norma de referência, no sentido em que quanto maior o acesso a essa norma (via a escolarização) e quanto menor o trânsito desses falantes de português pelas línguas autóctones, maiores as taxas de emprego do artigo definido diante de possessivos.

6. Resultados

A tabela 3, abaixo, apresenta os resultados em termos percentuais da ocorrência de artigos definidos diante de possessivos nos dados da norma urbana de Maputo.

Tabela 3: Distribuição das variantes

variantes	APL/T
Presença de artigo definido diante de possessivo	381/627 = 61%
Ausência de artigo definido diante de possessivo	246/627 = 39%

Fonte: Elaboração da autora

Como é possível observar na tabela acima, os dados de emprego de artigo definido diante de possessivo na norma urbana de Maputo se distribuem de maneira que a variável presença de artigo definido diante de possessivo seja mais produtiva nos dados, concentrando 61% das ocorrências, enquanto a ausência de definidos diante de possessivos corresponde à 39 % das ocorrências. Os resultados contrariam a hipótese inicialmente postulada, já que se esperava que a variável ausência de definidos fosse a norma na comunidade, hipótese essa que levou a selecionar a variável como valor de aplicação na análise probabilística.

Embora a distribuição das variantes contrarie a expectativa inicial, optou-se por manter, na análise probabilística, a ausência do definido diante do possessivo como variável de referência. Na análise, o programa indicou como restrições estatisticamente relevantes para a ausência do artigo definido as variáveis *escolaridade, tipo de sintagma, tipo de posse e faixa etária.*

Na tabela 4, a seguir, apresentam-se o *p-value*[4], a distribuição percentual, os pesos relativos e os *logodds*[5] para a variável escolaridade, a mais saliente na análise probabilística na investigação das condições que levam à ausência de definidos diante de possessivos na norma urbana de Maputo.

Tabela 4: Efeito da variável ***escolaridade***

variante	Apl/T	PR	*logodds*
básica	104/209 = 49,76%	.611	0.450
intermediária	85/206 = 41,26%	.569	0.278
superior	57/212 = 26,88%	.325	-0.729
p-value	9.27e-07		

Fonte: Elaboração da autora

[4] *P-value*, em inglês, é uma métrica estatística utilizada para avaliar se os padrões observados em um dado processo de variação têm probabilidade de serem reais ou se podem ser fruto do acaso. A relevância de cada variável investigada se dá pelo valor numérico expresso: valores maiores do que 0.05 indicam que uma determinada variável não tem efeito significativo; valores menores do que 0.05 indicam que uma determinada variável independente é estatisticamente significativa na investigação da variável linguística estudada.

[5] *Logodds* são medidas de probabilidade que indicam a estimativa de ocorrência para um determinado evento. (Devore, 2016). O programa *Rbrul* indica *logodds* positivos e negativos. Se positivo, a variante favorece a aplicação do fenômeno; se negativo, há o seu desfavorecimento.

Os resultados da tabela 4, acima, revelam uma estreita relação entre acesso à norma de referência e emprego de definidos diante de possessivos na norma de Maputo: quanto menos escolarizado o indivíduo, menor seu acesso à norma de prestígio da comunidade – o português europeu – e, por efeito, maiores são o índice percentual e o peso relativo para a variante ausência de definido. Os informantes com nível básico de instrução são os que mais tendem a não marcar a definitude por artigo (.611 de peso relativo). Na medida em que esses indivíduos progridem no acesso à norma de referência via escolarização, menores são os índices percentuais e probabilísticos para a ausência do definido diante de possessivo.

Os resultados para a segunda variável relevante estatisticamente para a ausência do artigo definido na norma de Maputo é uma restrição de natureza morfossintática, relativa à natureza do sintagma em figura o possessivo. Os resultados para a variável tipo de sintagma estão expressos na tabela 5, a seguir.

Tabela 5: Efeito da variável ***tipo de sintagma***

variante	Apl/T	PR	*logodds*
Sintagma Nominal (SN)	200/401 = 49,87%	**.636**	**0.558**
Sintagma Preposicional (SPrep)	57/212 = 26,88%	.364	-0.558
p-value	9.96e-07		

Fonte: Elaboração da autora

Na tabela 5, acima, fica nítida a diferença entre sintagmas nominais e sintagmas preposicionais: o SN é o contexto morfossintático em que há maior variação entre as variantes ausência e presença do definido, com um ligeiro favorecimento à ausência do artigo (49,87% de frequência e .636 de peso relativo). Já o SPrep é o contexto mais prototípico para o emprego do artigo definido diante de possessivos de definido, desfavorecendo o valor de aplicação *ausência de definidos* (.364 de peso relativo). Os exemplos entre (09) e (12), a seguir, ilustram a variação no emprego do definido diante do possesivo consoante o tipo de sintagma.

(09) É melhor ficar ***no meu lado*** (PMO-A-1-h)

(10) Eu insisto muito nisso ***em meus meninos*** (PMO-C-3-f)

(11) Moram ***meu irmão*** (o casado), mais uns dois que não casaram (PMO-A-2-f)

(12) Isso me criou formas de conseguir alcançar ***o meu objetivo*** (PMO-C-2-h)

Na tabela 6, a seguir, apresentam-se a distribuição e os valores estatísticos para a variável tipo de posse, uma restrição de natureza semântica.

Tabela 6: Efeito da variável ***tipo de posse***

variante	Apl/T	PR	*logodds*
Parte do corpo	2/3 = 66,66%	.658	0.656
Relações pessoais	41/74 = 55,40%	.628	0.526
Relações de parentesco	125/263 = 47,52%	.505	0.019
Posse alienável	25/74 = 33,78%	.441	-0.238
Posse abstrata	53/213 = 24,88%	.276	-0.293
p-value	2.19e-06		

Fonte: Elaboração da autora

Os resultados expressos na tabela 6, acima, revelam que noções mais referenciais do ponto de vista semântico, isto é, que delimitam referentes mais acessíveis do ponto de vista da determinação, são os contextos semânticos que mais favorecem à ausência do definido diante do possessivo. Posses relacionados ao corpo (.658) e às relações sociais pessoais (.628) são as variantes que mais propiciam a ocorrência de possessivos sem a marcação de definitude por artigo. No âmbito das relações de parentesco, o peso relativo que essa é uma relação semântica quase no limite da neutralidade (.505). Os resultados precisam ser tratados com cautela, dadas questões de distribuição dos dados.

Apesar das ressalvas, os resultados são coerentes com outras descrições em que o controle da semântica da posse (p. ex., Oliveira e Silva, 1982) foi uma restrição relevante para o emprego de artigos definidos diante de possessivos. Os exemplos entre (13) e (15), abaixo, ilustram contextos em que a ausência do definido diante de possesivo é favorecida em função do tipo de posse.

- **Parte do corpo**

(13) Você tá preocupado pra subir na chapa, alguém mete ***sua mão*** no bolso (PMO-A-1-h)

- **Relações pessoais**

(14) Em vez de (...) ir sair ***com meu marido*** (PMO-A-3-f)

- **Relações de parentesco**

(15) E ***meu pai*** não queria que eu estudasse à noite (PMO-A-3-f)

No que se refere à última variável selecionada, a faixa etária dos informantes, os resultados – expressos na tabela 7, a seguir – sugerem uma curva de mudança linguística na comunidade.

Tabela 7: Efeito da variável ***faixa etária***

variante	Apl/T	PR	*logodds*
18a.-35a.	111/227 = 48,89%	.592	0.373
36a.-55a.	35/103 = 33,98%	.518	0.073
Mais de 56a.	100/297 = 33,67%	.390	-0.446
p-value	0.0136		

Fonte: Elaboração da autora

Os resultados expressos na tabela acima indicam que os indivíduos mais jovens são os que mais favorecem possessivos sem marcação por artigo definido, padrão que decresce na medida em que se avança para as faixas mais altas da comunidade. A curva etária sugere que os jovens são responsáveis por implementar a variante ausência de definidos diante de possessivos. Tal curva vai de encontro à hipótese inicial, pois se esperava que os mais jovens fossem responsáveis por implementar o uso de artigo definido no contexto analisado, processo que ocorre em outras variedades parcialmente reestruturadas do Português (Baxter; Lopes, 2006, 2009).

As variáveis que poderiam revelar o efeito do contato entre línguas na configuração da norma urbana do português em Moçambique – o estatuto de aquisição do português e as línguas de comunicação – não foram selecionadas na análise probabilística, o que limita a correlação dos resultados a um efeito do processo de contato linguístico. Entretanto, apresenta-se a distribuição percentual das variantes controladas nessas duas variáveis, uma vez que a distribuição dos dados sugere padrões de comportamento da comunidade.

Tabela 8: Estatuto de aquisição do português

variante	Apl/T
Português como L1	140/392 = 35,71%
Português como L2	106/235 = 45,10%

Fonte: Elaboração da autora

Tabela 9: Línguas de comunicação

variante	Apl/T
Só Português	19/87 = 21.83%
Tanto Português quanto línguas locais	163/399 = 40,85%
Mais línguas locais do que o Português	64/141 =45,39%

Fonte: Elaboração da autora

Os resultados expressos nas tabelas 8 e 9, acima, indicam que quanto maior o trânsito do falante urbano de Maputo pelas línguas locais da comunidade, maior é a ocorrência de pronomes possessivos sem a marcação por artigo definido. Tal tendência é nítida no comportamento dos indivíduos que reconhecem ter o português como segunda língua (45,10% de ausência de artigo definido diante de possessivos) e no discurso dos indivíduos que se reconhecem efetivamente como bilingues, transitando entre o português e as línguas autóctones – 40,85% de ausência de definidos diante de possessivos no comportamento dos falantes que transitam igualmente entre o português e as línguas locais e 45,39% no discurso de quem se comunica mais em línguas locais do que em português. As diferenças percentuais entre as variantes controladas nessas duas variáveis podem indicar que questões relacionadas a outras condições, como distribuição dos dados e a própria configuração do *corpus*, podem mascarar o efeito das restrições que controlam o papel do contato entre línguas.

7. Considerações finais

A discussão encaminhada neste artigo permite afirmar que o comportamento do artigo definido diante de pronomes possessivos na variedade do português falada em Maputo, capital da República de Moçambique, se caracteriza como uma regra variável. Essa tendência é distinta da observada no português europeu, em que o emprego do artigo definido frente a pronomes possessivos se caracteriza como uma regra categórica (com violação em contextos discursivo-pragmáticos particulares). A variação do artigo definido na norma urbana de Maputo apresenta-se com uma configuração mais próxima às tendências observadas em normas do português brasileiro, em que o uso do artigo definido diante de possessivo é uma regra variável.

No que tange à regra variável, controlaram-se as variáveis linguísticas – tipo de sintagma, tipo de preposição, tipo de verbo, pessoa do discurso, tipo de posse, função sintática do sintagma em que figura o pronome possessivo e presença de elemento interveniente entre o possessivo e o nome –, e extralinguísticas – sexo, faixa etária, escolaridade, estatuto do português e línguas de comunicação – que poderiam condicionar a ausência do definido diante de possessivo, valor de referência para a análise estatística por conta da hipótese de interferência de traços das línguas bantu da comunidade.

A análise estatística revelou que, no que se refere às condições que favorecem a ausência do artigo definido diante de possessivo na norma de Maputo, são relevantes:

(i) a escolaridade do indivíduo: indivíduos pouco escolarizados são os que mais favorecem o emprego de sintagmas com possessivos sem a marcação de definitude por artigo, enquanto os indivíduos plenamente escolarizados são os que mais se aproximam da norma de referência, ao favorecerem o emprego do artigo definido;

(ii) a natureza do sintagma em que figura o possessivo: sintagmas nominais são os contextos morfossintáticos preferenciais para a variação, com favorecimento para a ausência do artigo definido. Os sintagmas preposicionais, por outro lado, bloqueiam à variação, com favorecimento para o emprego de artigos definidos;

(iii) a natureza semântica da posse: contexto de posse mais referenciais, associados ao corpo e as relações sociais/pessoais são os contextos semânticos que mais propiciam possessivos sem marcação de definitude por artigo, enquanto posses menos referenciais (genéricas e/ou abstratas) favorecem o emprego do definido;

(iv) os informantes mais jovens tendem a empregar menos o artigo definido do que os informantes mais velhos, em uma curva etária que sugere um processo de mudança em progresso no sentido de perda do definido diante de possessivos. Contudo, os limites da amostra não permitem atestar se de fato ocorre um processo de mudança na comunidade.

As variáveis que controlam o efeito do contato entre línguas não foram indicadas como relevantes estatisticamente para ausência de definidos diante de possessivos na norma urbana do português falada em Moçambique. No entanto, a distribuição percentual das variantes sugere que os falantes que se identificam como nativos de português e que o utilizam mais recorrentemente em suas interações cotidianas mais utilizam o definido diante de possessivo do que os indivíduos que reconhecem ter o português como segunda língua e/ou transitam mais consistentemente em contextos multilíngues.

Referências

ATANÁSIO, Nicolau. *Ausência de artigo no português de Moçambique*: análise de um corpus constituído por textos de alunos do ensino básico em Nampula. 2002. 178 f. Dissertação (Mestrado em Linguística Portuguesa) – Faculdade de Letras da Universidade do Porto/Universidade Pedagógica-Moçambique, Nampula, 2002.

AZEVEDO, Matheus de Araújo. *A variação no uso de artigo definido diante de possessivos pré-nominais no português falado em Luanda-Angola*: uma análise sociolinguística. 2025. 127 f. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Estadual de Feira de Santana, Feira de Santana, 2025.

BAXTER, A.; LOPES, N. O Artigo Definido em Variação com Zero no SN de Referência Específica: Re-estruturação em Três Variedades do Português. *In*: CONGRESSO DA ALFAL – ASOCIACIÓN DE LINGÜÍSTICA Y FILOLOGÍA DE AMÉRICA LATINA, 14., 2005, Monterrey. *Anais*... Monterrey: ALFAL, 2005.

BAXTER, Alan; LOPES, Norma. Bare definite reference NPs in afro-brazilian portuguese dialect. *Revista Internacional de Lingüística Iberoamericana,* Madrid, v. 4, n. 1, pp. 55-70,2006.

BAXTER, Alan; LOPES, Norma. O artigo definido. *In*: LUCCHESI, Dante; BAXTER, Alan;RIBEIRO, Ilza (org.). *O Português Afro-Brasileiro***.** Salvador: EDUFBA, 2009. pp. 319-330.

CALLOU, Dinah; OLIVEIRA E SILVA, Giselle Machline de. O uso do artigo definido em contextos específicos. *In*: HORA, Demerval da (org.). *Diversidade Lingüística no Brasil***.** João Pessoa: Idéia, 1997. pp. 11-27.

CASTRO, Ivo. *Curso de história da língua Portuguesa*. Lisboa: Universidade Aberta, 1991.

CHIMBUTANE, Feliciano. Portuguese and african languages in Mozambique. *In*: LÓPEZ, Laura; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de. (eds.). *The Portuguese language continuum in Africa and Brazil***.** Amsterdam/Philadelphia, John Benjamins Publishing Company, 2018. pp. 89-110.

DEVORE, Jay. *Probability and Statistics for Engineering and the Sciences***.** 9th edition**.** Boston: Cengage Learning, 2016**.**

FERRAZ, Luiz Ivens. The origin and development of four creoles in the Gulf of Guinea. *In*: WOOLFORF, Ellen; WASHABAUGH, William (eds.). *The social context of creolization*. Ann Arbor: Karoma, 1983. pp. 120-125.

FIGUEIREDO, Carlos Felipe Guimarães. Uso variável do artigo definido no português da comunidade de Almoxarife, São Tomé. *Journal of Ibero-Romance Creoles*. Lisboa. v. 9, n. 1, pp. 358-389, 2019.

FIRMINO, Gregório. *A situação do português no contexto multilíngue de Moçambique*. [*S. l.*], 2010. Disponível em: http://www.fflch.usp.br/dlcv/lpot/pdf/mes/06.pdf. Acesso em: 15 set. 2021.

GOMES, Danielle Kely; CORDEIRO, Lívia Rodrigues. O uso variável de artigo definido diante de pronome possessivo na variedade urbana do Português de São Tomé. *Revista do GEL*, v.18, n.3, pp. 95-114, 2021.

GONÇALVES, Perpétua. Panorama geral do Português de Moçambique. Revue belge de philologie et d'histoire, [*s. l.*], t. 79, fasc. 3, pp. 977-990, 2001

GONÇALVES, Perpétua. Opções metodológicas na pesquisa sobre o português em Moçambique. *In*: COLÓQUIO INTERNACIONAL VARIAR – VARIAÇÃO FONOLÓGICA E SINTÁTICA NAS LÍNGUAS ROMÂNICAS, 1., 2021. *Anais* [...]. [*S. l.*: s. n.], 2021. Disponível em: https://variar.wixsite.com/variar. Acesso em: 31 maio 2021.

HOLM, John. *Languages in contact*: The partial restructuring of vernaculars. Cambridge: Cambridge University Press, 2003.

INSTITUTO NACIONAL DE ESTATÍSTICA (Moçambique). *Recenseamento Geral da População e Habitação (CENSO)*. Maputo: INE, 2017. Disponível em: http://www.ine.gov.mz/iv-censo-2017. Acesso em: 20 dez. 2025.

INVERNO, Liliana; SWOLKIEN, Dominika. O artigo definido zero em dois contextos específicos no português do Brasil e no crioulo de Cabo Verde. *Biblos***.** n.s, I, pp. 179-192, 2003.

JANSON, Tore. Articles and plural formation in creoles: change and universals. *Lingua*, v. 64, n. 4, pp. 291-323, 1984.

LABOV, William. *Padrões sociolinguísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008. (Original de 1972).

LUCCHESI, Dante. The article systems of Cape Verde and São Tomé creole portuguese: general principles and specific factors. *Journal of Pidgin and Creole Languages*, v. 8, n. 1. pp. 81-108, 1993.

MAGALHÃES, Telma Vianna. O uso de artigo definido diante de pronome possessivo em textos portugueses do século XVI a XIX. *Leitura.* Maceió, v.1, n. 47, pp. 123-143, junho de 2011.

MENON, Odete Pereira da Silva. Presença/ausência de artigos diante de nomes próprios no português do Brasil (PB). *In*: ENCONTRO DA REDE SUL LETRAS, 5., 2016, Tubarão. *Anais* [...]. Tubarão: Unisul, 2016. pp. 729-738.

OLIVEIRA E SILVA, Giselle Machline de. *Estudo da regularidade na variação dos possessivos no português do Rio de Janeiro*. 1982. 458 f. Tese (Doutorado em Linguística). Rio de Janeiro: Faculdade de Letras/ UFRJ, 1982.

OLIVEIRA E SILVA, Giselle Machline de. Realização facultativa do artigo definido diante de possessivo e de patronímico. *In*: SILVA, Giselle Machline de Oliveira; SCHERRE, Maria Marta Pereira (org.). *Padrões sociolinguísticos*: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996, pp. 119-145.

PISSURNO, Karen Cristina da Silva. O perfil multilíngue de Moçambique. *In*: BRANDÃO, Silvia Figueiredo. (org.). *Duas variedades africanas do português*: variáveis fonético-fonológicas e morfossintáticas. São Paulo: Blucher, 2018, pp. 75-91.

RINKE, Esther. A combinação de artigo definido e pronome possessivo na história do português. *Estudos de Lingüística Galega*. v. 2, pp. 121-139, 2010

SCHEI, Ane. O artigo definido frente a pronomes possessivos na literatura brasileira do século XIX. *Revista da ABRALIN*, v. 8, n. 2, pp. 15-44, jul/dez. 2009.

VIEIRA, Silvia Rodrigues; PISSURNO, Karen Cristina da Silva (org.). *Corpus Moçambique-Port.* Rio de Janeiro: Faculdade de Letras/UFRJ, 2016. Disponível em: https://corporaport.letras.ufrj.br/. Acesso em: 15 set. 2019.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução de Marcos Bagno. São Paulo: Parábola, 2006. [1968].

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72211

Construções conclusivas no português escrito: variação e gênero textual

Conclusive constructions in written Portuguese: variation and text genre

Mayra França Floret[1]

RESUMO

Este artigo investiga a variação no uso de construções conclusivas do português conectadas pelos esquemas [Preposição Demonstrativo Substantivo], [Preposição Demonstrativo] e [Preposição Substantivo Preposição Demonstrativo], à luz dos Modelos Baseados no Uso (Boyland, 2009; Bybee, 2010; Bybee; Beckner, 2010) e da Gramática de Construções Baseada no Uso (Goldberg, 1995, 2006; Diessel, 2015, 2019). Parte-se da hipótese de que o emprego dessas construções não é aleatório, mas influenciado pelas especificidades dos gêneros textuais. Para testá-la, foi constituída uma amostra composta por textos de três gêneros distintos: redações do Enem, textos de blog e artigos científicos. A análise considera a frequência e a distribuição das construções nos diferentes gêneros, bem como a variável domínio da causalidade (Sweetser, 1990) e a porção de texto em que as construções são usadas. Os dados foram analisados estatisticamente através do programa GoldVarbX (Sankoff; Tagliamonte; Smith, 2005). Os resultados indicam uma distribuição desigual das construções entre os gêneros, confirmando a hipótese colocada. Além disso, verifica-se a centralidade de um dos esquemas analisados, a forte associação de um gênero com essas construções e a correlação entre o domínio da causalidade e a porção do texto em que elas ocorrem.

PALAVRAS-CHAVE: Construções conclusivas. Gêneros textuais. Redação do Enem. Textos de blog. Artigo científico.

ABSTRACT

This article investigates variation in the use of conclusive constructions in Portuguese connected by the schemas [Preposition Demonstrative Noun], [Preposition Demonstrative], and [Preposition Noun Preposition Demonstrative], within the framework of Usage-Based Models (Boyland, 2009; Bybee, 2010; Bybee; Beckner, 2010) and Usage-Based Construction Grammar (Goldberg, 1995, 2006; Diessel, 2015, 2019). The study is based on the hypothesis that the use of these constructions is not random but is influenced by the specific characteristics of text genres. To test this hypothesis, a sample composed of texts from three distinct genres was compiled: Enem essays, blog texts, and scientific articles. The analysis examines the frequency and distribution of the constructions across genres, as well as the variable domain of causality (Sweetser, 1990) and the portion of the text in which the constructions occur. The data were statistically analyzed using the GoldVarbX program (Sankoff; Tagliamonte; Smith, 2005). The results indicate an uneven distribution of these constructions across genres, thus confirming the proposed hypothesis. In addition, the findings reveal the centrality of one of the analyzed schemas, a strong association between one genre and these constructions, and a correlation between the domain of causality and the textual position in which they occur.

KEYWORDS: Conclusive constructions. Text genres. Enem essay. Blog texts. Scientific article.

1. Introdução

Na rede de construções do português, encontram-se diversas formas responsáveis pela expressão da relação de conclusão, entendida como parte do domínio mais amplo da causalidade.

[1] Universidade do Estado do Rio de Janeiro (UERJ), mayrafloret@yahoo.com.br, https://orcid.org/0009-0000-2272-9124.

Recebido em 19/01/2026 - Aceito em 17/06/2026

Essas construções estabelecem conexões entre dois segmentos de texto, podendo ser representadas pelo esquema mais geral [Segmento 1 CONECTOR Segmento 2] ↔ Conclusão (Floret, 2022). Nesse esquema, o *slot* CONECTOR pode ser preenchido por inúmeras formas.

Neste estudo, focalizamos as construções cujo conector é formado pela combinação de preposição, demonstrativo e substantivo, como *dessa forma*. Durante a análise dos dados, observou-se que outras construções semelhantes também podem ser empregadas para estabelecer a mesma relação. Por essa razão, o objeto de estudo desta pesquisa são as construções ligadas pelo esquema central [Preposição Demonstrativo Substantivo] (Bybee, 2006, 2010) e seus associados [Preposição Demonstrativo][2], exemplificado por *com isso* e [Preposição Substantivo Preposição Demonstrativo], representado por *em consequência disso*.

A análise dessas construções a partir de uma perspectiva baseada no uso implica considerar os contextos em que elas ocorrem, em especial os gêneros textuais. Entendidos como formas textuais socialmente situadas, associadas a propósitos comunicativos específicos e a práticas discursivas recorrentes, os gêneros orientam escolhas e afetam a frequência e a distribuição de recursos linguísticos (Marcuschi, 2002). Desse modo, é fundamental que o estudo da variação de uso de construções conclusivas considere o gênero textual, uma vez que diferentes gêneros apresentam diferentes funcionalidades e possibilidades de organização discursiva.

Do ponto de vista da tradição gramatical, o enfoque é predominantemente estrutural. Nesse cenário, a conclusão é tratada como uma relação de coordenação, expressa por conjunções coordenativas que ligam um segmento a outro anterior, como *logo*, *portanto*, *por conseguinte*, *pois* e *por isso* (Cegalla, 2008; Rocha Lima, 2011; Cunha; Cintra, 2017). No entanto, os dados analisados neste estudo indicam que construções formadas por combinações de preposição, demonstrativo e substantivo, pouco ou nada contempladas nessas descrições, são frequentes no português escrito e desempenham papel relevante na expressão da relação conclusiva.

Diante desse cenário, este estudo investiga a variação no uso de construções conclusivas conectadas pelos três esquemas construcionais mencionados de acordo com o gênero textual. O objetivo central é verificar em que medida o gênero textual e, consequentemente, o registro, seja ele mais ou menos formal, influenciam o uso e a distribuição desses esquemas, bem como os domínios funcionais por eles instaurados. Parte-se da hipótese de que o emprego desses esquemas não é aleatório, sendo condicionado principalmente pelas características e funcionalidades próprias de cada gênero.

Para testar essa hipótese, constituiu-se uma amostra composta por três gêneros textuais distintos: redações do Enem, textos de blog e artigos científicos. A análise considera a frequência e a distribuição das construções nos diferentes gêneros e o funcionamento de variáveis relacionadas tanto à forma quanto ao significado dessas construções. Para os limites deste artigo, focalizamos os resultados relacionados ao domínio da causalidade (Sweetser, 1990) e sua correlação com a porção

[2] *Por isso*, apesar de preencher esse esquema, será desconsiderado nesta pesquisa, visto que já foi objeto de estudo anteriormente (Floret, 2022, 2023).

de texto em que a construção é empregada. Os dados encontrados foram analisados estatisticamente por meio do programa GoldVarbX (Sankoff; Tagliamonte; Smith, 2005).

De modo geral, os resultados apontam para uma distribuição desigual dessas construções entre os gêneros, bem como para a centralidade de um dos esquemas analisados, além de diferenças quanto às relações semânticas privilegiadas em cada gênero. Esses achados sugerem que o gênero textual se configura como um fator decisivo para a variação de uso dessas construções, e que o registro, tomado de forma isolada, se mostra apenas parcialmente relevante.

Este artigo está organizado da seguinte forma: na seção seguinte, apresentamos o referencial teórico que fundamenta a análise. Em seguida, descrevemos a forma como se constituiu a amostra e os procedimentos metodológicos adotados. Posteriormente, discutimos os resultados encontrados e, por fim, apresentamos as considerações finais.

2. Construções conclusivas, uso e gênero textual

Este trabalho se fundamenta pelos princípios dos Modelos Baseados no Uso (MBU) (Boyland, 2009; Bybee, 2010; Bybee; Beckner, 2010), perspectiva que atribui papel central à experiência do falante com a língua e considera o uso como fator determinante para a organização da gramática. Nessa abordagem, padrões linguísticos emergem a partir do uso, o que torna indispensável a consideração de fatores como frequência, contexto e distribuição dos dados. A gramática, assim, é entendida como uma estrutura dinâmica, sujeita à variação e à mudança ao longo do tempo.

Esses pressupostos se articulam aos da Gramática de Construções Baseada no Uso (GCBU) (Goldberg, 1995, 2006; Diessel, 2015, 2019), que parte de dois princípios fundamentais. O primeiro estabelece que a língua é constituída por um inventário de construções, definidas como pareamentos forma - significado (Goldberg, 2006; Traugott; Trousdale, 2013). O segundo princípio postula que essas construções não existem de forma isolada, mas se interconectam em rede, formando um *constructicon* (Goldberg, 2006; Diessel, 2015; Hoffmann, 2023).

Nessa mesma direção, Bybee (2006, 2010) propõe que cada ocorrência de uso é armazenada na memória do falante, sendo conectada a outras com base em semelhanças formais e funcionais. A recorrência de determinados padrões fortalece sua representação na memória, ao passo que variações próximas, embora não idênticas, tendem a ser organizadas como exemplares relacionados. Esse processo resulta em uma rede de exemplares, inter-relacionados entre si, ainda que se diferenciem de forma sutil.

Portanto, é nessa perspectiva teórica que se insere este estudo sobre determinadas construções conclusivas do português, entendidas como parte da relação de causalidade. De forma geral, a relação conclusiva envolve a relação entre dois segmentos discursivos, podendo ser representada pelo esquema abstrato [Segmento 1 CONECTOR Segmento 2] ↔ Conclusão. De forma prototípica, essa relação é inferencial, associada ao encadeamento entre argumentos ou premissas e uma conclusão possível a partir deles (Oliveira, 2011; Floret, 2022).

Entretanto, como demonstram diversos estudos (Kury, 1993; Matos, 2003; Marques, 2014; Marques; Pezatti, 2015; Floret, 2022), a relação conclusiva também pode codificar relações de causa e consequência. Nesse caso, o segmento introduzido pelo conector expressa um efeito decorrente de uma causa previamente apresentada, configurando uma relação menos subjetiva, uma vez que não exige, necessariamente, um raciocínio inferencial por parte do falante.

Essa distinção se relaciona com a proposta de Sweetser (1990), que descreve a causalidade como um domínio instanciado em três níveis: epistêmico, referencial (do conteúdo) e interacional. O domínio epistêmico envolve relações em que fatos ou evidências sustentam crenças ou conclusões, implicando maior grau de subjetividade (Lopes *et al*., 2001). Já o domínio referencial diz respeito a relações entre eventos ou estados de coisas verificáveis no mundo, com menor envolvimento do falante. O domínio interacional, por sua vez, refere-se a usos em que construções conclusivas funcionam como justificativas para atos de fala diretivos. Em função das características da amostra analisada neste estudo, não foram identificadas ocorrências pertencentes a esse último domínio.

No que se refere ao aspecto formal das construções conclusivas, a tradição gramatical tende a descrevê-las como casos de coordenação, em que dois segmentos são ligados por conjunções coordenativas conclusivas, como *logo*, *portanto*, *por conseguinte*, *pois* e *por isso* (Cegalla, 2008; Rocha Lima, 2011; Cunha; Cintra, 2017). Essa abordagem, no entanto, contempla apenas um conjunto restrito de formas para o preenchimento do *slot* CONECTOR no esquema conclusivo apresentado anteriormente.

Tendo em vista uma perspectiva baseada no uso, essa descrição mostra-se limitada, uma vez que a rede de construções conclusivas do português inclui também estruturas formadas por combinações de preposição, demonstrativo e substantivo, amplamente atestadas no uso, mas pouco ou nada exploradas pelas gramáticas tradicionais. Neste trabalho, considera-se que tais construções se organizam em uma rede de exemplares, na qual o esquema [Segmento 1 [Preposição Demonstrativo Substantivo] Segmento 2] ↔ Conclusão funciona como o exemplar central, dada sua maior frequência, produtividade e consequente força representacional na amostra analisada.

Além desse esquema, identificam-se ainda esquemas relacionados, que apresentam variações formais sutis, mas preservam o mesmo valor funcional. São eles: [Segmento 1 [Preposição Demonstrativo] Segmento 2] ↔ Conclusão e [Segmento 1 [Preposição Substantivo Preposição Demonstrativo] Segmento 2] ↔ Conclusão. Essas construções estão fortemente relacionadas ao exemplar central, já que compartilham características formais e funcionais, mesmo que sejam menos frequentes e menos produtivas.

Como já dito, uma abordagem baseada no uso pressupõe que a descrição da língua leve em conta não apenas aspectos formais das construções, mas também os contextos discursivos em que elas são empregadas. Nesse sentido, o gênero textual constitui uma variável central, uma vez que organiza práticas comunicativas recorrentes e orienta escolhas linguísticas associadas a finalidades comunicativas específicas. Conforme propõe Marcuschi (2002), os gêneros textuais correspondem a

textos efetivamente produzidos em situações sociais concretas, caracterizando-se por sua dinamicidade, historicidade e adequação a necessidades comunicativas. Por serem socialmente situados, os gêneros condicionam tanto a frequência quanto a distribuição de recursos linguísticos.

No contexto dos Modelos Baseados no Uso e da Gramática de Construções Baseada no Uso, considerar o gênero textual implica reconhecer que diferentes padrões construcionais podem ser mais frequentes e mais produtivos em determinados textos, em função de seus objetivos discursivos, de seu grau de formalidade e das expectativas dos interlocutores. Assim, a variação no uso de construções conclusivas não é entendida como aleatória, mas como sensível às práticas discursivas próprias de cada gênero.

Dessa forma, a amostra analisada neste estudo é composta por textos representativos de três gêneros textuais distintos: textos de blog, redações do Enem e artigos científicos. Os textos de blog emergem no contexto da expansão da internet e das plataformas digitais, configurando-se como um gênero marcado pela flexibilidade temática, pela diversidade estrutural e por um forte caráter interacional. Em geral, esses textos possibilitam uma relação mais próxima entre autor e leitor, evidenciada por marcas de subjetividade e por maior informalidade. Além disso, por serem produzidos em ambientes digitais, os blogs podem incorporar recursos multimodais e permitir a interação por meio de comentários, o que reforça seu caráter dialógico (Santos, 2003; Komesu, 2005; Floret, 2025b).

Em contraste, a redação do Enem constitui um gênero institucionalizado, definido a partir das exigências do Exame Nacional do Ensino Médio. Trata-se de um texto dissertativo-argumentativo, com estrutura relativamente rígida, tema determinado e critérios avaliativos bem estabelecidos. A produção desse gênero está diretamente relacionada ao contexto escolar e à preparação dos candidatos para o exame, o que implica normatização quanto ao uso da língua, à organização textual e à progressão argumentativa (Cavalcante; Silva, 2023; Floret, 2025b). Nesse gênero, espera-se que o candidato desenvolva uma argumentação clara e coerente, sustentada por informações, dados ou fatos legitimados, e que apresente uma proposta de intervenção relacionada ao tema abordado.

Cabe destacar, ainda, que o uso de conectores desempenha papel central na redação do Enem, especialmente no que diz respeito à articulação entre as partes do texto. A cartilha do participante do Enem (Brasil, 2024) enfatiza explicitamente a importância dos operadores argumentativos para o estabelecimento de relações lógicas, como causa, consequência e conclusão, no atendimento à competência responsável pela coesão e pela coerência textual. Essa exigência contribui para a alta recorrência de conectores nesse gênero, uma vez que o uso adequado desses elementos é diretamente associado à avaliação positiva do texto.

Por fim, os artigos científicos configuram-se como um gênero associado à divulgação do conhecimento acadêmico, caracterizado por elevado grau de formalidade, organização relativamente estável e compromisso com a exposição de dados empíricos, procedimentos metodológicos e argumentos fundamentados teoricamente. Embora também exista espaço para a subjetividade, os

artigos tendem a privilegiar relações mais objetivas e menos subjetivas, o que pode impactar tanto a escolha quanto a frequência de determinados tipos de construções conclusivas.

A consideração desses três gêneros, com características discursivas e funcionais distintas, permite investigar de que modo a variação no uso de construções conclusivas está associada não apenas a fatores estruturais, mas também às práticas comunicativas específicas em que essas construções se inserem. Na próxima seção, detalhamos a forma como a amostra composta por esses três gêneros foi organizada, bem como quais foram os procedimentos metodológicos adotados.

3. Amostra e metodologia

Como já mencionado, a hipótese central deste estudo é a de que as construções conclusivas ligadas por [Preposição Demonstrativo Substantivo], [Preposição Demonstrativo] ou [Preposição Substantivo Preposição Demonstrativo] são usadas de forma variada a depender do gênero textual considerado. Por conta disso, estabelecemos que os dados seriam coletados em uma amostra composta por gêneros textuais diferentes.

A hipótese colocada surge a partir de resultados encontrados em estudos anteriores (Biber *et al*., 1999; Liu, 2008), que correlacionaram o uso de conectores conclusivos a gêneros textuais acadêmicos. A partir disso, decidiu-se que um dos gêneros seria o artigo científico, de registro formal. No lado oposto de um continuum de formalidade, o outro gênero escolhido foram os textos extraídos de blogs. Por fim, o terceiro gênero a compor a amostra foram as redações do Enem, também de registro formal, o que nos permitiria verificar se é, de fato, o gênero textual, ou, de maneira mais genérica, o registro, que se mostra mais relevante no que diz respeito ao uso dessas construções.

Tendo em vista os três gêneros escolhidos, a escolha dos textos da amostra foi determinada pelas redações do Enem. Diferentemente dos outros gêneros, que são de fácil acesso online, as redações não são disponibilizadas ao público. No entanto, alguns candidatos que recebem nota mil decidem compartilhar seus textos com sites de notícia ou de divulgação de conteúdo, e esses textos são divulgados como forma de modelo para os futuros candidatos. Sendo assim, encontramos, no site G1[3], 101 redações que receberam nota mil, produzidas entre os anos de 2014 e 2023, que contabilizam 42 mil palavras. Logo, determinamos que os textos de blog e os artigos científicos também deveriam ser produzidos nesse mesmo intervalo de tempo.

Para a amostra de blog, selecionados oito categorias temáticas abordadas. São elas: beleza e cuidados pessoais, moda, esporte, fofoca, TV e cinema, religião, tecnologia e viagem. Em seguida, escolhemos três blogs representativos de cada assunto e, deles, extraímos três textos diferentes. Com isso, obtivemos uma amostra composta por 72 textos, totalizando 48 mil palavras. No anexo 1, estão detalhados todos os blogs e os links de acesso aos textos utilizados.

Por fim, para a amostra de artigos científicos, consideramos as grandes áreas de conhecimento reconhecidas pela CAPES, quais sejam, Linguística, Letras e Artes, Ciências exatas e da Terra, Ciências

[3] Disponível em: https://g1.globo.com/educacao/enem/. Acesso em: 26/07/2024.

biológicas, Engenharias, Ciências da saúde, Ciências agrárias, Ciências sociais e Ciências humanas. Para cada uma dessas áreas, selecionamos três textos, o que resultou em uma amostra composta por 24 artigos e 142 mil palavras. As referências aos textos da amostra encontram-se no anexo 2.

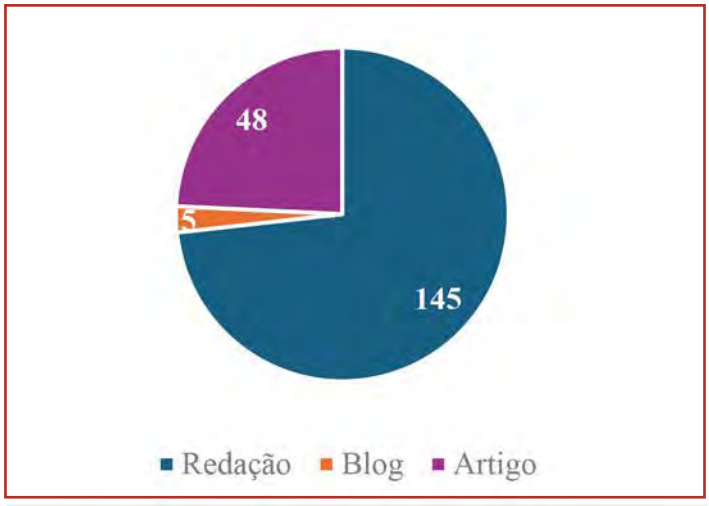
Como ficou claro, as amostras possuem extensões variadas devido às características do próprio gênero. Enquanto uma redação do Enem é limitada a uma página, um artigo científico chega a ter mais de vinte. Sendo assim, é necessário considerar o número de palavras antes de fazer afirmações mais amplas sobre a distribuição das construções nos diferentes gêneros.

Selecionada a amostra, coletamos todas as construções conclusivas ligadas por algum dos esquemas que são objeto deste estudo. Em seguida, esses dados foram analisados com base em algumas variáveis, tanto relacionadas à forma quanto ao significado e, por fim, quantificados pelo programa GoldVarbX (Sankoff; Tagliamonte; Smith, 2005). Na próxima seção, apresentamos os resultados relacionados às variáveis que geraram resultados mais relevantes.

4. Variação no uso de construções conclusivas em diferentes gêneros textuais

Através da análise de uma amostra composta por três gêneros textuais (textos de blog, redação do Enem e artigo científico), foram encontradas 198 construções conclusivas ligadas pelo esquema principal [Preposição Demonstrativo Substantivo] ou por um dos esquemas relacionados [Preposição Demonstrativo] e [Preposição Substantivo Preposição Demonstrativo]. Esses dados se distribuem de forma muito diversa entre os gêneros analisados, como mostra o gráfico 1.

Gráfico 1: Distribuição dos dados por gênero textual



Fonte: Elaboração da autora

De acordo com o gráfico, a maioria dos dados se concentra no gênero redação do Enem. As 145 construções encontradas nesses textos correspondem a 73,2% do total de dados encontrados na pesquisa. Esse resultado já era esperado, tendo em vista a associação entre o gênero e o uso de conectores, já mencionada anteriormente. Por outro lado, a frequência das construções analisadas

é muito baixa em textos de blog, representando apenas 2,5% do total de dados. Os outros 24,3% correspondem ao gênero artigo científico, que apresentou frequência mediana.

No entanto, esses resultados não podem ser interpretados de forma isolada. É necessário levar em consideração o número de palavras que compõe a amostra de cada gênero, já que os três possuem extensões muito diferentes entre si. Um alto número de dados poderia ser justificado por um número muito maior de palavras na amostra de um gênero específico, por exemplo.

Uma outra questão a se considerar é se a frequência de dados também indica variabilidade dos conectores encontrados. Em outras palavras, é importante verificar se os dados encontrados para cada gênero se distribuem entre diversos conectores, ou se tendem a se concentrar em alguns poucos. Na tabela 1, organizamos o número de dados por gênero, número de palavras e de conectores diferentes encontrados.

Tabela 1: Gênero, número de palavras e conectores encontrados

	Número de palavras	Número de dados	Conectores diferentes encontrados
Redação	42 mil	145	18
Blog	48 mil	5	4
Artigo	132 mil	48	10

Fonte: Elaboração da autora

A tabela 1 confirma a correlação entre o gênero redação do Enem e o uso de conectores. Mesmo em uma amostra menor (42 mil palavras), o número de construções encontradas supera com vantagem os demais gêneros. Por outro lado, o baixíssimo número de dados encontrados em textos de blog se distribui em uma amostra maior do que a de redação, composta por 48 mil palavras.

Já a amostra de artigos científicos, que tem mais do que o triplo de palavras em relação às redações (132 mil), possui cerca de um terço dos dados encontrados nelas (48), o que reforça a baixa frequência dessas construções no gênero. Em termos de proporção (número de dados/número de palavras), obtivemos os seguintes resultados, em ordem decrescente: redação do Enem (0,34%), artigo científico (0,03%) e textos de blog (0,01%). Portanto, esses resultados mostram, mais uma vez, que a redação do Enem é um ambiente linguístico favorável ao uso de conectores, em especial, os conclusivos analisados.

Ademais, a tabela 1 também mostra que muitos conectores diferentes foram encontrados na amostra. Apesar disso, a maioria dos dados tende a se concentrar em alguns conectores específicos. A tabela 2 mostra a distribuição dos dados por esquema e por conector.

Tabela 2: Distribuição por esquema e por conector

Conector	Redação Enem	Blog	Artigo
[Preposição Demonstrativo Substantivo]	**135**	**3**	**39**
Dessa forma	48	1	17
Desse modo	37	-	6
Dessa maneira	19	-	2
Nesse sentido	13	1	11
Nesse contexto	5	-	-
Nessa perspectiva	3	-	-
Nesse âmbito	2	-	-
Nesse cenário	2	-	1
Nesse viés	2	-	-
Por essa razão	2	-	1
Por esse motivo	1	1	1
Diante desse fato	1	-	-
[Preposição Substantivo Preposição Demonstrativo]	**5**	**0**	**1**
Por consequência disso	2	-	-
Em decorrência disso	1	-	-
Em consequência disso	1	-	-
Em vista disso	1	-	-
Em função disso	-	-	1
[Preposição Demonstrativo]	**5**	**2**	**8**
Com isso	4	2	7
Diante disso	1	-	-
Mediante isso	-	-	1
TOTAL	**145**	**5**	**48**

Fonte: Elaboração da autora

Como mostra a tabela 2, o esquema que abarca a grande maioria dos dados é [Preposição Demonstrativo Substantivo], o exemplar central. O esquema representa 93,1% dos dados de redações do Enem, 81,2% dos dados de artigos científicos e 60% dos dados de textos de blog. Os outros dois esquemas menos frequentes são exemplares relacionados a ele, formados por diferentes combinações das mesmas categorias. Dentre eles, o esquema [Preposição Substantivo Preposição Demonstrativo] licencia mais conectores diferentes, mas [Preposição Demonstrativo] abarca um maior número de dados por conta do representante *com isso*.

Além disso, a tabela mostra que foram encontrados 20 conectores diferentes no estudo. Quase todos aparecem pelo menos uma vez no gênero redação do Enem, exceto *em função disso* e *mediante*

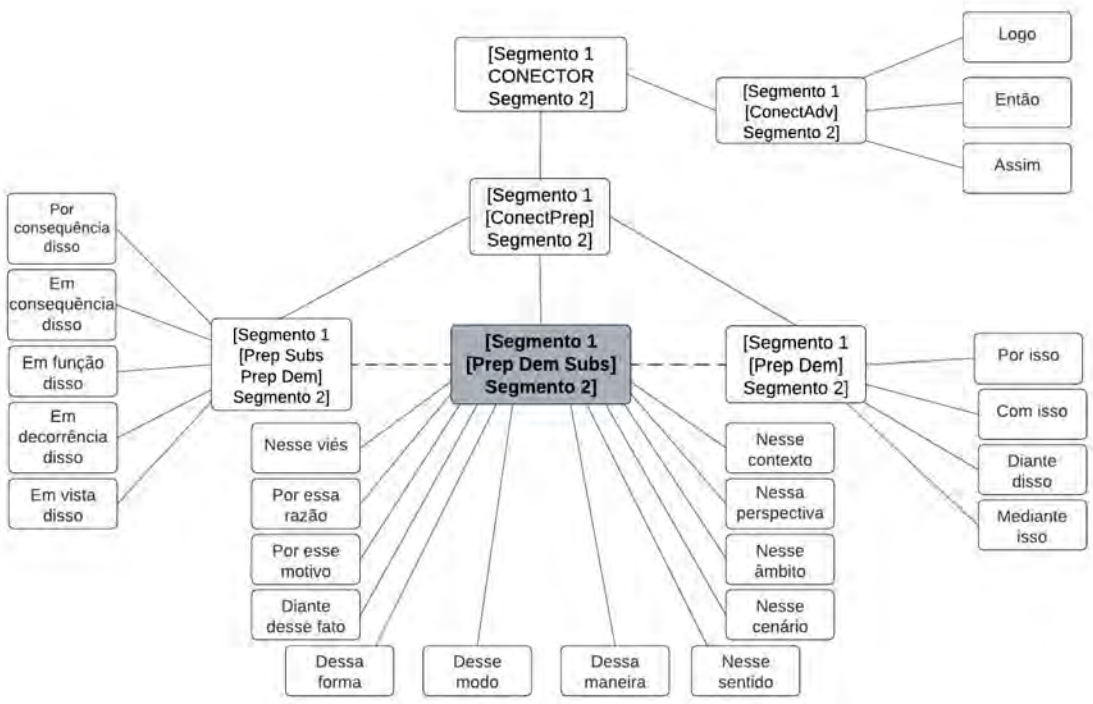
isso. O conector mais frequente é *dessa forma*, que aparece nos três gêneros e totaliza 66 dados. Em seguida, *desse modo* aparece 43 vezes, *nesse sentido* aparece 25 vezes e *dessa maneira* aparece 21 vezes. É importante mencionar novamente o conector *com isso*, que, mesmo não sendo um dos mais frequentes (13 ocorrências), registrou frequência mais elevada em artigos científicos do que em redações do Enem, apesar de haver uma grande diferença entre as quantidades de dados dos dois gêneros.

O exemplo (1) ilustra o uso do conector mais frequente do estudo.

> (1) A infância é uma fase de formação e aprendizagem, sendo necessário, portanto, que os bons costumes sejam cultivados. É, também, uma fase em que tudo é novo e interessante. **Dessa forma**, os produtos apresentados em comerciais inevitavelmente seduzirão meninos e meninas que, por sua vez, passarão a pautar sua felicidade naquilo que podem adquirir. (Redação do Enem, 2014).

Como já colocado anteriormente, este estudo parte do pressuposto de que nosso conhecimento linguístico pode ser capturado por uma rede de construções que estabelecem relações entre si (Goldberg, 2006; Diessel, 2015; Hoffmann, 2023). Sendo assim, na figura 1 organizamos os 20 conectores encontrados neste estudo em uma representação em rede junto a outros conectores conclusivos.

Figura 1: Rede de conectores conclusivos



Fonte: Floret, 2025a.

Na primeira linha da figura, encontra-se o esquema mais abstrato que representa as construções conclusivas ligadas por conector. A partir dele, há dois esquemas menos abstratos, já que, nesses casos, o tipo de conector é especificado, se adverbial ou preposicional. Para este estudo, focalizamos

apenas os preposicionais. Dentre as construções ligadas por conectores preposicionais, destacamos, no centro da imagem, os três esquemas estudados. O central está representado por um retângulo cinza e, ligados a ele por uma linha tracejada, estão os dois relacionados. A esses três esquemas, estão ligados os 20 conectores encontrados neste estudo, além de *por isso*, objeto de estudos anteriores (Floret, 2022, 2023). Os quatro conectores mais frequentes estão na linha mais baixa da rede.

Como é possível verificar na figura 1, apesar de diferentes, todos esses conectores compartilham semelhanças entre si e se ligam a esquemas mais abstratos, menos especificados. A figura também mostra que construções conclusivas podem ser preenchidas por diversos conectores. No entanto, ela ilustra apenas algumas das opções disponíveis na língua. Além disso, é importante ressaltar que é possível expressar essa relação por justaposição, sem o uso de conectores.

Partindo de estudos anteriores que mostraram que o uso de conectores conclusivos é mais evidente em gêneros textuais acadêmicos (Biber *et al*., 1999; Liu, 2008), esperávamos encontrar essa mesma correlação neste estudo. No entanto, o que os dados mostram é que, apesar de esses conectores serem utilizados em artigos científicos, essa correlação não é tão notória. Isso pode ser explicado pelo fato de outros conectores conclusivos serem mais usados nesse gênero, já que analisamos apenas um grupo limitado. É possível, ainda, que outros gêneros acadêmicos, como resenhas, relatórios, projetos de pesquisa, teses, entre outros, apresentem essa correlação de forma mais clara.

No que diz respeito a este estudo, algumas das variáveis analisadas apresentaram resultados mais relevantes. Uma delas foi o domínio da causalidade. Os dados encontrados foram analisados de acordo com a tripartição proposta por Sweetser (1990). A conclusão prototípica é entendida como uma relação no domínio epistêmico, já que envolve raciocínio inferencial por parte do emissor. Nesse caso, uma conclusão é tomada a partir de algum fato, premissa ou argumento. No entanto, construções conclusivas também instauram relações no domínio referencial (do conteúdo) quando relacionam causas e efeitos. Nesse domínio, há pouco ou nenhum envolvimento do emissor, visto que se relacionam fatos verificáveis no mundo real. Se comparadas, relações de domínio epistêmico são mais subjetivas do que as de domínio referencial. Por fim, o domínio interacional engloba os casos em que justificativas são atribuídas a atos de fala diretivos. Tendo em vista a natureza da amostra, não foram encontrados dados codificados nesse domínio.

A tabela 3 mostra o funcionamento dessa variável na amostra selecionada.

Tabela 3: Domínio da causalidade

	Redação	**Blog**	**Artigo**
Epistêmico	74 = 51%	4 = 80%	15 = 31,2%
Referencial	71 = 49%	1 = 20%	33 = 68,8%
TOTAL	145	5	48

Fonte: Elaboração da autora

A tabela 3 permite verificar que os três gêneros incluem relações tanto no domínio epistêmico quanto no referencial. Entretanto, há algumas tendências diferentes em cada um deles. No gênero redação do Enem, é possível verificar equilíbrio entre os dois domínios. Apesar de haver uma pequena vantagem para o domínio epistêmico, os dois tipos de relação coexistem de forma equiparada, o que pode ser explicado por características do próprio gênero.

A proposta de redação do Enem pede que o candidato escreva um texto dissertativo-argumentativo e apresente uma proposta de intervenção para lidar com a situação ilustrada. Nesse caso, ele precisa construir uma argumentação robusta, fundamentada em fatos concretos e verificáveis, o que acaba favorecendo a ocorrência de relação no domínio referencial. Porém, também é necessário que ele defenda uma tese e sugira uma forma de lidar com o problema em questão, o que, por sua vez, torna mais provável o emprego de relações mais subjetivas, no domínio epistêmico. Sendo assim, a fim de alcançar uma boa nota no exame, é fundamental que o participante produza um texto que inclua trechos mais e menos subjetivos. O excesso de subjetividade ou a falta dela certamente resultaria em uma penalização na nota (Floret, 2025a).

Nos textos extraídos de blog, encontramos poucos dados, o que acaba dificultando a identificação de tendências mais gerais. No entanto, o que encontramos está de acordo com o que se espera do próprio gênero textual. Uma das características centrais desse gênero é a aproximação do autor com o leitor, em muitos casos, funcionando como uma conversa informal entre os dois. O autor fala diretamente com o leitor que, por sua vez, tem a opção de interagir através dos comentários. Nesse caso, a ampla vantagem, apesar de poucos dados, para o domínio epistêmico, pode ser explicada pelo fato de conversas informais entre pessoas próximas exigir pouca preocupação com o estabelecimento de fatos verificáveis, e consequentemente abrir maior espaço para a expressão da subjetividade, da opinião, do pouco compromisso com o que pode ser provado ou não.

Por fim, no gênero artigo científico, observa-se predomínio de relações codificadas no domínio referencial, menos subjetivas. Essa predominância pode ser explicada pelo fato de o gênero se constituir como um espaço para a exposição de dados empíricos, informações factuais e à revisão de literatura. Ainda assim, o artigo não se limita à apresentação de informações, funcionando também como um espaço argumentativo, no qual se formulam interpretações e conclusões a partir do que foi exposto. Esse aspecto explica um índice considerável de construções no domínio epistêmico, ainda que sejam a minoria.

Sendo assim, os resultados indicam que as características dos gêneros textuais analisados desempenham papel central na distribuição de construções conclusivas que instauram diferentes domínios ao longo dos textos. A função comunicativa de cada gênero, o grau de envolvimento entre autor e leitor e o tipo de conteúdo privilegiado favorecem o uso de construções com maior ou menor grau de subjetividade.

Para além da correlação do gênero textual com o domínio da causalidade, uma outra variável analisada neste estudo foi a porção de texto em que a construção ocorre. No caso das redações,

codificamos os dados por parágrafo. Nos textos de blog, por parte do texto (primeiro, segundo e terceiro terços do texto). Nos artigos científicos, por seção. O objetivo principal de considerar essa variável era verificar se partes específicas do texto tendem a concentrar, também, relações de domínios específicos. Nos textos de blog e nos artigos científicos, essa correlação não foi observada. Em cada parte dos textos, a maioria dos dados encontrados coincidia com a relação mais frequente do gênero textual, seja ela epistêmica ou referencial.

Por sua vez, no gênero redação do Enem, verificamos que o relativo equilíbrio entre domínio referencial e epistêmico não é homogêneo ao longo do texto, mas, ao contrário, essas relações se distribuem de forma diferenciada, a depender do parágrafo considerado. A tabela 4 mostra a correlação entre parágrafo e domínio da causalidade no gênero.

Tabela 4: Parágrafo e domínio da causalidade em redações do Enem

	1	**2**	**3**	**4**
Epistêmico	13 = 56,5%	11 = 25,6%	19 = 40,4%	31 = 96,9%
Referencial	10 = 43,5%	32 = 74,4%	28 = 59,6%	1 = 3,1%
	23	43	47	32

Fonte: Floret (2025a)

De acordo com a tabela 4, construções conclusivas como as analisadas são mais frequentes nos parágrafos de desenvolvimento (2 e 3), especialmente no terceiro. Como já abordado em Floret (2025a), essa é a parte do texto em que o candidato reúne a maior quantidade de informação e as principais ideias, consolidando sua argumentação, o que fundamentará sua defesa de tese.

No entanto, a informação mais relevante da tabela consiste na distribuição dos dados por domínio e por parágrafo. Como vimos, no gênero de modo geral há equilíbrio entre domínio referencial e epistêmico, com uma pequena vantagem para o segundo. Porém, ao observar cada parágrafo individualmente, fica claro que isso se confirma apenas no primeiro, na introdução. Nesse parágrafo, o candidato indica o direcionamento da argumentação e expõe a tese que será defendida. Essa dinâmica favorece o equilíbrio entre relações mais ou menos subjetivas.

Por outro lado, os parágrafos de desenvolvimento são mais objetivos, já que são o espaço onde o candidato constrói sua argumentação. Espera-se que sejam expostos dados e informações verificáveis, o que favorece a expressão de relação de causa e consequência, com menor envolvimento do falante.

Finalmente, na conclusão, todos os dados, com exceção de um, instauram relação no domínio epistêmico, o que mostra a alta subjetividade do parágrafo. Como dito anteriormente, uma das exigências da redação do Enem é que o candidato apresente uma proposta de intervenção de acordo com o tema. Essa sugestão é pessoal, elaborada pelo candidato, e tende a ficar ao final do texto, depois da argumentação.

Sendo assim, os parágrafos que compõem a redação do Enem favorecem o uso de relações epistêmicas ou referenciais, de acordo com suas funções e particularidades. Por conta disso, construções conclusivas ligadas por [Preposição Demonstrativo Substantivo], [Preposição Demonstrativo] e [Preposição Substantivo Preposição Demonstrativo] tendem a se distribuir de forma diferenciada ao longo do texto.

De modo amplo, o estudo permitiu confirmar a hipótese de que as construções conclusivas analisadas se organizam de maneira distinta a depender do gênero textual. Em relação ao registro, não é possível fazer essa afirmação de forma genérica.

Nos textos de blog, o gênero com registro mais informal da amostra, verificou-se baixa frequência dessas construções. No entanto, os outros dois gêneros, de registro formal, apresentaram distribuição consideravelmente diferenciada. A amostra de artigos científicos, mesmo sendo a maior, registrou um baixo número de dados se comparada à de redações do Enem, a menor amostra e com alto número de dados. Sendo assim, o registro, por si só, não explica a diferença de distribuição desses termos, mas sim as características e demandas do próprio gênero textual.

5. Considerações finais

Este estudo teve como objetivo central verificar a influência do gênero textual e, consequentemente, do registro, no uso de construções conclusivas ligadas por três esquemas, que envolvem combinações de preposição, demonstrativo e, na maioria dos casos, substantivo. Partiu-se da hipótese de que o emprego dessas construções varia de acordo com o gênero textual.

Para testar a hipótese, construiu-se uma amostra composta por três gêneros textuais: redações do Enem, textos de blog e artigos científicos. A análise dos dados mostrou que o gênero exerce influência no emprego e na distribuição dessas construções nos diferentes textos, corroborando resultados encontrados em pesquisas anteriores (Biber *et al*., 1999; Liu, 2008).

De modo geral, os resultados evidenciam uma distribuição diferenciada das construções conclusivas entre os gêneros analisados, com forte concentração nas redações do Enem, frequência intermediária nos artigos científicos e ocorrência bastante reduzida nos textos de blog. Além disso, observou-se a predominância clara do esquema [Preposição Demonstrativo Substantivo], que se configura como exemplar central da rede, presente de forma majoritária nos três gêneros. Os demais esquemas analisados apresentam menor frequência e têm representação mais fraca se comparados ao exemplar central.

A análise também indicou que os gêneros diferem quanto ao domínio instaurado pelas construções conclusivas. Enquanto as redações do Enem apresentam um equilíbrio entre relações epistêmicas e referenciais, os textos de blog tendem a favorecer relações mais subjetivas, associadas ao domínio epistêmico, e os artigos científicos privilegiam relações referenciais, mais alinhadas à exposição de dados empíricos e informações factuais. Esses padrões confirmam que o uso dessas construções está intimamente relacionado às finalidades comunicativas de cada gênero.

Além disso, verificou-se que, no gênero redação do Enem, o parágrafo atua como fator relevante na distribuição de construções associadas aos domínios referencial e epistêmico. Os resultados indicam que não apenas o gênero em sua totalidade, mas também a função desempenhada por cada uma de suas partes na organização textual, resulta na maior ou menor recorrência de relações específicas.

No que diz respeito ao registro, os resultados indicam que ele não se mostrou decisivo para explicar o uso das construções analisadas. Os textos de blog, associados a um registro mais informal, apresentaram baixa recorrência das construções, enquanto nas redações, caracterizadas pelo registro formal, foram encontrados muitos dados. No entanto, os artigos científicos, também formais, apresentaram frequência de uso pouco significativa, especialmente se considerada a extensão da amostra em comparação com os outros dois gêneros.

Esses resultados sugerem que o registro, isoladamente, pode ser insuficiente ou pouco relevante para explicar a frequência de uso dessas construções nos textos. Contudo, é importante ressaltar que os artigos científicos podem apresentar um número expressivo de construções conclusivas ligadas por outros conectores, que não foram incluídos nesta pesquisa.

Desse modo, o estudo confirma a importância do gênero textual, para além do registro, como fator decisivo no uso e na distribuição das construções conclusivas analisadas. Os resultados sugerem que a frequência dessas construções está relacionada não apenas ao grau de formalidade, mas principalmente aos objetivos comunicativos e às características próprias de cada gênero. Nesse sentido, pesquisas futuras podem ampliar a investigação, incluindo outros gêneros textuais e outros conectores conclusivos, a fim de aprofundar a compreensão sobre os fatores que afetam tais usos e sua variação no português escrito.

Referências

BIBER, Douglas; JOHANSSON, Stig; LEECH, Geoffrey; CONRAD, Susan; FINEGAN, Edward. *Longman grammar of spoken and written English*. Harlow: Pearson Education Limited, 1999.

BOYLAND, Joyce Tang. Usage based models of language. *In*: EDDINGTON, David. (ed.) *Experimental and quantitative linguistics*. Munique: Lincom, 2009, pp. 351-419.

BRASIL. *A redação do Enem 2024*: cartilha do participante. Brasília, DF: Inep/MEC, 2024.

BYBEE, J. From usage to grammar: the mind's response to repetition. *Language*, v. 82, n.4. pp. 711-733, 2006.

BYBEE, Joan. *Language, usage and cognition*. Nova Iorque: Cambridge University Press, 2010.

BYBEE, Joan; BECKNER, Clay. Usage-based Theory. *In*: HEINE, Bernd; NARROG, Heiko. (eds.). *The Oxford handbook of Linguistic analysis*. Oxford: Oxford University Press, 2010, pp. 827-855.

CAVALCANTE, Francisco Mailson de Lima; SILVA, Ananias Agostinho. O gênero redação do ENEM: um estado do conhecimento. *Revista Eletrônica de Estudos Integrados em Discurso e Argumentação*, v. 23, n. 2, pp. 51-70, 2023.

CEGALLA, Domingos Paschoal. *Novíssima gramática da Língua Portuguesa*. São Paulo: Companhia Editora Nacional, 2008.

CUNHA, Celso; CINTRA, Lindley. *Nova gramática do português contemporâneo.* Rio de Janeiro: Lexikon, 2017.

DIESSEL, Holger. Usage-based construction grammar. *In*: DABROWSKA, Ewa; DIVJAK, Dagmar. (ed.) *Handbook of Cognitive Linguistics*. Berlim: Mouton de Gruyter, 2015, pp. 295-321.

DIESSEL, Holger. *The grammar network*: How linguistic structure is shaped by language use. Cambridge: Cambridge University Press, 2019.

FLORET, Mayra França. *A trajetória das construções conclusivas com portanto, por isso, logo e então*. Tese (Doutorado em Linguística), Faculdade de Letras, UFRJ, Rio de Janeiro, 2022.

FLORET, Mayra França. Construções conclusivas com por isso: do português arcaico ao português atual. *Domínios de Lingu@gem*, Uberlândia, v.17, p.e1726, 2023.

FLORET, Mayra França. A influência dos parágrafos no uso de construções conclusivas em redações do ENEM. *In*: WIEDEMER, M. L.; OLIVEIRA, M. C. (org.). *Gramática de construções e interfaces linguísticas*. São Carlos: Pedro & João Editores, 2025a, pp. 279-300.

FLORET, Mayra França. Uso de conectores conclusivos preposicionais em diferentes gêneros textuais: do blog à redação do ENEM. *Domínios de Lingu@gem*, Uberlândia, v.19, p.e019072, 2025b.

GOLDBERG, Adele. *A construction grammar approach to argument structure*. Chicago: University of Chicago Press, 1995.

GOLDBERG, Adele. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Express, 2006.

HOFFMANN, Thomas. Constructicon in progress: a short introduction to the constructionist approach to syntactic analysis (CASA). *Yearbook of the German Cognitive Linguistics Association*, v. 11, n. 1, pp. 7-22, 2023.

KOMESU, Fabiana Cristina. Blogs e as práticas de escrita sobre si na internet. *In*: MARCUSCHI, Luiz Antônio; XAVIER, Antônio Carlos. (org.). *Hipertexto e gêneros digitais*: novas formas de construção do sentido. Rio de Janeiro: Editora Lucerna, 2005.

KURY, Adriano da Gama. *Novas lições de análise sintática*. São Paulo: Ática, 1993.

LIU, Dilin. Linking adverbials: an across-register corpus study and its implications. *International Journal of Corpus Linguistics*, v. 13, n. 4, pp. 491-518, 2008.

LOPES, Ana Cristina Macário; PEZATTI, Erotilde Goreti; NOVAES, Norma Barbosa. As construções com portanto no Português europeu e no Português brasileiro. *Scripta*, Belo Horizonte, v. 5, n. 9, pp. 203-218, 2001.

MARCUSCHI, Luiz Antônio. Gêneros textuais: definição e funcionalidade. *In*: DIONÍSIO, Angela Paiva; MACHADO, Anna Rachel R.; BEZERRA, Maria Auxiliadora. (org.). *Gêneros Textuais e Ensino*. Rio de Janeiro: Editora Lucerna, 2002.

MARQUES, Norma Barbosa Novaes. *A relação conclusiva na Língua portuguesa: funções resumo, conclusão e consequência*. Tese (Doutorado em Estudos Linguísticos), Instituto de Biociências, Letras e Ciências Exatas, UNESP, São José do Rio Preto, 2014.

MARQUES, Norma Barbosa Novaes; PEZATTI, Erotilde Goreti. *A relação conclusiva na Língua portuguesa*: funções resumo, conclusão e consequência. São Paulo: Editora UNESP, 2015.

MATOS, Gabriela. Estruturas de coordenação. *In*: MATEUS, Maria Helena Mira *et al*. *Gramática da Língua portuguesa*. Lisboa: Caminho, 2003, pp. 549-592.

OLIVEIRA, Maria do Carmo Pereira. *A sintaxe da coordenação e os conectores conclusivos*. Tese (Doutorado em Linguística). Faculdade de Letras, Universidade do Porto, Porto, 2011.

ROCHA LIMA, Carlos Henrique da. *Gramática normativa da língua portuguesa*. Rio de Janeiro: José Olympio editora, 2011.

SANKOFF, David.; TAGLIAMONTE, Sali A.; SMITH, Eric. *Goldvarb X*: A variable rule application for Macintosh and Windows. University of Toronto, 2005.

SANTOS, Edna Cristina S. A linguagem dos blogs: um gênero textual emergente. *Cadernos de Linguagem e Sociedade*, 6, 2003.

SWEETSER, Eve. *From etymology to pragmatics:* metaphorical and cultural aspects of semantic structure. Cambridge: Cambridge University Press, 1990.

TRAUGOTT, Elizabeth Closs; TROUSDALE, Graeme. *Constructionalization and constructional changes*. Oxford: Oxford University Press, 2013.

Próximas páginas anexo 1 e anexo 2.

Anexo 1: Amostra de textos de blog

Categoria	Blog	Texto
Beleza e cuidados pessoais	Coisas de Diva	1) https://www.coisasdediva.com.br/2019/12/qual-o-melhor-effaclar-bb-blur-ou-blemish-blur-2-0-quintal/ 2). https://www.coisasdediva.com.br/2015/01/saiba-o-que-e-cronograma-capilar/ 3). https://www.coisasdediva.com.br/2023/11/foliculite-o-que-e-causas-e-tratamento/
	Juro valendo	1) https://jurovalendo.com.br/2014/03/31/5-dicas-pra-quem-quer-ficar-loira/ 2) https://jurovalendo.com.br/2018/11/08/actine-sabonete-liquido/ 3) https://jurovalendo.com.br/2017/11/07/saxenda-para-emagrecer/
	Dia de beauté	1) https://diadebeaute.com/2021/09/28/a-medicina-nao-foi-feita-para-o-corpo-feminino/ 2) https://diadebeaute.com/2022/10/18/testamos-a-nova-experiencia-da-authentic-beauty-concept-com-foco-no-couro-cabeludo/ 3) https://diadebeaute.com/2019/04/25/cuidados-com-o-cabelo-cacheado-os-produtos-favoritos-de-voces/
Moda	Thássia Naves	1) https://thassianaves.com/o-inverno-da-linda-de-morrer/ 2) https://thassianaves.com/bttips-hotel-marina/ 3) https://thassianaves.com/as-estampas-desejo-desse-verao-no-oqvestir/
	Blog da Mariah	1) https://www.blogdamariah.com.br/2021/02/a-nova-colecao-de-brunello-cucinelli/ 2) https://www.blogdamariah.com.br/2017/04/42276-2/++3) https://www.blogdamariah.com.br/2020/10/malwee-aposta-em-moda-inteligente/ 3) https://www.blogdamariah.com.br/2020/10/malwee-aposta-em-moda-inteligente/
	Garotas estúpidas	1) https://garotasestupidas.com/luisa-sem-mascara/ 2) https://garotasestupidas.com/a-revolucao-de-2019-entenda-por-que-este-foi-um-ano-de-mudancas/ 3) https://garotasestupidas.com/celebs-brasileiras-mais-fashionistas-os-looks-da-festa-de-totalmente-demais/
Esporte	Gazeta esportiva	1) https://www.gazetaesportiva.com/mais-esportes/judo-brasileiro-fecha-2015-vencendo-venezuela-no-super-desafio/ 2) https://www.gazetaesportiva.com/times/manchester-city/gabriel-jesus-lembra-que-pintou-ruas-de-verde-e-amarelo-na-copa-de-2014/ 3) https://www.gazetaesportiva.com/times/sao-paulo/sao-paulo-ganha-quatro-seguidas-no-brasileiro-pela-1a-vez-desde-2014/
	BBC esportes	1) https://www.bbc.com/portuguese/noticias/2014/12/141219_retrospectiva_esporte_5momentos_pai 2) https://www.bbc.com/portuguese/noticias/2016/01/160115_olimpiada_bom_mau_negocio_rm 3) https://www.bbc.com/portuguese/geral-58293067
	R7 esportes	1) https://esportes.r7.com/lance/futebol/ranking-do-ano-veja-qual-foi-o-melhor-e-o-pior-momento-do-sao-paulo-em-2023-28122023/ 2) https://esportes.r7.com/lance/apos-sucesso-em-2022-ronaldo-planeja-comprar-novos-clubes-08122022/ 3) https://esportes.r7.com/futebol/ricardo-oliveira-foi-atraves-da-varzea-que-cheguei-ao-profissional-06072022/

Fofoca	O Fuxico	1) https://ofuxico.com.br/noticias/andressa-urach-foi-o-nome-mais-buscado-no-google-em-2014-espero-que-seja-com-meu-testemunho/ 2) https://ofuxico.com.br/redes-sociais/juliette-revela-cor-da-calcinha-de-usou-na-virada-de-2020-vai-esgotar/
	Portal Uai	1) https://www.uai.com.br/app/noticia/mexerico/2016/05/13/noticias-mexerico,179827/silvio-santos-e-a-personalidade-mais-admirada-no-brasil-segundo-pesqu.shtml 2) https://www.uai.com.br/app/entretenimento/famosos/2023/12/12/not-famosos,334004/no-aniversario-de-belo-horizonte-veja-12-famosos-que-nasceram-na-capital.shtml 3) https://www.uai.com.br/app/noticia/mexerico/2019/11/09/noticias-mexerico,253095/elasticidade-e-dentes-pequenos-veja-o-que-os-famosos-aprontaram.shtml
	Purepeople	1) https://www.purepeople.com.br/noticia/sandy-da-a-luz-primeiro-filho-com-o-musico-lucas-lima-em-sao-paulo_a20334/1 2) https://www.purepeople.com.br/noticia/-a-fazenda-monick-se-retrata-com-marcos-apos-briga-desculpa-pelo-cuspe_a205742/1 3) https://www.purepeople.com.br/noticia/cleo-pires-explica-corpo-20-kg-mais-magro-e-lembra-criticas-trauma_a310860/1
TV e cinema	Adoro cinema	1) https://www.adorocinema.com/noticias/filmes/noticia-141974/ 2) https://www.adorocinema.com/noticias/filmes/noticia-152229/ 3) https://www.adorocinema.com/noticias/filmes/noticia-154595/
	Cinema } vírgula	1) https://cinemavirgula.blogspot.com/2015/12/critica-star-wars-o-despertar-da-forca.html 2) https://cinemavirgula.blogspot.com/2023/11/a-sete-palmos-uma-das-melhores-series.html 3) https://cinemavirgula.blogspot.com/2017/12/3-series-estreantes-da-netflix-que.html
	Pedro Pitanga	1) https://www.pedropitanga.com.br/duas-series-e-um-documentario-para-te-inspirar-na-quarentena/ 2) https://www.pedropitanga.com.br/grace-frankie-no-netflix/ 3) https://www.pedropitanga.com.br/4-filmes-para-quem-ama-ficcao-cientifica-e-espaco/
Religião	Religio samente (Folha de S.Paulo)	1) https://religiosamente.blogfolha.uol.com.br/2015/01/21/da-pra-tolerar/ 2) https://religiosamente.blogfolha.uol.com.br/2014/09/26/marina-contra-os-politicos-evangelicos/ 3) https://religiosamente.blogfolha.uol.com.br/2014/07/16/as-regras-do-templo-de-salomao/
	Guiame	1) https://guiame.com.br/gospel/mundo-cristao/sinais-do-fim-em-2022-fome-atingiu-27-milhoes-de-criancas-apos-secas-e-inundacoes.html 2) https://guiame.com.br/gospel/mundo-cristao/pastor-perdoa-homens-que-mataram-sua-esposa-gravida-quero-que-eles-sigam-jesus.html 3) https://guiame.com.br/gospel/noticias/igreja-nao-tera-que-pagar-multa-por-fazer-cultos-na-pandemia-decide-tribunal-nos-eua

Religião	Sofá gospel	1) https://sofagospel.blogspot.com/2016/12/o-que-esperar-para-2017.html 2) https://sofagospel.blogspot.com/2018/01/dicas-para-ter-sucesso-em-tudo-que-fizer.html 3) https://sofagospel.blogspot.com/2019/09/nao-culpe-o-tempo.html
Tecnologia	Tecnoblog	1) https://tecnoblog.net/arquivo/148983/ces-2014-bitcoin/ 2) https://tecnoblog.net/noticias/bluetooth-tem-falhas-que-poem-em-risco-aparelhos-lancados-desde-2014/ 3) https://tecnoblog.net/noticias/huawei-200-milhoes-smartphones-2018-eua/
	TechTudo	1) https://www.techtudo.com.br/noticias/2015/12/bola-fora-confira-os-principais-fiascos-tecnologicos-de-2015.ghtml 2) https://www.techtudo.com.br/noticias/2022/12/techtudo-completa-12-anos-e-comemora-com-premios-em-2022.ghtml 3) https://www.techtudo.com.br/noticias/2019/12/premio-melhores-do-ano-techtudo-anuncia-eletronicos-vencedores-em-2019.ghtml
	Canal Tech	1) https://canaltech.com.br/meio-ambiente/2020-o-ano-dos-furacoes-satelites-indicam-temporada-intensa-no-atlantico-170389/ 2) https://canaltech.com.br/saude/nao-pessoas-vacinadas-contra-covid-nao-brilham-no-escuro-200717/ 3) https://canaltech.com.br/smartphone/baixa-procura-pelo-iphone-x-pode-levar-apple-a-reduzir-preco-do-smartphone-105810/
Viagem	Prefiro viajar	1) https://prefiroviajar.com.br/brasil/oliale-casa-da-alemanha-olimpiadas-rio-2016 2) https://prefiroviajar.com.br/dicas-de-viagem/curiosidades/cingapura-ou-singapura-qual-correto 3) https://prefiroviajar.com.br/mundo/navigli-o-bairro-boemio-de-milao
	Mala de viagem	1) https://blogmaladeviagem.com.br/essenza-hotel-jericoacoara/ 2) https://blogmaladeviagem.com.br/coimbra-o-que-fazer-2-dias/ 3) https://blogmaladeviagem.com.br/melhores-destinos-dezembro/
	Apaixonados por viagens	1) https://www.blogapaixonadosporviagens.com.br/2021/09/ilha-do-japones-em-cabo-frio-paraiso-ou-furada.html 2) https://www.blogapaixonadosporviagens.com.br/2014/07/berlim-esta-em-festa-este-ano-de-2014.html 3) https://www.blogapaixonadosporviagens.com.br/2016/10/4-edicao-do-mondial-de-la-biere-comecou.html

Anexo 2: Amostra de artigos científicos

Área de conhecimento	Referências
Ciências agrárias	FILHO, A. C.; TOEBE, M.; ALVES, B. M.; BURIN, C.; SANTOS, G. O.; FACCO, G.; NEU, I. M. M. Dimensionamento amostral para avaliar caracteres morfológicos e produtivos de aveia preta em épocas de avaliação. **Ciência Rural**, v. 45, n. 1, p. 9-13, 2015. SANTOS, D. W. F. N. Desempenho operacional de máquinas florestais em função do número de dias da escala de trabalho. **Revista de Ciências Agrárias**, v. 64, p. 1-7, 2021. PONCIANO, V. F. G.; PONCIANO, I. M.; VALICHESKI, R. R.; OLIVEIRA, S. S. C.; CRUZ, S. J. S.; FILHO, W. C. V. Retenção de água no solo sob diferentes adubos verdes como cobertura vegetal em neossolo litólico distrófico. **Ciência Agrícola**, v. 20, n. 2, p. 1-19, 2022.
Ciências biológicas	BURGOS, T. N.; SCHUROFF, P. A.; LOPES, A. M.; LIMA, N. R.; PELAYO, J. S. Água de consumo humano proveniente de poços rasos como fator de risco de doenças de veiculação hídrica. **Rev. Ciênc. Saúde**, v. 16, n. 1, p. 34-38, 2014. RAPATÃO, V. S.; PEIRÓ, D. F. Ilustração científica na Biologia: aplicação das técnicas de lápis de cor, nanquim (pontilhismo) e grafite. **Revista da Biologia**, v. 16, n. 1, p. 7-14, 2016. PASSOS, G. A. G.; DIAS, L. P.; LIMA-NEIVA, V.; MIRANDA, A. V.; COSTA, J. Levantamento preliminar das espécies de Lepidoptera encontradas no Campus da Fiocruz, Manguinhos (RJ), Brasil. **Acta Biol. Par.**, v. 47, p. 33-55, 2018.
Engenharias	REJAILI, R. P. A. Análise de métodos de detecção de vazamento com foco em operações de offloading. **Mecatrone**, n. 1, artigo 3, p. 1-11, 2015. DARIFF, A. P.; DAMO, A. C. M.; KEMPKA, A. P. Clarificação de efluente proveniente de uma agroindústria do oeste catarinense utilizando extrato de acácia negra (Acácia mearnsii) como coagulante. **ENGEVISTA**, v. 19, n. 1, p. 86-98, 2017. CINTRA, W. A.; MENDES, A. F. Estudo do comportamento de vigas em concreto armado quando submetidas a grandes vãos. **CREARE – Revista das Engenharias**, v. 2, n. 1, 1. ed., 2019.
Ciências exatas e da Terra	BRITO, A. P.; SILVA, N. C.; TOMASELLA, J.; FERREIRA, S. J. F.; MONTEIRO, M. T. F. Análise do índice de anomalia de chuva e tendência de precipitação para estações pluviométricas na Amazônia Central. **Revista Brasileira de Meteorologia**, v. 37, n. 1, p. 19-30, 2022. LIMA, C. L. S. O papel do amortecimento e das condições iniciais nas oscilações subamortecidas. **Revista Brasileira de ensino de Física**¸v.45, e20220221, 2023. LIMA, N. K.; ARTONI, A. J.; RODRIGUES, P. G.; ZEFERINO, C. P.; BRENNECKE, K.; FERRAZ, L. P.; MAGNO, M. P.; SOARES, V. E. Parâmetros hematológicos e bioquímicos de coelhos suplementados com probióticos na dieta. **Revista VIDA**, v. 1, n. 1, p. 36-48, 2023.

Ciências humanas	SOUZA, M. A.; BUSSOLOTTI, J. M.; RIBEIRO, S. L. S.; CUNHA, V. M. P. O uso de plataformas digitais e flipped em uma disciplina no Mestrado Profissional em Educação. **Revista Ciências Humanas - UNITAU**, v. 12, n. 2, 24. ed., p. 189-207, 2019. MELLO, C. M. G.; PEREIRA, P. O.; RODRIGUES, D. C. G. A. A problemática do bullying no contexto escolar: proposta de aplicativo como ferramenta de resiliência. **Revista de Ciências Humanas**, v. 21, n. 1, p. 29-46, 2020. VALDIER, M. A jornada do herói: uma proposta para o letramento literário. **Revista de Ciências Humanas**, v. 3, n. 21, p. 1-13, 2021.
Linguística, Letras e Artes	MARINHO, E. S.; FERRARI, L. Mesclagem conceptual em piadas curtas. **Revista LinguíStica**, v. 12, n. 1, p. 147-160, 2016. MARTINS, A. L. O anjo e a besta: a antropologia pascaliana no laboratório de Machado de Assis. **Aletria**, v. 29, n. 4, p. 65-82, 2019. GOBBI, M. A.; BÚFALO, J. Mário de Andrade: turista aprendiz em terra arrasada, da Amazônia a São Paulo, até dizer chega. **Revista Artes de educar**, v. 8, n. 1, p. 34-53, 2022.
Ciências da saúde	FIALHO, H.; MOURA, N.; KOSER, M.; TEIXEIRA, A.; AZEVEDO, R. O caminho para esquizofrenia pela teoria genética de novos genes: uma revisão sistemática. **Rev.Ciênc.Saúde**, v. 16, n. 1, p. 24-33, 2014. LIRA, G. N. W.; PADILHA, W. W. N. Qualidade dos registros enviados ao sistema de informação da atenção básica pelas equipes da Atenção Básica da Paraíba, em 2019. **Revista Brasileira de Ciências da Saúde**, v. 24, n. 4, p. 527-536, 2020. TAVARES, P. C. H. A influência da neurociência cognitiva sobre a Musicoterapia: bases para a Musicoterapia Neurológica. **RCS Revista Ciências da Saúde**, v. 1, n. 1, p. 8-28, 2023.
Ciências sociais	GUIMARÃES, N. A.; BRITO, M. M. A.; BARONE, L. S. Mercantilização no feminino: a visibilidade do trabalho das mulheres no Brasil. **Revista Brasileira de Ciências Sociais**, v. 31, n. 90, p. 17-39, 2016. MONSMA, K. Como pensar o racismo: o paradigma colonial e a abordagem da sociologia histórica. **Revista de Ciências Sociais**, v. 48, n. 2, p. 53-82, 2017. PASSIANI, E. Figuras do intelectual: gênese e devir. **Sociologias**, n. 47, p. 16-47, 2018.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72230

Variação dos pronomes de 2ª pessoa do singular com função de sujeito nas músicas de Poesia Acústica: análise sociofuncionalista sobre a influência da variável gênero

Variation of second-person singular subject pronouns in Poesia Acústica songs: a sociofunctional analysis of the influence of the gender variable

Dennis Castanheira[1]

Júlia Saadi[2]

RESUMO

Este trabalho tem como objetivo investigar a variação pronominal de 2ª pessoa do singular entre os pronomes "tu", "você" e "cê" com função de sujeito nas músicas do projeto *Poesia Acústica*, examinando o papel da variável gênero. A pesquisa é fundamentada na perspectiva sociofuncionalista, integrando pressupostos da Sociolinguística Variacionista (Labov, 2008 [1972]) e do Funcionalismo Norte-americano (Givón, 1995), que consideram tanto condicionadores linguísticos quanto sociais na constituição da gramática. Também dialoga com estudos de gênero que discutem a relação entre papéis sociais, performances discursivas e práticas linguísticas (Ribeiro, 2018). Partimos da hipótese de que h]á uma relação entre a variação dos pronomes de 2ª pessoa do singular com função de sujeito e a variável gênero, refletindo construções identitárias e discursivas. O *corpus* é constituído por todas as músicas do projeto, nas quais foram identificadas e classificadas as ocorrências dos pronomes. A metodologia é qualitativa e quantitativa, em que buscamos mapear os usos caso a caso e também sistematizá-los e tabulá-los por meio do Microsoft 365, mais especificamente, o Excel. Nossos resultados indicam diferenças significativas entre os gêneros, sugerindo que a variação pronominal, nesse *corpus*, opera também como estratégia discursiva vinculada às identidades e às performances dos artistas.

PALAVRAS-CHAVE: Variação pronominal de 2ª pessoa do singular. Sujeito. Sociofuncionalismo. Gênero. Poesia Acústica.

ABSTRACT

This study aims to investigate second-person singular pronominal variation among the pronouns *"tu"*, *"você"* and *"cê"* functioning as subjects in the songs of the *Poesia Acústica* project, examining the role of the gender variable. The research is grounded in the sociofunctional perspective, integrating assumptions from Variationist Sociolinguistics (Labov, 2008 [1972]) and North American Functionalism (Givón, 1995), which consider both linguistic and social constraints in the constitution of grammar. It also engages with gender studies that discuss the relationship between social roles, discursive performances, and linguistic practices (Ribeiro, 2018). We start from the hypothesis that there is a relationship between the variation of second-person singular subject pronouns and the gender variable, reflecting identity and discursive constructions. The *corpus* consists of all the songs from the project, in which occurrences of the pronouns were identified and classified. The methodology is both qualitative and quantitative, seeking to map usages on a case-by-case basis and also to systematize and tabulate them using Microsoft 365, specifically Excel. Our results indicate significant differences between genders, suggesting that pronominal variation in this *corpus* also operates as a discursive strategy linked to the artists' identities and performances.

KEYWORDS: Pronominal variation of second-person singular forms. subject. Sociofunctionalism. Gender. Poesia Acústica.

[1] Universidade Federal Fluminense (UFF), denniscastanheira@gmail.com, https://orcid.org/0000-0001-9092-5936.

[2] Universidade Federal Fluminense (UFF), juliasaadi@id.uff.br, https://orcid.org/0009-0004-9063-0029.

Recebido em 19/01/2026 - Aceito em 11/06/2026

1. Introdução

A variação pronominal de segunda pessoa do singular no Português do Brasil (PB) constitui um dos fenômenos mais recorrentes e discutidos nos estudos de variação e mudança linguística (Paredes Silva, 2003; Mota, 2008; Ferreira, 2010; Scherre *et al.*, 2015; Silva, 2025). Diferentemente de outras línguas, o PB apresenta um quadro altamente heterogêneo, no qual diferentes formas pronominais com função sintática de sujeito ("tu", "você", "ocê" e "cê") coexistem de modo variável. Essa diversidade se manifesta tanto nos usos dos pronomes quanto nas flexões verbais associadas, configurando um campo rico para investigar o funcionamento social, histórico e discursivo da língua portuguesa.

Como fica evidente na discussão de Freitag (2015), nos estudos variacionistas, a variável gênero tem sido tradicionalmente considerada como um fator externo potencialmente relevante para essa variação, ainda que seus efeitos nem sempre sejam evidentes em todas as comunidades de fala. Pesquisas clássicas, como sumarizado por Paiva (2003), apontam que mulheres, em determinadas circunstâncias, tendem a favorecer variantes prestigiadas, enquanto homens podem aderir com maior frequência a formas socialmente menos monitoradas; entretanto, estudos posteriores problematizaram essa generalização ao demonstrar que os usos se relacionam a práticas sociais, papéis identitários e performances situadas (Coates, 2006; Freitag, 2015).

Considerando essas discussões revisionais, nosso objetivo é analisar, à luz do Sociofuncionalismo, a influência dessa variável na variação dos pronomes de 2ª pessoa do singular com função sintática de sujeito nas músicas do projeto Poesia Acústica. Nossa hipótese é que há uma relação entre a variação dos pronomes de 2ª pessoa do singular com função de sujeito e a variável gênero, refletindo construções identitárias e discursivas. Isso será explorado ao longo das próximas seções.

Destacamos que o projeto combina elementos do rap, trap, funk, reggae e da poesia falada, apresenta forte alternância de artistas, protagonismo diversificado e ambiente textual híbrido. Esse cenário favorece o surgimento de marcas linguísticas associadas à fala cotidiana e representativas do PB. Assim, o *corpus* constitui uma fonte relevante para observar como diferentes artistas mobilizam estratégias discursivas e como a variável gênero pode influenciar a seleção de variantes.

Do ponto de vista teórico, o estudo se fundamenta no Sociofuncionalismo, perspectiva que articula princípios da Sociolinguística Variacionista e do Funcionalismo Norte-americano. Essa abordagem entende que o funcionamento da língua resulta da interação entre fatores estruturais e pragmáticos, sendo sensível às pressões discursivas, à frequência e à especialização de formas (Tavares, 2013; Tavares; Görski; Paza, 2025). Nesse viés, a variação é um reflexo das nuances comunicativas e a gramaticalização é um dos pontos centrais, tendo as motivações discursivas mapeadas.

Por fim, destacamos que o trabalho será organizado em algumas seções. Inicialmente, teremos um capítulo teórico voltado para os pressupostos do Sociofuncionalismo e para as discussões sobre a variável gênero. Em seguida, vamos expor como as gramáticas normativas e de uso apresentam os pronomes estudados, discutiremos os processos de mudança e gramaticalização envolvidos na consolidação da forma "você" e abordaremos os usos atuais dos pronomes de segunda pessoa. Logo

após, descreveremos a metodologia adotada para a composição e análise do *corpus* e, na sequência, exploraremos a análise dos dados, articulando os resultados quantitativos e qualitativos. Por fim, apresentaremos as considerações finais, acompanhadas da bibliografia utilizada ao longo da pesquisa.

2. Sociofuncionalismo e estudos de gênero

Segundo Castanheira (2018) e Sousa, Kanthack e Silva (2023), o Sociofuncionalismo engloba os pressupostos da Sociolinguística e do Funcionalismo norte-americano por meio do estudo da língua em uso de modo contextualizado. Essa articulação não é nova e vem sendo desenvolvida há algumas décadas em alguns centros de pesquisa brasileiros, como a Universidade Federal do Rio de Janeiro, a Universidade Federal de Santa Catarina, a Universidade Federal Fluminense, a Universidade Federal do Ceará e a Universidade Federal do Sudoeste da Bahia.

Görski e Tavares (2013), Tavares (2013) e Tavares e Görski (2015) defendem que é possível relacionar as duas teorias, já que há diversos pontos estruturantes em convergência – ainda que seja necessário realizar algumas adaptações e concessões. Exemplos disso são: foco no uso nas análises da língua; mudança linguística vista como gradual; relação de complementariedade entre a sincronia e a diacronia; análise dos vários níveis da língua; vinculação entre língua e sociedade; pertinência dos fatores interacionais na variação e na mudança.

Para Tavares (2003), os pressupostos centrais da abordagem sociofuncionalista se voltam para a função a que serve a gramática, que nunca deve ser vista como algo pronto e constituído, mas como aquela que sofre alterações pelos usos dos falantes. Segundo a autora, é fundamental que haja, por parte do estudioso, análise das inter-relações entre condicionamentos internos e externos à língua, ponderando quais são as motivações em cada situação comunicativa e de que forma contribuem com as atualizações constantes. Assim, as estruturas são ligadas às motivações funcionais.

Além disso, são estudadas as diferentes etapas do processo de gramaticalização, partindo da ideia de que essa diversidade é influenciada pelas transformações que cada elemento sofreu ao longo de sua trajetória até alcançar a função que desempenha atualmente. Isso engloba, por exemplo, a gramaticalização do pronome "você" (Lopes *et al.*, 2025), que passou de pronome de tratamento para pronome pessoal e continua sofrendo redução, resultando nas formas "ocê" e "cê". Desse modo, considerando a língua como um fenômeno social além de um instrumento de interação, e também da coletividade, é necessário ponderar os condicionadores linguísticos e sociais que estão atrelados a uma situação comunicativa.

Além da gramaticalização, outro ponto bastante relevante é a marcação, em que são estabelecidos critérios, como a complexidade estrutural, a frequência e a complexidade cognitiva, que delineiam a distinção entre as categorias marcadas e não-marcadas. Pensando na variação dos pronomes de segunda pessoa do singular, a partir desse princípio, seria possível compreender qual é a estrutura mais e menos saliente (Cunha; Costa; Cezario, 2003). Segundo Givón (1995), a marcação é um fenômeno dependente do contexto, devendo ser explicada com base em fatores comunicativos e cognitivos. Por

isso, algumas formas são marcadas em algumas situações e não são em outras.

No que tange à variável gênero, a discussão ocorre porque existe uma grande diferença nos valores e nos comportamentos que se esperam de um homem e de uma mulher, e "o cotejamento das estatísticas destas diferenças por sexo sinaliza padrões de gênero" (Freitag, 2015, p. 18). A autora ressalta, a partir dos dados do IBGE, que, no Brasil, o lento processo de mudança dos padrões culturais de gênero diminuiu as tradicionais dificuldades com a entrada das mulheres no mercado de trabalho, reduziu a taxa de fecundidade e elevou a escolaridade das mulheres ao longo das últimas três décadas.

Segundo Paiva (2003), as investigações variacionistas inclinam-se para algumas noções no tocante à relação do gênero com os fenômenos linguísticos estudados. As mulheres tendem a ser menos inovadoras do que os homens quando são formas fora do padrão por sofrerem uma pressão do conservadorismo, contudo tendem às inovações quando são formas sem estigma. Nesse caso, utilizar a variante de prestígio seria uma forma de atingirem um maior *status* social, normalmente atrelado aos homens.

Por outro lado, Freitag (2015) afirma que, se a Sociolinguística tem como premissa o estudo da relação entre língua e sociedade, é preciso considerar que a atual realidade é bem diferente daquela apresentada nos anos de 1960, por exemplo, em que a mulher era vista como aquela que precisava manter o papel de mãe e educadora. Conforme Freitag (2015), autores como Severo (2006), criticam as elucidações já consagradas, tendo em vista que, embora versem sobre as diferenças que permeiam as questões linguísticas e sua relação com o gênero, não são suficientemente significativas no que tange às motivações para tais distinções e que houve muitas mudanças sociais ao longo dos anos.

Ao discutir os usos linguísticos das mulheres, também é preciso pensar que não se trata de um grupo homogêneo, pelo contrário, há que se considerar outros fatores, como escolaridade, raça, classe social e faixa etária. Isso indica que é preciso adotar um olhar interseccional, em que as mulheres sejam vistas a partir de suas múltiplas complexidades e características e também inseridas em um contexto discursivo que deve ser discutido e sistematizado (Ribeiro, 2018; Castanheira; Martins, 2024). Ressaltamos, contudo, que nem sempre é possível considerar todas essas complexidades devido ao *corpus* ou às informações disponíveis e, por isso, é preciso que a investigação considere o máximo de questões dentro das possibilidades.

Freitag (2015) ressalta que, para entender as relações de gênero, é preciso considerar a história do feminismo. Para Martins (2015), a disseminação das ideias de emancipação feminina trouxe consigo uma enorme quantidade de questionamentos sobre a materialidade desse sujeito. É necessário questionar, por exemplo, quem seriam esses novos sujeitos, quem são as mulheres aptas a adentrar no universo do reconhecimento de direitos, quais são os limites da identidade política das mulheres no processo de reivindicação e exercício de direitos.

Há, nesse âmbito, de acordo com Coates (2006) e Freitag (2015), diferentes olhares para o tratamento da relação entre linguagem e gênero: (a) déficit – nessa visão os usos linguísticos das mulheres constroem uma imagem de fraqueza, passando uma imagem de não-assertividade; (b)

dominância – sob esse olhar, as mulheres são um grupo oprimido e, por isso, a dominância masculina é reiterada pelos participantes de uma interação; (c) diferença – nessa visão, as diferenças culturais geram distinções entre homens e mulheres e (d) construção social – sob esse olhar, a identidade de gênero é uma construção social, "não é algo estático, mas é algo construído no cotidiano" (Freitag, 2015, p. 24).

Oushiro (2015), em um estudo que visa à discussão do papel que exerce a variável em gravações de 118 paulistanos, com ênfase nas concordâncias verbais de números em sintagmas verbais (SVs) de primeira pessoa do plural (1PP) e de terceira pessoa do plural (3PP), demonstrou a importância de considerar os grupos sociais ao definir as tendências de uso das formas padrão. No tocante à 3PP, por exemplo, a classe média apresenta uma maior adesão às formas padrão e a diferença entre gêneros é relativamente menor entre os informantes com Ensino Fundamental e Superior, em relação aos que cursaram até o Médio; já em relação à 1PP, as mulheres que pertencem à classe baixa com menor escolaridade e mobilidade apresentam um padrão inverso.

Além disso, o estudo ratifica a importância de compreender os diferentes papéis sociais de homens e mulheres e o seu pertencimento a distintos subgrupos para definir melhor as possíveis correlações e também haver um olhar multidimensional. Oushiro (2015) delineia alguns motivos possíveis que poderiam motivar a escolha das mulheres pelas variantes de prestígio, como, por exemplo, superar a falta de privilégios na sociedade ou por questões de *status* social no geral. A autora também salienta que os homens se reafirmam a partir da sua profissão ou renda.

Em um outro estudo mais recente, Castanheira e Martins (2024) objetivam à análise de letras de funk cantadas por mulheres para compreender se há relação entre os seus perfis sociais e musicais e as formas variáveis usadas. Semelhanças e diferenças entre os perfis das artistas foram percebidas por Castanheira e Martins (2024) ao procederem à análise das letras de músicas. Nas três letras, há casos de variação linguística, algumas mais ou menos aceitas socialmente. Além disso, há, em todas as letras, formas não padrão de maior estigma, como a não marcação da concordância. Embora tais usos possam estar relacionados a diferentes motivações, segundo os autores, é preciso destacar que há usos singulares que podem ser explicados pela construção da identidade das artistas.

A utilização do termo "otário", por Mc Carol, e a expressão "bom de cama", empregada por Luísa Sonza, são pontos que apresentam relevância dentro de uma análise. Assim, não apenas os casos de variação, mas a construção do contexto discursivo nas letras, como é feito pelo Sociofuncionalismo, são utilizados. Castanheira e Martins (2024) também conclui que, no caso do *feat* de Jojo Toddynho e Anita, fica nítida a diferença lexical que acompanha os perfis de ambas: enquanto a Anitta utiliza variantes mais aceitas e constrói uma imagem "familiar", Jojo Toddynho utiliza palavrões, de forma a afrontar a sociedade com seu discurso.

Destacamos, por fim, que, neste artigo, o Sociofuncionalismo é utilizado para fundamentar o trabalho realizado, o que se evidencia pela discussão da gramaticalização das formas "você" e "cê", que resulta em sua variação com "tu", pela análise do papel da variável gênero, pela análise do

contexto discursivo das músicas do projeto Poesia Acústica e pela discussão de questões semânticas e morfossintáticas ligadas à pragmática. Nesse âmbito, o Sociofuncionalismo engloba a interface variação-discurso, já sistematizada em reflexões teóricas prévias, aqui retomadas, como Tavares (2013), Castanheira (2018) e Tavares, Görski e Paza (2025).

3. Variação pronominal de 2ª pessoa do singular: o que dizem as gramáticas e as pesquisas linguísticas?

Com a finalidade de compreender como os gramáticos e os linguistas estudam os pronomes sujeito de segunda pessoa do singular, serão retomadas as visões de algumas Gramáticas Tradicionais (GT), de algumas Gramáticas Descritivas (GD) e de alguns estudos linguísticos de cunho socio(funcionalista). Isso possibilitará um panorama mais amplo do tema a fim de retomar o que já foi feito e também dialogar com os resultados a serem discutidos ao longo deste artigo.

Rocha Lima (2011 [1992], p. 156), com relação à segunda pessoa, menciona que se trata daquela com quem se fala, com singular "tu" e plural "vós", afirma que o verbo declara as pessoas a que o predicado faz referência e demonstra quais são as terminações referentes às pessoas verbais – no caso da segunda pessoa do singular ("tu"), acrescenta-se o "s" ao final do verbo ("tu amas"). Mais adiante, o autor cita o pronome "você", classificando-o como um pronome de tratamento, sendo esse familiar. O autor ainda frisa que, ainda que se refira ao interlocutor, exige concordância verbal com a terceira pessoa do singular.

Por sua vez, Cunha e Cintra (2016 [1985]) caracterizam os pronomes pessoais como aqueles que denotam as três pessoas gramaticais, ou seja, indicam quem fala, com quem se fala, de quem se fala. No caso da segunda pessoa do singular, afirmam que pertencem ao segundo caso, isto é, com quem se fala, e afirmam que "tu" é a forma singular de segunda pessoa, ao passo que "vós" é a forma plural. Já na seção denominada "pronomes de tratamento", Cunha e Cintra (2016 [1985]) abordam o pronome "você" e afirmam que, no caso do PB, substituiu o "tu" em algumas situações.

Já Bechara (2009) define os pronomes pessoais como os que designam as pessoas do discurso, classifica "tu" como segunda pessoa do singular e "vós" como segunda pessoa do plural. A respeito de "você", o autor menciona, ao longo da seção, que pertence ao que intitula como "formas substantivas de tratamento" ou "formas pronominais de tratamento", e, desse modo, é empregado no tratamento familiar. Ainda, conforme Bechara (2009), "você", hoje usado familiarmente, é a redução da forma de reverência "vossa mercê", e "vocês" assume o plural de "tu", pelo fato de o pronome "vós" atualmente ser considerado um pronome em desuso – utilizado apenas nas orações e estilo solene.

Acerca das GD, destacamos Neves (2011), que afirma que os pronomes de segunda pessoa, seja do singular ou plural, fazem referência a um dos interlocutores, ou seja, a uma das pessoas que pertence ao circuito de comunicação. Ao apresentar o quadro pronominal, a autora cita "tu" e "você", como pronomes de segunda pessoa do singular. Posteriormente, esclarece que as formas "você" e "vocês" se referem à segunda pessoa, porém levam o verbo para a terceira pessoa, da mesma forma

que funciona com os pronomes de tratamento e ressalta que o emprego de "você" é muito mais difundido que o de "tu", ao se referir a um interlocutor.

Por outro lado, Castilho (2014) ressalta que os pronomes pessoais são muito suscetíveis a mudanças e que a sua reorganização no PB, principalmente em sua modalidade falada, acarreta fortes consequências na estrutura sintática da língua. Na apresentação dos pronomes pessoais do PB na atualidade, o autor considera "você", "ocê" e "tu" como pronomes de segunda pessoa do singular. Castilho (2014) também esclarece que "você" surgiu a partir da gramaticalização da expressão de tratamento "Vossa Mercê" e que pode ser explicada por questões fonológicas, sintáticas e pragmáticas.

Em relação aos estudos linguísticos, destacamos os trabalhos de Silva (2025), que esclarece que a forma "você" se relaciona diretamente à expressão de tratamento "Vossa Mercê", explicitada anteriormente como aquela que era utilizada para se referir a pessoas com uma elevada posição social. Ao longo do tempo, essa expressão sofreu um processo de desgaste fonético e redução lexical fruto da sua gramaticalização. Vale frisar que esse processo de transformações fonéticas foi se consolidando paulatinamente, sendo alimentado não só pelo fato de que a forma "Vossa Mercê" baixava na escala social, mas, ao mesmo tempo, por ser alvo de muita repetição a toda hora e por todo mundo.

Com o surgimento e a consolidação do uso de "você", no entanto, não houve a eliminação do pronome "tu". Este ainda é amplamente utilizado em diversas comunidades linguísticas, sendo muitas vezes acompanhado de formas verbais da terceira pessoa. Essa coexistência entre diferentes formas pronominais evidencia um sistema variado no PB, no qual múltiplas possibilidades são ativadas de acordo com fatores socioculturais e contextuais, como o grau de intimidade entre os falantes.

Paredes Silva (2003) investiga a trajetória do pronome "tu" na fala carioca, partindo da ideia de que, embora o português disponha de duas formas para a segunda pessoa ("tu" e "você"), descrições recentes restringiam o uso do "tu" quase exclusivamente ao Sul do país. No Rio de Janeiro, a substituição de "tu" por "você" ocorreu progressivamente ao longo do século XX, e a autora demonstra que, apesar desse apagamento histórico, há sinais de movimento de retorno do "tu" a partir da segunda metade do século XX, inicialmente associado a personagens de baixa condição social e a usos não padrão.

De acordo com a autora, a análise confirma tendências já instituídas na dramaturgia a partir dos anos 1950, que passaram a colocar o "tu" na fala de personagens jovens de classes sociais mais baixas, revelando um uso emergente e não padrão. Paredes Silva (2003) ainda argumenta que essa retomada se articula com outros fenômenos de variação de segunda pessoa e que a entrada histórica de "você" no sistema pronominal gerou desequilíbrios que ainda reverberam na fala carioca. Assim, a autora conclui que a nova configuração do sistema, marcada pelo uso socialmente estigmatizante do "tu" sem concordância, pode representar uma mudança linguística em desenvolvimento.

Por outro lado, Scherre *et al.* (2015), a partir de um mapeamento sociolinguístico do português brasileiro, compilam dados de diversas pesquisas feitas ao longo dos anos sobre o uso dos pronomes "tu", "você", "cê" e "ocê", não só com o objetivo de apresentar os resultados globais de diversos

trabalhos sociolinguísticos sobre os usos desses pronomes, mas remodelar mapa anterior para representar os seis subsistemas propostos no eixo geográfico ou horizontal. Nessa investigação, os trabalhos utilizados foram orientados pela perspectiva da Teoria da Variação e da Mudança linguística de base laboviana.

A partir dos resultados, as autoras propuseram a organização do uso desses pronomes em seis subsistemas: (i) somente "você"; (ii) predominância de "tu" com baixa marcação de concordância; (iii) predominância de "tu" com alta concordância; (iv) alternância entre "tu" e "você" com concordância baixa; (v) alternância entre "tu" e "você" com concordância intermediária; e (vi) uso de "você" e "tu" sem marcação de concordância. Esses são encontrados de diferentes maneiras nas localidades do país, como o Rio de Janeiro, e podem caracterizar alguns grupos a partir do seus perfis sociais.

Diante dos estudos expostos sobre a variação pronominal de 2ª pessoa do singular, podemos afirmar que a forma "você" passou por gramaticalização e está em variação com outras e que há diferentes usos em distintas regiões do Brasil, bem como há a necessidade de discutir o papel de cada variável de modo mais detalhado. Neste trabalho, faremos isso por meio da análise das músicas de Poesia Acústica nas próximas seções.

4. Metodologia

Na presente pesquisa, a metodologia empregada articula procedimentos quantitativos e qualitativos, tal como realizado em diversas pesquisas sociofuncionalistas, que reconhecem a importância simultânea da frequência e dos contextos discursivos para a constituição das formas gramaticais (Tavares, 2013; Castanheira, 2018). Isso indica que todos os dados foram analisados caso a caso e, posteriormente, quantificados e sistematizados para mapeamento de frequência de uso.

O *corpus* utilizado é composto por músicas do projeto Poesia Acústica, consolidado como uma das produções mais expressivas da cena musical urbana contemporânea no Brasil, caracterizado por um formato híbrido que articula elementos do rap, do trap, do funk e do reggae a arranjos predominantemente acústicos, especialmente o violão. Sua estrutura em formato de *cypher*, isto é, a reunião de múltiplos intérpretes em uma mesma faixa, cada um com a sua contribuição poética, de forma independente, aproxima a proposta tanto da poesia falada quanto da canção popular, resultando em composições extensas.

A historicidade do projeto remonta ao período que compreende os anos de 2016 e 2017, quando a *Pineapple Storm Records* idealizou a proposta como uma série capaz de reunir artistas de diferentes estilos dentro de um mesmo arranjo minimalista. Com o passar dos anos, o projeto foi expandido para múltiplas plataformas, especialmente o *YouTube* e os serviços de *streaming*, consolidando-se como um produto de multimídia de grande alcance e impacto cultural. Dentre os artistas mais recorrentes ou emblemáticos, estão Xamã, Lourena, MC Cabelinho, Filipe Ret, Azzy e L7NNON, formando um grupo heterogêneo.

Justificamos a escolha do *corpus* pelo fato de que as músicas apresentam um estilo discursivo híbrido, o que viabiliza a observação de fenômenos linguísticos que emergem na oralidade mais espontânea. Outro ponto é que percebemos que há poucos trabalhos variacionistas sobre músicas, o que indica que é preciso estudar e discutir com maior detalhamento os usos em amostras como essas. Por fim, destacamos que o projeto Poesia Acústica apresenta significativa diversidade de vozes, permitindo analisar a variável gênero enquanto condicionador extralinguístico ligado à performance e aos papéis sociais.

Nesse âmbito, foram utilizadas, ao longo da análise, 17 músicas, que é o quantitativo total de canções do projeto, focalizando os trechos em que há marcas explícitas de referência à segunda pessoa do singular. Os exemplos apresentados foram transcritos conforme a versão original presente nas plataformas digitais, preservando sua autenticidade. Esclarecemos, ainda, que as músicas são compostas pelos mesmos artistas que as cantam e que há menos mulheres do que homens nas músicas, o que impactou a construção das nossas tabelas na seção a seguir.

Após a constituição do *corpus*, procedemos à identificação de todas as ocorrências dos pronomes de 2ª pessoa do singular com função de sujeito encontrados: "tu", "você" e "cê". Ressaltamos que, no caso dos pronomes "você" e "cê", optamos por observar de forma isolada a variante "cê" (e não agrupá-la com o "você") para compreender melhor os usos que os cantores faziam dessa forma, uma vez que apresentou uma alta frequência. Enfatizamos, ainda, que não consideramos os casos de sujeito não expresso, ou seja, analisamos apenas os dados com usos dos pronomes com papel de sujeito morfossintaticamente expressos.

Para melhor quantificar e, posteriormente, analisar essas ocorrências, utilizamos o Microsoft 365, mais especificamente o Excel, para listar, tabular e organizar as ocorrências. A partir disso, consideramos os seguintes grupos de fatores: (a) gênero, (b) concordância e (c) tipo de referência. Buscamos, nesse âmbito, observar as tendências nos dois gêneros e comparar com resultados de trabalhos anteriores (Scherre *et al*., 2015; Franceschini; Lorengian-Penkal, 2015; Oushiro, 2015).

Por fim, destacamos que a variável gênero foi tratada associando as discussões dos estudos variacionistas tradicionais de modo integrado às atualizações teóricas recentes. Desse modo, não assumimos os comportamentos linguísticos masculino e feminino como categorias homogêneas, mas como indicadores que podem emergir a partir de configurações sociais contextualizadas e situadas, relacionadas à identidade, ao estilo discursivo e à performance artística, tal como discutido também em Castanheira e Martins (2024).

5. Análise

Nesta seção, será apresentada a análise realizada nesta pesquisa. Considerando a hipótese que orienta este estudo de que há uma relação entre a variação dos pronomes de 2ª pessoa do singular com função de sujeito e a variável gênero, refletindo construções identitárias e discursivas, os dados serão discutidos de maneira qualitativa e quantitativa.

Inicialmente, apresentaremos a variação pronominal de 2ª pessoa do singular em relação aos dois gêneros, conforme os exemplos a seguir, com trechos das músicas analisadas:

(1) "Baby o que você queria, eu fiz
Tudo que você pediu, eu dei
Mas **tu** jogou o meu amor no abismo
Dividiu o que era dois pra três
E eu bem que tentei te fazer feliz [...]" (Maria – Poesia Acústica 5)

(2) "Quando **você** reparar, quando **você** me notar
Já vai ser tarde demais, não vai dar tempo mudar
Quero alguém pra me valorizar
Dizer que eu amo sem me preocupar [...]" (Lourena – Poesia Acústica 9)

(3) "[...] No dia em que tudo perder o valor
Talvez **cê** me diga, me diga
O que é saudade
Sem que alguém lhe culpe [...]" (Bob – Poesia Acústica 4)

Esses exemplos ilustram a variação aqui discutida, sistematizada na tabela 1:

Tabela 1: variação pronominal de 2ª pessoa do singular (tu x você x cê)

Homem		**Mulher**	
Tu	85/237	Tu	15/75
	35,86%		20%
Você	101/237	Você	49/75
	42,62%		65,33%
Cê	51/237	Cê	11/75
	21,52%		14,67%
Total	237/237	Total	75/75
	100%		100%

Fonte: Elaboração dos autores

Inicialmente destacamos que a tabela 1 é organizada de modo a separar a frequência de uso de homens e mulheres, visto que há uma quantidade desigual de artistas dos dois gêneros, havendo muito mais homens do que mulheres no projeto, o que poderia enviesar nossa análise. Nesse âmbito, é preciso analisar a variação pronominal em cada gênero para, posteriormente, comparar os percentuais e não os números isolados. Isso possibilita um percurso mais seguro analiticamente.

A tabela 1, que apresenta a frequência em relação aos dados da variação dos pronomes "tu", "você" e "cê", demonstra tendências que podem parecer similares, mas que, na verdade, apresentam

especificidades. Nos trechos cantados por homens, o pronome "você" teve 42,62%, seguido do pronome "tu", com 35,86% e do "cê", com 21,52%. Já nos trechos cantados por mulheres, o "você" também foi o mais usado, com 65,33%, seguido do "tu", com 20% e do "cê", com 14,67%. Isso indica que, na leitura vertical e separada das colunas, ambos utilizam mais o pronome "você", seguido de "tu" e "cê", havendo, então, as mesmas tendências de uso das variantes.

Contudo, em uma visão comparativa, percebemos algumas especificidades que precisam ser destacadas. Em relação ao pronome "você", apesar de ser mais frequente nos dois gêneros, tem uso muito mais expressivo nos trechos cantados por mulheres – 65,33% contra 42,62% nos trechos cantados pelos cantores. Na observação do pronome "tu", a diferença também se evidencia, já que os homens utilizaram em 35,86% e as mulheres em apenas 20%. Em relação ao "cê", os homens utilizam em 21,52% e as mulheres em 14,67%.

Tais diferenças podem ser explicadas pelas características desses pronomes de modo correlacionado à construção de gênero. Conforme Oushiro (2015), os dados podem refletir os distintos papéis sociais atrelados aos homens e às mulheres. No que concerne às mulheres, há uma preferência pela variante de prestígio discursivo, com mais polidez, "você", uma vez que o uso do pronome "tu", conforme demonstrado em estudos como o de Resende (2023), pode conferir ao discurso uma certa inadequação ou indelicadeza.

Durante a pesquisa realizada por Resende (2023), a qual versa sobre os usos dos dois pronomes na fala do município de Bagé (RS), alguns entrevistados se sentiram à vontade para tecer comentários a respeito de sua visão de língua e o falar imposto. Segundo a autora, uma das informantes relatou que, em uma conversa com seu chefe, se referiu a ele pelo pronome "tu", e, logo em seguida, ficou reflexiva e se sentiu mal pela forma com que se expressou, pensando ter sido grosseira ao fazer uso desse pronome, o que reforça a ideia de que o "tu" pode representar indelicadeza.

Destacamos, nesse âmbito, que os comportamentos dos indivíduos são influenciados pelas normas e expectativas culturais de cada sociedade. Desde cedo, somos socializados a seguir padrões de gênero que orientam desde a maneira de se vestir até a forma de se expressar. Enquanto os homens muitas vezes têm maior liberdade para adotar comportamentos mais agressivos ou desafiadores, as mulheres são frequentemente ensinadas a serem mais submissas, cuidadosas e reservadas, o que também se reflete na comunicação. Sob esse viés, as mulheres tendem a internalizar essas expectativas e adaptar seu discurso, enquanto os homens podem usar a linguagem de maneira mais fluida e legítima.

Também devemos relacionar tais debates a estudos, como Scherre *et al.* (2015), em que os autores demonstram que o pronome de 2ª pessoa do singular "você" é o mais comum no PB atual. Por esse motivo, o "você" pode ser interpretado pelas mulheres como a forma menos marcada e, por isso, a que deve ser usada para evitar julgamentos em detrimento de "tu" e "cê". Enquanto o "tu" tende a expressar maior grosseria, o "cê" é uma redução fruto da gramaticalização do "você" e ligado ao menor monitoramento; portanto, seu uso também pode fazer com que não sejam bem-vistas.

Isso é reforçado pelo estudo de Carvalho e Lopes (2022), em que as autoras demonstram que, quanto à percepção dos cariocas, o "você" é bem avaliado, enquanto o "tu" tende a ser visto como mais adequado quando usado em relações simétricas. Tendo isso em consideração, podemos afirmar que as cantoras, para não serem julgadas de modo negativo, usam o pronome "você" frequentemente, como demonstra a tabela 1. Como as mulheres são historicamente mais cobradas a seguir as normas sociais do que os homens, sua expressão é julgada como reflexo da sua educação e da sua moral. Por outro lado, o uso que se faz dessas mesmas variantes pelos homens tende a ser visto como sinal de virilidade ou autenticidade.

Para analisar com maior detalhamento a questão do "tu", observamos a relação de concordância entre o pronome e o verbo no *corpus*, conforme tabela 2:

Tabela 2: concordância verbal em relação ao "tu"

Homem		**Mulher**	
Padrão	--	Padrão	--
	--		--
Não padrão	85/85	Não padrão	15/15
	100%		100%
Total	85/85	Total	15/15
	100%		100%

Fonte: Elaboração dos autores

A tabela 2 demonstra o quantitativo de concordância com a segunda pessoa do singular relacionada aos usos do pronome "tu", tanto de homens como mulheres. Evidenciamos que ambos os gêneros, ao longo de todas as músicas aqui analisadas, não realizaram a concordância verbal, conforme os parâmetros da GT (gramática tradicional). Nos termos de subsistemas pronominais propostos por Scherre *et al.* (2015), esses usos revelam um enquadramento no sexto subsistema pronominal, marcado pela alternância entre os pronomes "você" e "tu" sem concordância.

Esses resultados podem ser relacionados com os dados da pesquisa de Silva (2019), em que há uma predominância de não ativação da flexão verbal de segunda pessoa, no uso do pronome "tu" pelos entrevistados. Ao investigar os falares da variedade dos falantes de Rio Branco (Acre), a autora explica que a utilização de "tu" com concordância verbal de terceira pessoa consiste em um subsistema sem possibilidade de concordância, uma vez que, ainda que o pronome mantenha uma especificação formal e semântico-discursiva de segunda pessoa, quando ativa a concordância com o verbo, a concordância verbal corresponde à terceira pessoa.

O uso do "tu" sem concordância por parte de homens e mulheres pode ter ocorrido também pelo *corpus* elegido, o que sugere uma motivação de natureza funcional. O Poesia Acústica é um projeto conhecido como um rap de base acústica, que também incorpora elementos de *trap*, *reggae* e do *funk*, e, por isso, é híbrido em relação à modalidade, com características da fala e da escrita. Com isso,

nesse contexto, os usos de "tu" sem concordância se tornam mais esperados e menos marcados, tendo em vista a natureza discursiva desses gêneros musicais, geralmente associados à linguagem cotidiana, e ao hibridismo típico das músicas. A mistura de ritmos no Poesia Acústica não só reflete a cultura urbana, mas também um tipo de expressão mais próxima à realidade dos falantes.

Nesse cenário, a ênfase na autenticidade comunicativa faz com que essa forma de expressão se torne uma característica esperada e valorizada, refletindo a identidade do grupo e da familiaridade com o público jovem e periférico. Além disso, esse uso linguístico pode ser entendido como uma forma de resistência à norma-padrão, conferindo ao discurso uma dimensão de liberdade e subversão que está em sintonia com as temáticas de contestação e afirmação presentes nas músicas de Poesia Acústica, como discutido em Castanheira e Martins (2024), acerca do funk.

Então, as explicações para as formas não padrão de concordância nos usos do pronome "tu" em Poesia Acústica em ambos os gêneros podem ser, possivelmente, as características do *corpus*, formado por músicas que mesclam diferentes gêneros musicais, que, no geral, são mais próximos à fala cotidiana e à representação identitária típica do projeto musical aqui discutido e dos gêneros acionados, que tende a representar a ruptura com os modelos tradicionais de comportamento social e de usos gramaticais. Por isso, o "tu" sem concordância é menos marcado nesse contexto.

Outro ponto é que a tabela 2 deve ser analisada de modo integrado à tabela 1. Embora ambos os gêneros não utilizem o "tu" com flexão de 2ª pessoa, é preciso considerar que, no caso dos homens, o uso é mais comum do que das mulheres. Por ser algo mais marcado nesse contexto discursivo e também uma característica já comum no PB, seria improvável que qualquer cantor ou cantora de Poesia Acústica utilizasse formas, como "tu és". Nesse âmbito, o fato de as mulheres usarem o "tu" menos do que os homens está diretamente relacionado ao maior percentual de uso do "você", provavelmente para que o julgamento seja menor, já que esse é mais bem avaliado (Carvalho; Lopes, 2022).

Outro ponto analisado em nosso trabalho foi o tipo de referência. Vejamos, então, exemplos dos usos das formas por homens e mulheres em cada tipo de referência genérica e específica, respectivamente:

(a) Homem:

I. Referência genérica

(4) "Ret, boladão
Sem rodeio, só visão
Apenas faça
Tu pode não ter talento, mas raça é obrigação
Sem perdão, louco pra sempre
Tô deixando que os bobo se queime [...]" (Filipe Ret – Poesia Acústica 9)

(5) “De segunda a sexta eu faço referência
Só não me compara
Não há nada igual que **você** conheça
As leis estão aí, então desobedeça
Tudo que vai e passa vira experiência
Uma dose de sanidade na independência, fé.” (Hariel – Poesia Acústica 7)

(6) “No dia em que tudo perder o valor
Talvez cê me diga, me diga
O que é saudade
Sem que alguém lhe culpe
Ou que alguém lhe julgue
Eu vi a inveja comprar o homem
Pro homem comprar algo inútil
Chega a ser cômico, ilógico
Vê que **cê** pode ser o próximo [...]” (Bob – Poesia Acústica 4)

II. Referência específica

(7) “E se hoje eu te ligar, é melhor **tu** não atender
Já separei tuas coisas, não consegui devolver
Se quiser, tu vem buscar, sabe, eu vou te receber
Tudo aquilo p’a falar, tanta coisa que eu queria esquecer [...]”
(Derxan – Poesia Acústica 14)

(8) “E seja para mim só
Seja pra mim só (oh, fé)
Eu sei que **você** é só sua
Assim como eu sou só meu [...]” (Dudu – Poesia Acústica 6)

(9) “Quero sentir a sensação da primeira vez que te vi
Explicar pro coração que cê faz falta nem sempre dá
Te vejo nos melhores sonhos, tenho planos pra nós
Se te beijo, ouço o som dos anjos
Com seu rosto, sua voz, tô indo te buscar
Ter onde voltar, pra onde chegar, chegar
Cê me deixa sem ar
Pra nunca faltar motivos pra comemorar.” (MC Chris – Poesia Acústica 11)

(b) Mulher:

II. Referência específica

(10) "Melodia dos teus suspiros, sua voz no meu ouvido
Poder dizer que tá comigo, quero te apresentar
Pra minha mãe, levar a sogrinha pra jantar
Eu entreguei tudo que podia, tudo que até não tinha
Tudo que um dia me fez infeliz
E eu não tinha nada, nadava contra a maré
Fiz todo o caminho a pé, **tu** não me deixou desistir." (Elana Dara – Poesia Acústica 8)

(11) "Não vou falar de amor, **você** não merece
Me esquece, sai fora, eu tô tão de boa
E já passou da hora de você cuidar da sua vida
Só liga se for pra me pagar." (Marina Sena – Poesia Acústica 12)

(12) "Ah, se perguntarem o que é amor pra mim
Eu não sei responder, não sei explicar
Mas tudo fica sem graça, amor, se **cê** não tá." (Duquesa – Poesia Acústica 17)

Conforme indicam os exemplos acima, os homens utilizaram as três variantes no que tange aos dois tipos de referência, genérica e específica, já as mulheres utilizaram apenas a referência específica. A principal diferença entre as duas referências é que, a partir da referência genérica, primordialmente nesse contexto de músicas, é realizada uma atribuição a um referente de caráter mais amplo e geral, ao passo que, na referência específica, há uma designação de um referente determinado.

Em relação à referência genérica associada aos três casos de usos dos pronomes de segunda pessoa do singular, no caso dos homens, não há um referente específico ao qual os três cantores se dirigem, mas sim a qualquer interlocutor que possa ouvir a música e refletir sobre os temas ou ideias mencionados. A escolha do uso do pronome "tu", no trecho de autoria do Filipe Ret, "Apenas faça / **Tu** pode não ter talento, mas raça é obrigação" faz com que a mensagem se aplique a qualquer ouvinte ou pessoa que se identifique com a ideia de que, mesmo sem talento, a "raça" (esforço, persistência) é um requisito essencial.

Do mesmo modo, Hariel, a partir do uso do pronome "você", se refere a qualquer ouvinte que possa se identificar com a ideia de ser incomparável e com a crítica à padronização, em "Só não me compara / Não há nada igual que você conheça / As leis estão aí, então desobedeça" e, por último, Bob, utilizando o pronome "cê", atinge aquele interlocutor que possa ouvir a música e refletir sobre o tema da saudade, em "No dia em que tudo perder o valor / Talvez **cê** me diga, me diga / O que é saudade", e, em seguida, realiza uma crítica direta ao consumismo movido pela inveja e à ideia de que as pessoas compram coisas inúteis apenas para se encaixar em um padrão social ou para competir

com os outros, em “Eu vi a inveja comprar o homem / Pro homem comprar algo inútil / Chega a ser cômico, ilógico / Vê que **cê** pode ser o próximo.”

Em contrapartida, nos usos da referência específica por homens e mulheres, no caso do pronome “tu”, existe uma pessoa a que o pronome faz alusão, o parceiro romântico, como nos trechos “E se hoje eu te ligar, é melhor **tu** não atender / Já separei tuas coisas, não consegui devolver”, de autoria de Derxan, e “E eu não tinha nada, nadava contra a maré / Fiz todo o caminho a pé, **tu** não me deixou desistir”, de autoria de Elana Dara. No primeiro caso, tendo em vista o contexto em que a forma foi utilizada, a pessoa à qual o falante se refere foi alguém que provavelmente teve uma relação pessoal e íntima, provavelmente uma ex-namorada ou alguém com quem esteve envolvido emocionalmente, e parece haver uma tentativa de desapego. Já no segundo caso, o sujeito compartilha aspectos de um relacionamento que se tornou significativo o suficiente para ser integrado à sua vida familiar e afetiva.

Ainda sobre as referências específicas, no caso do pronome “você”, de forma similar, os cantores Dudu e Marina Sena aludem a pessoas com quem o eu possui uma relação direta. Nos trechos “Seja pra mim só (oh, fé) / Eu sei que **você** é só sua”, existe um tom de reflexão sobre a autonomia do outro e o desejo de exclusividade, mas com respeito pela individualidade da pessoa. Por outro lado, nos trechos de Marina Sena, “Não vou falar de amor, **você** não merece / Me esquece, sai fora, eu tô tão de boa”, há uma ênfase no término de uma relação, com a expressão de desaprovação e distanciamento.

Por, no caso do pronome “cê”, também a partir das referências específicas, podemos pensar que, tanto no trecho de MC Chris, como no de Duquesa, há uma referência a alguém que se ama e que se quer bem. Quando MC Chris afirma “Explicar pro coração que cê faz falta nem sempre dá / Te vejo nos melhores sonhos, tenho planos pra nós / Se te beijo, ouço som dos anjos”, na verdade, está ao mesmo tempo dizendo que ele tem urgência de um certo alguém, e que tudo que há de bom se encontra ali. Por outro lado, Duquesa afirma não saber exatamente explicar o sentimento amor, mas acaba resumindo-o em um alguém específico quando diz que “Mas tudo fica sem graça, amor, se **cê** não tá.”

Cabe examinarmos, então, a tabela 3, em que quantificamos as ocorrências desses dois tipos de referências relacionadas ao uso dos pronomes de segunda pessoa do singular em cada gênero:

Tabela 3: tipo de referência na variação pronominal de 2ª pessoa do singular

HOMEM			**MULHER**		
Tu	Específica	80/85	Tu	Específica	15/15
		94,12%			100%
	Genérica	5/85		Genérica	--
		5,88%			--
Você	Específica	98/101	Você	Específica	49/49
		97,03%			100%
	Genérica	3/101		Genérica	--
		2,97%			--

Cê	Específica	48/51	Cê	Específica	11/11
		94,12%			100%
	Genérica	3/51		Genérica	--
		5,88%			---

Fonte: Elaboração dos autores

A tabela 3 apresenta, relacionados à variação dos pronomes "tu", "você" e "cê", os percentuais dos dados relativos ao tipo de referência utilizada pelos homens e pelas mulheres: específica ou genérica. De maneira geral, percebemos que, nos usos do pronome "tu", com 94,12%, homens, e 100%, mulheres, bem como no uso do pronome "você", com 97,13%, homens, e 100%, mulheres, ou também no uso do pronome "cê", com 94,12%, homens, e 100%, mulheres, há uma tendência pelo tipo de referência específica. Tal ponto está relacionado a diversas questões, como exploraremos a seguir.

É importante destacar que, em relação aos homens, os percentuais são variáveis, mas, no caso das mulheres, são categóricos, o que demonstra uma diferença, e se alinha com a teoria das regras propostas por Labov (2003). Segundo o autor, as regras categóricas são aquelas que se manifestam de forma consistente, sem exceções, nos dados de fala, enquanto as regras variáveis são caracterizadas por alternâncias que dependem de fatores contextuais, sociais ou situacionais. Labov (2003) também destaca que uma estrutura linguística ser admitida como uma opção gramatical não significa que seja categoricamente usada ou que tenha um comportamento variável em todas as circunstâncias, tendo em vista que a quantidade de dados e os contextos específicos em que essas formas aparecem são essenciais para determinar a natureza da variação em uma língua.

Apesar de não haver dados sobre o uso de referência genérica em relação às mulheres, observamos que os homens utilizam ambas, mas muito mais a referência específica. Esse padrão corrobora os desdobramentos em Silva (2019), os quais fazem alusão a um emprego desses pronomes tanto no caso de uma referência genérica ou específica. Com base nos dados que evidenciam a variação entre "tu" e "você" na variedade rio-branquense, e considerando os dois tipos de referência, a autora enfatiza que o fenômeno é dependente de uma escolha funcional, já que o uso de uma variante ou outra dependerá da escolha do falante, baseada nos propósitos comunicativos que ele busca alcançar durante o processo de interação em que está envolvido.

Destacamos, ainda, que a referência específica aponta para um referente claramente identificado, facilmente reconhecido, ao passo que a referência genérica não faz alusão a um referente determinado ou específico. O uso elevado, em ambos os gêneros, da referência específica parece ter uma motivação funcional, relacionada ao *corpus*. A amostra muito se aproxima do diálogo, uma vez que são diversos cantores (homens e mulheres) que participam de cada uma das músicas, e esses artistas cantam e compartilham suas contribuições a partir de uma troca que se estabelece. Há, então, alternância de cantores, os quais trazem suas percepções, experiências e visões no tocante a diversos assuntos.

Desse modo, as músicas representam diálogos entre tais artistas, o que pode explicar o maior uso de referências mais específicas.

Já o uso da referência genérica, especificamente por parte dos homens, pode ter um valor discursivo. O funcionamento dessa referência, nesse *corpus*, pode estar atrelado à interlocução com o público; uma espécie de "conselho", uma voz que se dirige àquele que ouve. Inclusive, o fato de os homens utilizarem esse tipo de referência, e não as mulheres, parece se relacionar com uma questão de cunho social: a heteronormatividade, bastante evidente nas músicas de Poesia Acústica, predominantemente dominadas por homens.

Embora o projeto Poesia Acústica conte com algumas mulheres, a representatividade feminina ainda é pequena. Essa disparidade reflete questões de cunho social e discursivo: uma estrutura de poder desigual a qual implica restrição do protagonismo feminino, silenciando experiências e perspectivas valiosas. O uso da referência genérica por parte dos homens representa o reforço da ideia do homem como conselheiro do seu público. Esse cenário, portanto, corrobora a hipótese de que a desigual representatividade entre homens e mulheres influencia os usos dos pronomes de segunda pessoa do singular, refletindo padrões discursivos caracterizados por assimetrias de gênero, apesar do pouco quantitativo de dados.

Souza (2023) discute a influência dos gêneros rap, trap e funk sobre a imagem da mulher. A autora analisa algumas letras de música desses gêneros, sendo todas cantadas por homens, e analisa como há relações de poder em diversos trechos das letras. A falta de representatividade feminina é algo que chama a atenção da autora, bem como a objetificação feminina, corroborando a lógica de que a mulher é submetida a expectativas que buscam moldar seu comportamento, suas escolhas de vestuário, sua linguagem, sua postura e até mesmo companhias que mantém.

Sob esse viés, os homens estão, como figuras centrais e dominantes nas músicas de Poesia Acústica, ocupando um lugar de maior representatividade, voz e imagem de sabedoria na sociedade. Por vezes, ao longo das músicas, utilizam a referência genérica para orientar o público, a partir de suas vivências e experiências, reforçando a questão da identidade e da performance. Alguns trechos, como, por exemplo, "[...] no dia em que o mundo perder a cor, talvez você entenda a verdade, no dia em que tudo perder o valor […]" (Poesia Acústica 4, de autoria do cantor Bob), ou "[...] sem rodeio, só visão, apenas faça, tu pode não ter talento, mas raça é obrigação […]" (Poesia Acústica 9, de autoria do cantor Filipe Ret), demonstram esse teor de "conselho", a partir do uso da referência genérica, mais marcada nesse *corpus*.

Os resultados apresentados nesta seção confirmam a hipótese desta pesquisa, demonstrando como o uso da linguagem está, habitualmente, atrelado às diferentes expectativas sociais e culturais para os gêneros, refletindo as dinâmicas de poder e os seus papéis na sociedade. A variação pronominal entre "você", "tu" e "cê" com função de sujeito, então, está diretamente relacionada a questões identitárias e discursivas, já que há diferenças que se evidenciam pelos resultados qualitativos e quantitativos e que indicam que essa temática deve ser ainda mais explorada.

6. Considerações finais

O presente trabalho teve como objetivo investigar a variação dos pronomes de segunda pessoa do singular com função sintática de sujeito, "tu", "você" e "cê", nas músicas do projeto Poesia Acústica, a partir de uma perspectiva sociofuncionalista, observando especialmente o papel desempenhado pela variável gênero na distribuição dessas formas. Delineamos a análise a partir do *corpus* constituído por todas as músicas que compõem o projeto, nas quais foram identificadas, classificadas e interpretadas as ocorrências pronominais. A partir disso, buscamos compreender se o gênero dos artistas influi na escolha das variantes e de que maneira esse comportamento linguístico se manifesta nos versos analisados.

Ao examinarmos as músicas, observamos que as formas consideradas, "tu", "você" e "cê", ocorrem no *corpus*, ainda que com distribuições distintas, já que, apesar de o "você" ser o mais frequente, isso não ocorre de modo igualitário entre homens e mulheres. Os dados coletados permitiram-nos confirmar a predominância da forma "você" no conjunto total das ocorrências, acompanhando tendências já descritas para comunidades urbanas do PB, como evidenciado, dentre outros, por Scherre *et al.* (2015). Contudo, evidenciaram também a vitalidade do "tu" e a expressividade de "cê".

Ao analisarmos o papel da variável gênero, foco central deste estudo, observamos que a distribuição das variantes não é uniforme entre homens e mulheres. A tabela 1 revelou diferenças quantitativas em relação a homens e mulheres na variação pronominal, com uso mais expressivo de "você" nos trechos cantados por mulheres, e a tabela 2 demonstrou que os usos de "tu" são sempre sem concordância. Além disso, percebemos, nesse *corpus*, os usos dos pronomes pelas mulheres sempre com valor diretivo, ao contrário dos homens que também têm usos genéricos, possivelmente pela construção do seu papel social nas músicas como conselheiros e como grupo dominante.

Esses resultados permitem retomar a hipótese de pesquisa formulada neste trabalho. Partimos da hipótese de que há uma relação entre a variação dos pronomes de 2ª pessoa do singular com função de sujeito e a variável gênero, refletindo construções identitárias e discursivas. Nossa hipótese foi confirmada, porque, no *corpus* examinado, de fato, surgem diferenças na distribuição das variantes correlacionadas ao gênero. Esse efeito deve ser interpretado como um reflexo de práticas discursivas e de modos de construção identitária de homens e mulheres.

Enfatizamos, ainda, que, devido à limitação do *corpus*, não foi possível realizar outros cruzamentos a fim de averiguar, de maneira multifatorial e interseccional, questões raciais e de classe acerca das mulheres do projeto Poesia Acústica. Com poucas participantes, não houve como discutir, de modo integrado, tais pontos, mapeando possíveis outras influências. No entanto, já conseguimos atestar a relevância da variável gênero, o que representa um avanço para uma agenda de trabalho ainda pouco discutida e com necessária revitalização, como apontam os trabalhos recentes.

Destacamos, por fim, que o *corpus* escolhido constitui um material linguístico ainda pouco explorado por estudos variacionistas e que a abordagem sociofuncionalista adotada ao longo do trabalho foi especialmente produtiva para este tipo de análise. Ao considerar simultaneamente

questões linguísticas, sociais e discursivas, essa perspectiva permitiu observar não apenas a frequência das formas, mas também suas funções e valores no contexto de uso, destacando questões, como a gramaticalização do "você" e do "cê", a marcação situacionalmente analisada e a interface variação-discurso. Dessa maneira, o estudo conseguiu integrar a dimensão quantitativa e qualitativa, oferecendo uma dimensão mais ampla e consistente dos dados encontrados.

Referências

BECHARA, Evanildo. *Moderna Gramática Portuguesa***.** 37. ed. Rio de Janeiro: Nova Fronteira, 2009.

CASTANHEIRA, Dennis. Sociofuncionalismo: caminhos na interface variação-discurso. *Falange Miúda*, v. 3, pp. 87-95, 2018. Disponível em: https://periodicos.upe.br/index.php/refami/article/view/506. Acesso em: 04 abr. 2026.

CASTANHEIRA, Dennis; MARTINS, Yasmim. "Meu namorado é mó otário": análise sociofuncionalista de funks cantados por mulheres. *Revista E-scrita*, v. 15, pp. 155-168, 2024. Disponível em: https://revistasuniabeu.com.br/index.php/revista-e-scrita/article/view/72. Acesso em: 04 abr. 2026.

CARVALHO, Bruna Brasil de Albuquerque; LOPES, Célia Regina dos Santos. "Esse uso está adequado?": a percepção dos cariocas sobre as formas tu e você. *Organon*, Porto Alegre, v. 37, n. 73, pp. 59-79, 2022. Disponível em: https://seer.ufrgs.br/index.php/organon/article/view/122710/85280. Acesso em: 22 jun. 2025.

CASTILHO, Ataliba. *Nova Gramática do Português Brasileiro*. São Paulo: Contexto, 2014

COATES, J. Gender. *In*: LAMAS, C.; MULLANY, L.; STOKWEEL, P. *The Routhledge Companion of Sociolinguistics*. Nova York: Francis & Taylor, 2006, p. 62-68.

CUNHA, Celso; CINTRA, Lindley. *Nova Gramática do Português Contemporâneo*. 7ª ed. Rio de Janeiro: Lexikon, 2016 [1985].

CUNHA, Maria Angélica Furtado; COSTA, Marcos.; CEZARIO, Maria Maura. Pressupostos teóricos fundamentais. *In*: CUNHA, Maria Angélica Furtado da; OLIVEIRA, Mariangela Rios; MARTELOTTA, Mário Eduardo. (org). *Linguística funcional*: teoria e prática. Rio de Janeiro: DP&A, 2003, pp. 21-47.

FERREIRA, Ediene. O uso dos pronomes tu e você em textos orais na cidade de Santarém. *Margens*, v. 6, n. 10, pp. 279-293, 2010. Disponível em: https://periodicos.ufpa.br/index.php/revistamargens/article/view/2825. Acesso em: 01 jun. 2026.

FREITAG, Raquel. (Re)discutindo sexo/gênero na sociolinguística. *In*: FREITAG, Raquel; SEVERO, Cristiane. (org). *Mulheres, Linguagem e Poder* – Estudos de Gênero na Sociolinguística Brasileira. São Paulo: Blucher, 2015, pp. 17-74.

FRANCESCHINI, Lucelene Teresinha; LOREGIAN-PENKAL, Loremi. A variável sexo/gênero e o uso de tu/você no sul do Brasil. *Signum*: Estudos da Linguagem, [*S. l.*], v. 18, n. 1, pp. 182-205, 2015. Disponível em: https://ojs.uel.br/revistas/uel/index.php/signum/article/view/20205. Acesso em: 18 ago. 2025.

GIVÓN, Talmy. *Functionalism and grammar*. Amsterdam: John Benjamins, 1995.

GÖRSKI, Edair Maria; TAVARES, Maria Alice. Reflexões teórico-metodológicas a respeito de uma interface sociofuncionalista. *Revista do GELNE*, v. 15, pp. 75-97, 2013.

LABOV, William. Padrões sociolinguísticos. Tradução de Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008 [1972].

LABOV, William. Some sociolinguistic principles. *In*: PAULSTON, Cristina Bratt; TUCKER, G. Richard (eds.). *Sociolinguistics*: the essential readings. Oxford: Blackwell, 2003. pp. 235-250.

LOPES, Célia Regina dos Santos; COELHO, Izete; RUMEU, Márcia; SOUZA, Christiane Nunes de. História dos pronomes de tratamento no português brasileiro. *In*: MARTINS, Marco Antonio; OSÓRIO, Paulo; COELHO, Izete (org.). *História do português brasileiro*: trajetórias e perspectivas. Berlin; Boston: De Gruyter, 2025. pp. 161-196.

MARTINS, Ana Paula Antunes. O Sujeito "nas ondas" do Feminismo e o lugar do corpo na contemporaneidade. *Revista Café com Sociologia*, v. 4, n. 1, pp. 231-245, 2015. Disponível em: https://revistacafecomsociologia.com/revista/index.php/revista/article/view/443. Acesso em: 20 jan. 2026.

MOTA, Maria Alice. *A variação dos pronomes 'tu' e 'você' no português oral de São João da Ponte (MG)*. 2008. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2008. Disponível em: https://repositorio.ufmg.br/bitstream/1843/AIRR-7DHJPA/1/mariaalice_mota_diss.pdf. Acesso em: 12 ago. 2025.

NEVES, Maria Helena de Moura. *Gramática de usos do português*. 2. ed. São Paulo: Editora Unesp, 2011.

OUSHIRO, Livia. Interação entre sexo/gênero e classe social no uso variável da concordância verbal. *In*: FREITAG, R. M. Ko.; SEVERO, C. G. (org.). *Mulheres, linguagem e poder*: estudos de gênero na sociolinguística brasileira. São Paulo: Blucher, 2015. pp. 151-168.

PAIVA, Maria da Conceição. de. A variável gênero/sexo. *In*: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (org.) *Introdução à Sociolinguística* – o tratamento da variação. São Paulo: Contexto, 2003. pp. 33-42.

PAREDES SILVA, Vera Lúcia. O retorno do pronome tu à fala carioca. *In*: RONCARATI, Cláudia; ABRAÇADO, Jussara. (org.) *Português brasileiro*: contato lingüístico, heterogeneidade e história. Rio de Janeiro: 7 Letras, 2003, pp. 160-169.

RESENDE, G. M. *Variação dos pronomes tu x você, na fala bageense*. 2023. Trabalho de Conclusão de Curso (Graduação em Licenciatura em Letras – Português e Literaturas de Língua Portuguesa), Universidade Federal do Pampa, Bagé, 2023. Disponível em: https://repositorio.unipampa.edu.br/items/fc57e214-b43e-4c08-9977-c00ca382c781. Acesso em: 08 jan. 2026.

RIBEIRO, Djamila. *Quem tem medo do feminismo negro?* São Paulo: Companhia das Letras, 2018.

ROCHA LIMA, Carlos Henrique. *Gramática Normativa da Língua Portuguesa*. 49. ed. Rio de Janeiro: José Olympio, 2011 [1972].

SEVERO, Cristine Görski. O papel do gênero/sexo nos estudos sociolinguísticos de variação/mudança. *Revista de Letras*, Curitiba, v. 8, n. 1, pp. 1-8, 2006. Disponível em: https://dialnet.unirioja.es/descarga/articulo/6137777.pdf. Acesso em: 10 nov. 2025.

SCHERRE, Maria Marta Pereira; DIAS, Edilene Patrícia; ANDRADE, Carolina Queiroz; MARTINS, Germano Ferreira. Variação dos pronomes “tu” e “você”. *In*: MARTINS, Marco Antonio; ABRAÇADO, Jussara. (org.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015. pp. 133-172.

SILVA, Marinete Rodrigues da. *Tu e você na variedade rio-branquense*: um caso de variação ou de escolha funcional? 2019. Tese (Doutorado em Estudos Linguísticos), Universidade Estadual Paulista, São José do Rio Preto, 2019. Disponível em: https://bdtd.ibict.br/vufind/Record/UNSP_a95e81ed06b088524bfefe990fe2a659. Acesso em: 10 jan. 2026.

SILVA, Suziane de Oliveira Porto. Tu, você ou cê? Um olhar sociolinguístico sobre o pronome de segunda pessoa do singular no português brasileiro. *Revista JRG de Estudos Acadêmicos*, v. 8, n. 19, pp. 1-11, 2025. Disponível em: https://revistajrg.com/index.php/jrg/article/view/2208. Acesso em: 05 jan. 2026.

SOUSA, Valeria Viana; KANTHACK, Gessilene; SILVA, Camilo Rosa. Cartografia do Sociofuncionalismo no Nordeste. *In*: FREITAG, Raquel; AZEVEDO, Isabel Cristina. (org.). *Cartografia de pesquisas linguísticas no Nordeste*: áreas emergentes, aplicações para o ensino e interfaces. Campinas: Pontes, 2023, v. 3, pp. 129-156.

SOUZA, Evellyn Larissa Fernandes Benites de. *A objetificação da mulher no rap, trap e funk sob a visão da criminologia feminista e cultural*. 2023. Trabalho de Conclusão de Curso, Faculdade Plesbiteriana Mackenzie, São Paulo, 2023. Disponível em: https://dspace.mackenzie.br/items/7b177340-5d74-4795-836f-b2694272d0af. Acesso em: 20 jul. 2025.

TAVARES, Maria Alice. *A gramaticalização de E, AÍ, DAÍ e ENTÃO*: estratificação/variação e mudança no domínio funcional da sequenciação retroativo – propulsora de informações –um estudo sociofuncionalista. 2003. Tese (Doutorado em Linguística), UFSC, Florianópolis, 2003.

TAVARES, Maria Alice. Sociofuncionalismo: um duplo olhar sobre a variação e a mudança linguística. *Interdisciplinar: Revista de Estudos em Língua e Literatura*, v. 17, pp. 27-47, 2013. Disponível em: https://periodicos.ufs.br/interdisciplinar/article/view/1312. Acesso em: 20 jun. 2025.

TAVARES, Maria Alice; GÖRSKI, Edair Maria. Variação e sociofuncionalismo. *In*: MARTINS, Marco Antonio; ABRAÇADO, Jussara. (org.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015, p. 249-270.

TAVARES, Maria Alice; GÖRSKI, Edair Maria; PAZA, Carla Regina Martins. Gramaticalização: por uma perspectiva socialmente situada. *Estudos da Língua(gem)*, v. 23, n. 1, pp. 1-24, 2025. Disponível em: https://periodicos2.uesb.br/estudosdalinguagem/article/view/18345. Acesso em: 09 jan. 2026.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72232

Objetos acusativos de 3ª pessoa na variedade carioca do português brasileiro

Accusative objects of 3rd person in carioca variety of Brazilian Portuguese

Adriana Martins Simões[1]

RESUMO

Neste artigo, o nosso objetivo é apresentar os resultados de nossa pesquisa sobre o objeto acusativo anafórico de 3ª pessoa na variedade carioca do português brasileiro a partir da análise de entrevistas do PEUL. Tendo em vista as pesquisas de Duarte (1986), Cyrino (1994), Casagrande (2012), Kato, Martins e Nunes (2023) e as nossas pesquisas (Simões, 2015, 2022), partimos da hipótese de que haveria restrições ou impossibilidade de realização do pronome lexical com SNs [-animados; -específicos] indefinidos ou quantificados. Como referencial teórico, aliamos a teoria gerativista (Chomsky, 1981, 1986) à sociolinguística (Labov, 2008; Weinreich; Labov; Herzog, 2009). Confirmamos a nossa hipótese e propomos nossa análise gerativista desenvolvida em Simões (2015) para as construções com pronomes lexicais e objetos nulos na variedade carioca do PB.

PALAVRAS-CHAVE: Português brasileiro. Objeto acusativo anafórico. Variação linguística. Concordância de objeto. coexistência de gramáticas.

ABSTRACT

In this paper, our aim is to present the results of our research about the 3rd anaphoric accusative object in the carioca variety of Brazilian Portuguese based on PEUL interviews. Considering the studies of Duarte (1986), Cyrino (1994), Casagrande (2012), Kato, Martins e Nunes (2023) and our research (Simões, 2015, 2022), our hypothesis was that there would be restrictions or impossibility to lexical pronouns expression if nouns are [-animated; -specific] indefinite or quantified. Regarding the theoretical approach, generative grammar (Chomsky, 1981, 1986) and sociolinguist perspective (Labov, 2008; Weinreich; Labov; Herzog, 2009) were combined. We confirmed our hypothesis and we proposed the generative analysis developed in Simões (2015) on lexical pronouns and null objects constructions in the carioca variety of Brazilian Portuguese.

KEYWORDS: Anaphoric accusative object. Brazilian Portuguese. Linguistic variation. Object agreement. Coexistence of grammars.

Introdução

Em sua pesquisa diacrônica, Tarallo (1983, *apud* 2018 [1993]) observou, no português brasileiro (doravante PB), a diminuição na realização do objeto acusativo de 3ª pessoa a partir da 1ª metade do século XIX, que acentuou-se na 2ª metade desse mesmo século, como é possível observar na tabela 1.

[1] Universidade Federal de Alfenas (UNIFAL-MG), adriana.simoes@unifal-mg.edu.br, https://orcid.org/0000-0003-2911-8873.

Recebido em 20/01/2026 - Aceito em 19/06/2026

Tabela 1: Preenchimento do objeto direto pronominal

Data	% de preenchimento do objeto direto
1725	89,2%
1775	96,2%
1825	83,7%
1875	60,2%
1982	18,2%

Fonte: Adaptado de Tarallo (1983 *apud* 2018, p. 40 [1993])

Os dados diacrônicos de Cyrino (1990, *apud* Cyrino, 2018 [1993]) também revelaram um aumento na frequência de objetos nulos no mesmo período em que Tarallo (1983, *apud* 2018 [1993]) constatou uma diminuição na realização pronominal. Além disso, Cyrino demonstrou que, a partir da 2ª metade do século XIX, diminuiu a frequência dos pronomes clíticos e ocorreu o surgimento do pronome lexical em função acusativa.

Em sua pesquisa diacrônica, Cyrino (1994) verificou também que, no século XVI, os objetos nulos mais frequentes eram os oracionais. A partir do século XVIII, essa frequência se ampliou cada vez mais, chegando a 90% no século XX. Quanto aos objetos nominais, a autora encontrou 8,8% de objetos nulos com SNs [-específicos] e 2,9% com SNs [+específicos] no século XVI. Houve um aumento na frequência desses objetos nulos com SNs [-específicos] no século XVII e com SNs [+específicos] no século XIX. Contudo, nos dados da autora, é no século XX que há um aumento substancial na frequências dos objetos nulos, sendo de 86,1% para os SNs [-específicos] e de 67,4% para os SNs [+específicos]. Essa autora defende que a expansão dos objetos nulos no PB se iniciou pelos objetos oracionais. Além disso, como os dados demonstram, os objetos nulos sempre teriam sido possíveis no PB e a partir de um determinado momento aumentou a sua frequência nessa língua.

Galves (2018 [1993]) propõe que a inversão na retenção pronominal do PB revelaria uma mudança qualitativa em sua gramática e não apenas quantitativa na possibilidade de objetos nulos. Sendo assim, essa autora argumenta que teria ocorrido uma fixação paramétrica diferente no PB e que isso teria levado a gramática dessa língua a diferenciar-se da gramática do português europeu e da de outras línguas românicas. Galves (2001, 2018 [1993]) defende que esse fenômeno teria se originado do enfraquecimento da concordância verbal no PB, iniciada pela perda da distinção entre a 2ª e a 3ª pessoas, devido à inserção do pronome *você* no paradigma pronominal.

Kato, Martins e Nunes (2023) discutem diferentes propriedades sintáticas que seriam comuns a variedades do PB e do português europeu, incluindo as construções com objetos nulos, e concluem que, considerando a perspectiva de gramática como conhecimento internalizado, proposta por Chomsky (1981, 1986), o PB e o português europeu apresentariam gramáticas diferentes. Para esses autores, o pronome clítico já não integraria a gramática internalizada do PB, mas seria decorrente do processo de escolarização. Outros trabalhos anteriores também tecem uma análise nesse sentido. Duarte (1986) observou que apenas brasileiros escolarizados usam esse pronome. Magalhães (2006) constatou que

o clítico já não seria usado por crianças adquirindo o PB como língua materna. Além disso, Galves (2001) e Kato (2005) propõem que esse pronome integraria a periferia marcada, na medida em que seria resultado de aprendizagem formal.

Em Simões (2015)[2], realizamos uma análise sobre a expressão do objeto acusativo anafórico de 3ª pessoa nas variedades de espanhol de Madri e Montevidéu. Nossos resultados revelaram que os objetos nulos nessas variedades foram favorecidos pelos SNs sem determinante, os introduzidos por quantificadores, pelo artigo indefinido, os [-animados] e, na variedade de espanhol, também os [-específicos]. Além disso, observamos a ocorrência dessa categoria vazia em construções de aspecto [-perfectivo], perífrase verbal, clítico dativo, predicação secundária, tópico e verbos cognitivos, contextos estes em que os objetos nulos são possíveis em outras variedades do espanhol que apresentam objetos nulos de maneira mais ampla (Landa, 1993; Suñer; Yépez, 1988).

Em nossa pesquisa posterior (Simões, 2022), na qual investigamos entrevistas da variedade de espanhol de Madri cujos participantes tinham escolaridade equivalente ao Ensino Médio, observamos que os objetos nulos foram favorecidos pelos SNs sem determinante, os quantificados, construções com perífrase verbal, SNs [-animados], construções em que não ocorre predicação secundária e construções com verbos de estado.

Em nossa pesquisa (Simões, 2015), realizamos uma análise qualitativa entre as variedades do espanhol investigadas e o PB, tendo em vista nossa intuição de falante paulistana e observamos, mediante a tradução dos dados de objeto nulo dessas variedades ao PB, que a variação entre o objeto nulo e o pronome lexical seria possível entre os antecedentes mais referenciais, que são os [+/-animados; +específicos]. Contudo, entre os antecedentes [-animados; -específicos], a variação entre objeto nulo e pronome lexical já seria mais restringida, sobretudo entre os SNs quantificados, indefinidos e em construções com verbo de estado, nas quais a tendência seria aceitar apenas o objeto nulo. Em Simões (2022) investigamos a variedade do PB de São Paulo, cujos falantes apresentavam elevado nível de escolaridade, e observamos objetos nulos quase categóricos.

Neste artigo, nosso objetivo é apresentar e discutir a análise sociolinguística que realizamos sobre o objeto acusativo anafórico de 3ª pessoa na variedade do PB carioca a partir da análise de entrevistas sociolingüísticas do *Programa de Estudos sobre o Uso da Língua* (PEUL), desenvolvido na Universidade Federal do Rio de Janeiro (UFRJ). Nessa análise, investigamos falantes de três faixas etárias e de três escolaridades. Nosso objetivo foi realizar uma análise quantitativa para o PB, analisando os mesmos contextos linguísticos de nossas pesquisas anteriores sobre as variedades do espanhol e a variedade paulistana do PB (Simões, 2015, 2022). Isso porque nos interessa compreender melhor as semelhanças e diferenças entre as gramáticas de variedades dessas línguas, já que está em nossa perspectiva voltar a estudar o processo de aquisição/aprendizagem do espanhol por brasileiros (Simões, 2012).

[2] Parece-nos importante mencionar nossas pesquisas sobre as variedades de espanhol de Madri e Montevidéu justamente porque nossas pesquisas quantitativas sobre as variedades do PB nascem desse estudo sobre a língua espanhola e analisamos os mesmos contextos linguísticos para as duas línguas.

Este artigo apresenta as seguintes seções: na primeira, discutimos as características da gramática do PB quanto ao objeto acusativo anafórico de 3ª pessoa; na seção seguinte, abordamos o referencial teórico; na terceira seção, apresentamos a metodologia; na quarta seção, que se divide em subseções, apresentamos a análise sociolinguística e discutimos os resultados; na quinta seção, está a análise teórica que propomos para as tendências encontradas; para concluir, apresentamos as considerações finais.

1. O objeto acusativo anafórico na gramática do português brasileiro

Conforme vimos, os estudos de Tarallo (2018 [1993]) e Cyrino (1994) revelaram que houve no PB uma diminuição dos pronomes clíticos, uma ampliação na frequência dos objetos nulos e o surgimento do pronome lexical em função acusativa.

No âmbito do PB contemporâneo, Cyrino (1994) observou que os objetos nulos são mais frequentes com antecedentes [-animados] e, sobretudo, os [-animados; -específicos], cuja frequência em seu estudo foi de 93,1%. Quanto aos antecedentes [-animados; +específicos], a frequência foi de 86,5%. Em relação aos antecedentes [+animados; -específicos], a frequência foi de 57,1%. No entanto, essa autora não encontrou nos dados que analisou nenhuma ocorrência de objeto nulo com antecedente [+animado; +específico].

Em um estudo sociolinguístico, Duarte (1989) verificou um índice de 76,3% de objetos nulos com antecedentes [-animados] [exemplo (1a)] e 23,7% com antecedentes [+animados] [exemplo (1b). Observem-se as construções a seguir:

(1a) O Armando leu a peça e aprovou *(e)*. (novela) (123)

(1b) A FEBEM é um dos elos dessa corrente que cria o menor infrator; não é ela o único responsável, o único elo que cria *(e)*, e como tal ela não consegue recuperar *(e)*. (entrevista) (46,47) (Duarte, 1986, p. 16)

O estudo de Duarte (1989) revela que a maior possibilidade de objetos nulos relaciona-se com os SNs [-animados]. O estudo de Cyrino (1994) confirma essa tendência, que se potencializa combinada aos SNs [-específicos].

Duarte (1986) observou também que as construções que apresentam objeto indireto exibem alta frequência de objetos nulos, sendo de 93% quando o antecedente do objeto acusativo constitui um SN [exemplo (2)]. Contudo, em relação às construções com complemento predicativo, a autora encontrou 52,9% de objetos nulos e 40,4% de pronomes lexicais [exemplo (3a)], de modo que esse não seria um contexto de elevada frequência da omissão do objeto. Duarte observou também elevada frequência de objetos nulos com verbos no infinitivo, sendo esta de 71,5% [exemplo (3b)]. Observem-se as construções:

(2) (Seu dinheiro é todo seu?)

Não, eu dou *(e)* pra eles, mas eles não fazem muita questão não. (SP82-3-69)

(3a) Eu não tenho nada pra reclamar dela não. Eu acho ela sensacional! (SP85-29-707)

(3b) Eu queria ter uma irmã. Eu acho *(e)* tão bom! (SP85-28-661) (Duarte, 1986, p. 23)

Ao estudar a aquisição do PB como língua materna, Lopes (2007) observou que os objetos nulos anafóricos começam a aparecer na produção das crianças quando aparece também a distinção entre aspecto perfectivo e imperfectivo. Os resultados dessa autora parecem revelar a atuação do aspecto na possibilidade de manifestação de objetos nulos. Observe-se a construção em (4), na qual o aspecto é imperfectivo:

(4) Não vou guardar. (AC., 3;7)

(a criança se refere a seus brinquedos) (Lopes, 2007, p. 118)

Ao aplicar testes de julgamento de gramaticalidade a falantes adultos do PB, Casagrande (2012) observou que não seria possível a retomada pelo pronome lexical com antecedentes [+/-animados; -específicos], tanto com aspecto perfectivo quanto imperfectivo, como se observa nos exemplos em (5) e (6). Sendo assim, o objeto nulo seria a forma de expressão para esse tipo de antecedente no PB. Entretanto, se o antecedente for [-animado; +específico], já seria possível a sua expressão tanto pelo objeto nulo quanto pelo pronome lexical, como se observa em (7) e (8), e o pronome seria a variante preferida em construções perfectivas. Nesse caso, o pronome seria favorecido pela perfectividade do verbo. Por outro lado, conforme a autora, quando se trata de antecedente [+animado; +específico], apenas o pronome lexical seria possível com aspecto perfectivo ou imperfectivo, como se observa em (9) e (10).

(5) Minha tia faz **bolo** e eu como **Ø**/***ele** todo dia.

(6) Minha tia tem **filho** e educa **Ø**/***ele** com carinho. (Casagrande, 2010, p. 112-130)

(7) Eu tenho **uma bicicleta** e o meu pai sempre **lava Ø**/**ela**.

(8) Minha mãe comprou **as bananas** e meu pai comeu **Ø**/**elas**. (Casagrande, 2010, p. 120-137)

(9) Minha mãe encontra **o meu pai** e sempre beija ***Ø**/**ele**.

(10) Minha mãe encontrou **o meu pai** e beijou ***Ø**/**ele**. (Casagrande, 2010, p. 124-139)

Os resultados de Casagrande (2012) sugeririam que a expressão do objeto pelo pronome lexical estaria relacionada à especificidade do antecedente, uma vez que apenas os SNs [+específicos] permitem a retomada pelo pronome, enquanto os antecedentes [-animados] permitem também o objeto nulo e os [-específicos] se restringem à categoria vazia. Ao contrário dos estudos de Duarte

(1986) e Cyrino (1994), essa autora propõe que o fator determinante para a expressão pronominal do objeto acusativo seria a especificidade.

A partir da análise de dados de aquisição, Casagrande (2012) observou que haveria uma correlação entre a perfectividade e a expressão do objeto pelo pronome. Observa-se que essa tendência se assemelha à encontrada por Lopes (2007).

Cyrino (2020) afirma que os objetos nulos no PB seriam possíveis quando se trata de antecedentes [+animados; -específicos], como se observa em (11). Contudo, antecedentes [+animados; +específicos] não poderiam se expressar mediante objetos nulos, como se observa pela agramaticalidade da construção em (12). Diante desses dados, a autora conclui que, no PB, o objeto nulo teria a propriedade de expressar antecedentes [-animados], enquanto os traços de definitude/especificidade, quando marcados de maneira negativa, anulariam a restrição de animacidade, que se manifesta nos SNs [+definidos] e [+específicos].

(11) Então dar aula no curso de Engenharia é o degredo né, então degredar um professor, então, manda ___ lá pra Engenharia. (Freire, 2000, cap. 3. Ex. (69), *apud* Cyrino, 2020, p. 392)

(12) *A Maria encontrou o Pedro mas não beijou ___. (Cyrino, 2020, p. 392)

Kato, Martins e Nunes (2023) analisam as possibilidades de objetos nulos no PB e no português europeu, como na construção em (13), na qual o antecedente aparece no discurso, especificamente na construção anterior, próximo ao objeto nulo.

(13) Olha [que vaso bonito]i! Vou comprar **Ø**i para a Maria. (Kato; Martins; Nunes, 2023, p. 265)

Esses autores afirmam que, em um contexto como o de (14)[3], o objeto nulo apenas poderia ser interpretado como referindo-se ao *João* ou à pessoa a que o pronome *ele* se refere, devido ao fato de o objeto nulo tanto no PB quanto no português europeu apresentar a restrição quanto ao valor do traço [pessoa]. O objeto nulo poderia, portanto, referir-se à 3ª pessoa, que compreende o valor [-pessoa], mas não à 1ª e 2ª pessoas do discurso, que compreende o valor [+pessoa]. Observem-se as construções em (14):

(14) [Contexto: As pessoas estão discutindo como vão se distribuir nos carros para irem à festa.]

a. A: — E {eu/você/ele/o João}?

b. B: — A Maria leva **Ø**.
Ø = {*eu/*você/ele/o João} (PE/PB)

b'. B: — A Maria leva {**-me/-te/-o**}. (PE)

b''. B: — A Maria {{**me/te**} leva / leva **ele**}. (PB) (Kato; Martins; Nunes, 2023, p. 270)

[3] Exemplos renumerados em relação ao texto original de Kato, Martins e Nunes (2023).

Kato, Martins e Nunes (2023) afirmam também que o objeto nulo poderia ter um antecedente [-humano], como ocorre em (15)[4]. Conforme esses autores, em um contexto como o de (15), o objeto nulo seria a forma preferida para a expressão do objeto anafórico tanto no PB quanto no português europeu.

(15) [Contexto: As pessoas estão discutindo a distribuição da comida para o piquenique.]

a. A: — E o bolo?

b. B: — A Maria leva **Ø**. (PE/PB)

b'. B: — ?A Maria leva-o. (PE)

b''. B: — ?A Maria leva **ele**. (PB) (Kato; Martins; Nunes, 2023, p. 271)

Kato, Martins e Nunes (2023, p. 270) argumentam que " (…) a interpretação de objetos nulos como [-hum] é muito menos restrita do que sua contraparte [+hum], pois parece não haver configuração estrutural onde um objeto nulo pode ser interpretado como [+hum], mas não como [-hum]". Essa característica revela o que os trabalhos com dados empíricos de Duarte (1986) e Cyrino (1994) já mostravam essa tendência em relação aos SNs [-animados], que correspondem ao contexto de maior frequência dos objetos nulos. Além disso, para Kato, Martins e Nunes (2023), a variação nos julgamentos de aceitabilidade de SNs [+humanos] tende a ser maior quando se trata de objeto nulo e a tendência seria os falantes preferirem a expressão por um pronome nesses casos, sendo a expressão do objeto realizada pelo pronome lexical no PB e pelo clítico no português europeu.

No âmbito do PB, segundo Kato, Martins e Nunes (2023), essa língua permite objetos nulos [+/-humanos] quando o antecedente aparece no discurso, como vimos em (14) e (15), e também quando o antecedente aparece como um tópico, como vemos em (16) e (17). Contudo, quando o antecedente aparece na posição de sujeito ou em outra função sintática, o objeto nulo se restringiria aos SNs [-humanos], como se observa em (18) e (19). Esses autores afirmam também que em contextos de ilha[5] as restrições aumentam, pois, nesses casos, apenas os SNs [-humanos] seriam possíveis em todos os contextos e os [+humanos] apenas nas construções de tópico. Embora esses autores considerem agramatical a construção em (21a)[6], na qual o objeto nulo teria como antecedente o SN [+humano] *esses alunos*, em nossa intuição essa construção seria bem formada.

(16a) [**Este livro**]i, todos os meus amigos leram **Ø**i.

(16b) [**Este livro**]i, todos os meus amigos leram **ele**i.

(17a) [**O Obama**]i, o João disse que todos queriam cumprimentar **Ø**i.

(17b) [**O Obama**]i, o João disse que todos queriam cumprimentar **ele**i. (Kato; Martins; Nunes, 2023, p. 275)

[4] Exemplos renumerados em relação ao texto original de Kato, Martins e Nunes (2023).

[5] Os contextos de ilhas correspondem a orações relativas, sujeitos sentenciais e orações adverbiais.

[6] Exemplos renumerados em relação ao texto original de Kato, Martins e Nunes (2023).

(18a) [**Esse prato**]i exige que o cozinheiro acabe de preparar **Ø**i na mesa.

(18b) [**Esse prato**]i exige que o cozinheiro acabe de preparar **ele**i na mesa. (Kato; Martins; Nunes, 2023, p. 273)

(19a) *[**Esse juiz**]i não gosta que os advogados contradigam **Ø**i.

(19b) [**Esse juiz**]i não gosta que os advogados contradigam **ele**i. (Kato; Martins; Nunes, 2023, p. 274)

(20a) O João trouxe [**os documentos**]i e deixou **Ø**i em cima da mesa.

(20b) O João trouxe [**os documentos**]i e deixou **eles**i em cima da mesa. (Kato; Martins; Nunes, 2023, p. 273)

(21a) *O João recomendou [**esses alunos**]i, mas ninguém contratou **Ø**i.

(21b) O João recomendou [**esses alunos**]i, mas ninguém contratou **eles**i. (Kato; Martins; Nunes, 2023, p. 274)

Em nossa pesquisa Simões (2015)[7], verificamos que, quando se trata de antecedente [+animado; +específico], com SN definido ou indefinido, seria possível a variação entre o objeto nulo e o pronome lexical, como nas construções em (22). Essa tendência se manifesta também com os antecedentes [+animados; -específicos], sendo o SN definido ou quantificado, como se observa nas construções em (23).

(22a) (…) **este** espera que eu te mostro **Ø**/**ele** (...) (referente: **o filho da informante**)
(*cf*. entrevista 16 – Madri)

(22b) **uma amiga minha** eu chamei **Ø**/**ela** quando ela estava vendo TV para dedicar um vídeo para ela (...)
(*cf*. entrevista 5 – Madri)

(23a) porque **a polícia** nós não chamamos **Ø**/**ela** porque eles não tinham entrado... não tinham roubado nada (...)
(*cf*. entrevista 18 – Montevidéu)

(23b) A: e você considera que tem **alguns amigos**?
B: sim... E é uma
A: ah... sim
B: está bom então
A: por sorte eu encontrei **Ø**/**eles**/**alguns**... viu? (...) (*cf*. entrevista 2 – Montevidéu) (Simões, 2015, p. 337 e 339)

No que se refere aos antecedentes [-animados; +específicos], também ocorreria variação entre o objeto nulo e o pronome lexical quando os antecedentes são SNs definidos e indefinidos, como em (24a) e (24b). Contudo, com um SN quantificado, como em (24c), apenas o objeto nulo seria possível.

[7] Conforme vimos na introdução, essa análise que fizemos sobre o PB integrou nossa tese de doutorado (Simões, 2015), na qual investigamos a expressão do objeto acusativo anafórico de 3a pessoas nas variedades de espanhol de Madri e Montevidéu e, considerando os dados de objetos nulos encontrados, tecemos uma comparação com o PB mediante a tradução desses dados ao PB, tendo em vista nossa intuição de falante paulistana. Nosso objetivo era verificar as possibilidades de objetos nulos e pronome lexical no PB, observando a variação ou a restrição de uma das variantes em um dado contexto.

(24a) (…) eu não sei do que é **esse sal**... me deram **Ø**/**ele**... eu tenho um pouco aí (...) (*cf.* entrevista 13 – Madri)

(24b) com uma barrinha de ferro que introduzia em **uma tábua**... eu tenho **Ø**/**ela** lá naquele quarto... depois eu te mostro **Ø**/**ela** (...) (*cf.* entrevista 12 – Montevidéu)

(24c) (…) eu levava no bolso **dois mil e alguma coisa**... se eu chego a dar **Ø**/***ele** para o meu filho... o cara sai frustrado de tudo... entende? (...) (*cf.* entrevista 16 – Madri) (Simões, 2015, p. 338-339)

Quanto aos antecedentes [-animados; -específicos] e definidos, a variação entre objeto nulo e o pronome lexical seria possível nas construções com verbos dinâmicos, como se observa em (25a) e (25b). Entretanto, se for um verbo de comunicação, a construção não seria gramatical com o pronome, como se observa em (25c). Em relação às construções com verbos de estado, como *querer*, o pronome torna o antecedente [+específico], como se observa em (25d).

(25a) (…) no arroz... por exemplo... n**o arroz**... primeiro eu refogo **Ø**/**ele** com óleo e cebola (...)
(*cf*. entrevista 8 – Madri)

(25b) **as estações** eram mais definidas... né? absolutamente... as pessoas associavam **Ø**/**elas** aos jogos (...)
(*cf*. entrevista 13 – Montevidéu)

(25c) A: **esses dados** interessam pra vocês?
B: sim
A: porque talvez é como uma novela
B: nos interessam e a gente adora
A: pois então eu te conto **Ø**/?**eles** (*cf*. entrevista 18 – Madri)

(25d) (…) recebemos **os apartamentos** e não era possível dizer "eu quero **Ø**/**ele** nesta rua nesta altura" (...) *(cf.* entrevista 18 – Madri) (Simões, 2015, p. 340-341)

Por fim, entre os antecedentes [-animados; -específicos] e indefinidos, apenas o objeto nulo seria gramatical na maioria das construções. Em (26a), o pronome lexical conduziria a uma interpretação [+específica] do SN antecedente. Nessa construção, aparece o verbo intensional *pedir*. Em relação a (26b), o pronome seria gramatical com o verbo dinâmico *fazer*, mas agramatical com o verbo *comprar.* Nessa última construção, o antecedente aparece em posição de tópico. Na construção com o verbo estativo *ter* o pronome também não é possível, como se observa em (26c). Quanto aos SNs quantificados, apenas em (26e) a construção é gramatical com o pronome.

(26a) A: (…) como eu peço **uma bolsa**… um horário... documentos... e vou
B: e vai pra fora
A: estou ocupada com A
B: o único que você perde... mas... você perde a bolsa ou a ajuda de...
A: não... mas eu não posso pedir **Ø**/?**ela**... eu é que não me atrevo (...) (*cf.* entrevista 18 – Madri)

(26b) (...) minha mãe não comprava **um sorvete** na sorveteria... "porque Miguel nos ajuda e você sabe que não... eu te faço **Ø**/**ele** em casa do que você quiser... de chocolate... de creme... mas **um sorvete de sorveteria** eu não posso te comprar **Ø**/?**ele**" (...) (*cf*. entrevista 14 – Montevidéu)

(26c) A: e com ela você tem **uma boa relação**?
B: como eu tinha **Ø**/?**ela** com o meu irmão... como eu tenho **Ø**/?**ela** com a minha cunhada (...) (*cf*. entrevista 8 – Montevidéu)

(26d) (…) depois que está **tudo** frito... eu misturo **Ø**/?**ele**... agrego **Ø**/?**ele** à massa (...) (*cf*. entrevista 3 – Madri)

(26e) A: (…) pois eu tenho **muitos exames muitos atestados de incapacidade** que refazer é muito difícil
B: é verdade... é uma confusão... vai muito tempo
A: e só eu poderia fazer **Ø**/**eles** / porque meus filhos não fizeram **Ø**/**eles** (*cf*. entrevista 18 – Madri)

(26f) A: e você comprou **alguma rifa**?
B: não… mamãe compra **Ø**/?**ela**/**alguma** na de arquitetura (*cf*. entrevista 20 – Montevidéu) (Simões, 2015, p. 341-342)

As tendências observadas em Simões (2015) nos permitiram concluir que, no PB, a possibilidade de expressão do objeto pelo pronome lexical se relacionaria com os antecedentes [+animados] e [+específicos], enquanto entre os antecedentes [-animados], apenas os [+específicos] definidos ou indefinidos aceitariam o pronome lexical em todas as construções. Esses resultados são um reflexo da Hierarquia Referencial de Cardinaletti e Starke (1994), já que as entidades [+animadas] e/ou [+específicas] tenderiam a expressar-se pelo pronome. Além disso, detectamos que haveria restrições relacionadas à estrutura do SN e à semântica do predicado verbal.

Em nosso estudo posterior (Simões, 2022), investigamos o objeto acusativo anafórico de 3ª pessoa na variedade do português brasileiro de São Paulo, mediante a análise de entrevistas do *Projeto História do Português Paulista* (Lima-Hernandes; Vicente, 2012). Ao analisar os dados, encontramos uma ocorrência quase categórica de objetos nulos, cinco ocorrências do pronome clítico e nenhuma do pronome lexical. Tendo em vista esses resultados, não pudemos usar o programa estatístico *Goldvarb X* para a realização da análise multivariada, obtenção dos pesos relativos e verificação dos contextos que favoreceriam o objeto nulo ou o pronome lexical. A ocorrência quase categórica de objetos nulos revelaria um estágio avançado de implementação dessa variante na variedade de São Paulo do PB contemporâneo. Além disso, o fator escolaridade também precisa ser considerado, já que a amostra analisada apresenta falantes com elevada escolaridade. No intuito de verificar as possibilidades do pronome lexical, optamos por verificar para cada ocorrência de objeto nulo e clítico se seria possível o pronome lexical, tendo em vista nossa intuição de falante nativa e paulistana do PB[8].

[8] Como existem limitações na utilização de dados da própria intuição nas pesquisas (*cf*. Kenedy, 2015), pretendemos complementar nosso estudo com um teste de aceitabilidade que será submetido a falantes nativos do PB e aplicado durante a pesquisa sobre a aquisição/aprendizagem do espanhol como língua estrangeira que iremos realizar.

Nessa pesquisa, observamos a ocorrência de objetos nulos referindo-se a SNs com os diferentes tipos de determinantes, sendo a manifestação dessa variante categórica no caso dos determinantes quantificados ou SNs sem determinante. Quanto ao pronome lexical, são justamente os SNs quantificados e, sobretudo, os sem determinante os contextos de restrição a essa variante.

Em relação ao traço de animacidade, houve 98.8% de objetos nulos com antecedentes [-animados] e 75% com antecedentes [+animados], tendência que corrobora os estudos de Duarte (1986) e Cyrino (1994). No caso do pronome lexical, observamos 83,3% de construções gramaticais com essa variante com SNs [+animados] e 56,2% com SNs [-animados].

No que se refere ao traço de especificidade, encontramos 94,2% de objetos nulos com SNs [+específicos] e 98,4% com SNs [-específicos]. Na esfera do pronome lexical, verificamos que essa variante seria gramatical em 94,2% dos SNs [+específicos] e apenas em 43,4% dos SNs [-específicos]. Esse resultado para o pronome lexical confirma o estudo de Casagrande (2012), no qual esse pronome estaria associado a antecedentes [+específicos].

Cruzamos os fatores de animacidade e especificidade e encontramos 98% de objetos nulos com os antecedentes [-animados; +específicos], 67% com os [+animados; +específicos], 99% com os [-animados; -específicos] e 83% com os [+animados; -específicos]. Esses resultados confirmam que essa variante está fortemente relacionada ao traço [-animado] do antecedente, como nos trabalhos de Duarte (1986) e Cyrino (1994). Por outro lado, os objetos nulos que se referem aos antecedentes [+animados; +específicos] divergem da análise de Cyrino (2020), pois, conforme a proposta dessa autora, haveria uma restrição à categoria vazia quando se trata de antecedente com esses traços semânticos. Quanto ao pronome lexical, consideramos gramaticais 93% das construções em que essa variante retoma antecedentes [-animados; +específicos] e todas as construções em que os antecedentes são [+animados; +específicos]. Entretanto, consideramos gramaticais apenas 42% das construções em que os antecedentes são [-animados; -específicos] e 67% daquelas em que os antecedentes são [+animados; -específicos]. Essas tendências reforçam a relação entre antecedentes [+específicos] na possibilidade do pronome lexical, como no estudo de Casagrande (2012), e que se amplia quando esse traço se combina com o [+animado].

Vejamos algumas construções desse estudo (Simões, 2022). Em (27a), há um SN definido [+animado; +específico], em (27b), um SN definido [-animado; +específico], em (27c), um SN definido [-animado; -específico], em (27d), um SN indefinido [-animado; -específico] e, em (27e), um SN quantificado [-animado; +específico]. Destacamos a construção em (27a), na qual o objeto nulo refere-se a um antecedente [+animado; +específico] e aparece o verbo cognitivo *conhecer*, que constitui um dos contextos que permitem objetos nulos [+animados] na variedade basca do espanhol (*cf*. Landa, 1995).

(27a) L3 e o senhor conheceu **sua esposa** em toda essa jornada quando?
L1 eu conheci... bom... aí primeiro eu me casei muito jovem não é minha filha sabe não é... me casei com quarenta e cinco quase quarenta e seis anos não é... eu conheci **Ø** num curso que eu fui fazer no Rio de Janeiro [...] (Informante 7)

(27a') ... e o senhor conheceu **sua esposa** em toda essa jornada quando?
... eu conheci... bom... aí primeiro eu me casei muito jovem não é minha filha sabe não é... me casei com quarenta e cinco quase quarenta e seis anos não é... eu conheci **ela** num curso que eu fui fazer no Rio de Janeiro

(27b) L2 [...] você acaba não achando muito tempo mas eu gostaria muito tenho ainda **esse plano** apesar de já tá já devia estar praticando **Ø** tenho ainda esse plano (não teria tempo) (Informante 10)

(27b') ... mas eu gostaria muito tenho ainda **esse plano** apesar de já tá já devia estar praticando **ele** tenho ainda esse plano (não teria tempo)

(27c) L1 [...] eu eu encaro sempre como encomenda... mesmo **as canções que eu faço pros meus discos**... também são encomendas eu tenho mais prazer de faze **Ø** [...] (Informante 6)

(27c') ... mesmo **as canções que eu faço pros meus discos**... também são encomendas eu tenho mais prazer de faze **elas**

(27d) L1 [...] do que que cria a compatibilidade entre melodia e... e letra... aí é **um estudo que dá pra fazer em pós-graduação**... não posso introduzir **Ø** em graduação porque... não faz parte do programa ((sorriu))... [...] (Informante 6)

(27d') ... aí é **um estudo que dá pra fazer em pós-graduação**... não posso introduzir **ele** em graduação

(27e) .....((L1 voltando)) eu não encontro a nova fase já saíram **dois números da nova fase** mas eu fiquei tão decepcionado que eu acho que eu não sei se eu joguei **Ø** fora... ((risos)) [...] (Informante 5)

(27e') ... já saíram **dois números da nova fase** mas eu fiquei tão decepcionado que eu acho que eu não sei se eu joguei **eles** fora... ((risos)) (Simões, 2022, p. 1582)

Em relação aos outros contextos linguísticos observados, verificamos que os objetos nulos apresentam alta implementação nas construções com verbos cognitivos, com verbos de comunicação e demais verbos, em construções com aspecto [+/-perfectivo], com verbos dinâmicos e estativos, em construções com ou sem elementos intensionais, em construções com ou sem objeto indireto, em construções com perífrase verbal, em construções com predicação secundária e em construções com tópico e outros contextos discursivos. Merece destaque o fato de que essa variante foi categórica nas construções com verbos de comunicação, aspecto perfectivo, objeto indireto, perífrase verbal e quando o antecedente é um tópico. Como vimos, Duarte (1986) já havia detectado uma elevada frequência de objetos nulos em construções com objeto indireto. Em relação às construções com tópico, o estudo de Kato, Martins e Nunes (2023) destacam estas como um importante contexto na possibilidade de manifestação dos objetos nulos.

Em relação ao pronome lexical, consideramos gramaticais 69,2% das construções com verbos cognitivos. Contudo, consideramos agramaticais 60% das construções com verbos de comunicação. Também consideramos gramaticais 90% das construções com aspecto perfectivo e apenas 54% das

construções com aspecto imperfectivo. Quanto ao aspecto lexical, consideramos gramaticais 62,2% das construções com verbos dinâmicos e 57,6% das construções com verbos estativos. Consideramos gramaticais 56,6% das construções com elementos intensionais, 41,2% das construções com objeto indireto, 75% das construções com predicação secundária, 59,6% das construções com predicação secundária, 47,1% das construções com tópico e 59,1% das construções com outros contextos discursivos. Em resumo, considerando nossa intuição, o pronome lexical seria melhor aceito nas construções com verbo de aspecto perfectivo, verbos dinâmicos e predicação secundária e teria baixa aceitação nas construções com predicados estativos, de comunicação, objetos indiretos e tópico. Vemos, portanto, que nosso estudo confirma a tendência encontrada em Casagrande (2012) sobre a relação entre aspecto perfectivo e pronome lexical, em Duarte (1986) sobre as construções com predicação secundária e a baixa aceitação em contextos de objeto indireto e tópico, que, como vimos, compreendem contextos de alta implementação do objeto nulo.

Vimos, nesta seção, que o estudo de Cyrino (1994) revela que os objetos nulos no PB seriam muito frequentes com os antecedentes [-animados] e, sobretudo, com os [-animados; -específicos]. Duarte (1986) encontrou também frequência elevada de objetos nulos com SNs [-animados], além de quando há objeto indireto e verbo no infinitivo. Por outro lado, nas construções com predicação secundária os objetos nulos não seriam tão frequentes, sendo um contexto importante para a ocorrência do pronome lexical. Vimos também que o trabalho de Casagrande (2012) revelou a relação entre os antecedentes [+específicos] e o pronome lexical, sendo que apenas o pronome lexical seria possível com os antecedentes [+animados; +específicos] e os [-animados; -específicos] se restringiriam a ocorrer com objetos nulos. Esses resultados de Casagrande (2012) estão em consonância com os trabalhos de Cyrino (1994, 2020) quanto à restrição dos objetos nulos aos SNs [+animados; +específicos]. Contudo, como vimos, embora Kato, Martins e Nunes (2023) demonstrem que há mais restrições para o objeto nulo quando se trata de antecedentes [+humanos], esses autores demonstram também a possibilidade de objetos nulos tanto com antecedentes [-humanos] quanto [+humanos][9] em contextos em que o antecedente aparece próximo no discurso ou como tópico, havendo restrições para os antecedentes [+humanos] que aparecem como sujeito ou objeto. Além disso, esses autores demonstram haver uma restrição aos objetos nulos em referência à 1ª e 2ª pessoas do discurso, mas não à 3ª, que pode manifestar-se com antecedente [-/+humano]. Por fim, nossos estudos (Simões, 2015, 2022) revelaram que haveria restrições ao pronome lexical relacionadas com a referencialidade do antecedente e sua estrutura, assim como com a semântica do predicado verbal da construção, objeto indireto e tópico. Na seção dedicada à discussão dos resultados de nossa pesquisa sobre a variedade carioca do PB, veremos o quanto todos esses trabalhos dialogam com essas tendências e análises.

[9] Ao contrário de Duarte (1986), Cyrino (1994, 2020) e Casagrande (2012), Kato, Martins e Nunes (2023) utilizam a classificação dos antecedentes em termos de [+/-humanos], que seria mais restringida que a classificação em [+/-animados], que abrange também os animais.

2. Referencial teórico

Consideramos, em nossas pesquisas, a concepção biológica de língua (Chomsky, 1981, 1986) aliada à concepção social (Labov, 2008; Weinreich; Labov; Herzog, 2009). Sendo assim, adotamos a perspectiva de gramática como um conhecimento internalizado na mente/cérebro dos falantes, além de considerarmos também alguns aspectos da sociolinguística, que nos permitem observar as propriedades dessa gramática internalizada.

Embora essas perspectivas teóricas apresentem bases epistemológicas diferentes e tenham diferenças em relação ao lugar das regras variáveis no sistema linguístico, seriam também complementares. A *Sociolinguística Paramétrica*, que nasce da junção dessas duas teorias, tem início na década de 80 no Brasil. Kato e Tarallo (1989) foram seus expoentes e conduziram uma pesquisa pioneira demonstrando a relevância de aliar as perspectivas intralinguística e interlinguística, já que a primeira fornece subsídios para uma melhor compreensão e desenvolvimento da segunda.

2.1. Gerativismo

Chomsky (1981, 1986) propõe que os seres humanos teriam em sua mente/cérebro uma capacidade linguística inata, denominada *Faculdade da Linguagem*. Seria mediante esse dispositivo que o conhecimento linguístico seria desenvolvido. A partir da interação com os dados linguísticos do ambiente, a criança adquirindo a sua língua materna fixa os parâmetros na *Gramática Universal*, que abriga também os princípios das línguas naturais.

O conhecimento linguístico constitui a língua-I, que tem como características ser internalizada, intensional e individual. A língua-I opõe-se à língua-E, que tem como características ser externa e extensional. Chosmky também diferencia a competência do falante do desempenho linguístico. Para ele, a primeira corresponde à gramática internalizada, enquanto o segundo corresponde à língua em uso.

Como resultado da fixação de parâmetros, o falante teria uma *gramática nuclear*, cuja origem advém do processo de aquisição da linguagem desde o nascimento. Contudo, essa gramática nuclear se expande e surge a *periferia marcada*, responsável por abrigar resíduos de mudança linguística e outros fenômenos.

Vimos que o clítico do PB já não se incorpora à gramática mediante aquisição natural (Magalhães, 2006), mas via aprendizagem formal, ao longo do processo de escolarização (Duarte, 1986). Devido a isso, Galves (2001) e Kato (2005) propõem que esse pronome integraria a periferia marcada. Por outro lado, o objeto nulo e o pronome lexical compreendem aquisição natural.

Lightfoot (2006) propõe que a língua-E seria o reflexo do *output* das gramáticas das comunidades linguísticas e do uso da língua no discurso e na variação social. Para nossa pesquisa, analisamos entrevistas sociolinguísticas, que correspondem à língua-E. Contudo, os objetos nulos e o pronome lexical correspondem à língua-I como conhecimento linguístico internalizado. Sendo assim, embora as entrevistas representem a língua em uso, nosso estudo visa a análise do conhecimento linguístico internalizado, a competência linguística, em oposição ao desempenho linguístico, que não foi considerado na análise.

2.2. Sociolinguística

Labov (2008) propõe que os sistemas linguísticos se caracterizam por serem heterogêneos. Isso significa que as línguas apresentam variação na medida em que se inserem em comunidades de fala. Para que haja variação é necessário que existam duas ou mais formas linguísticas, as variantes, que seriam equivalentes em significado ou valor funcional e que são compartilhadas pelos membros de uma comunidade de fala, embora cada membro apresente frequências diferentes dessas formas em sua produção.

A variação linguística ocorre porque não seria possível que a variante conservadora fosse substituída pela variante inovadora de maneira instantânea. No entanto, existem fenômenos de variação que ocorrem durante um longo período de tempo e que não culminam em mudança linguística. Nesses casos, se manifestaria o fenômeno de variação estável.

Para Weinreich, Labov e Herzog (2009), a variação e mudança linguísticas envolvem os seguintes princípios: (1) os fatores condicionantes, que compreendem os contextos linguísticos e sociais que atuam para favorecer uma variante ou outra; (2) a transição, que demonstra a mudança linguística a partir do estudo de diferentes faixas etárias e permite a observação da coexistência de duas variantes até a obsolescência de uma delas; (3) o encaixamento, que está relacionado às estruturas linguística e social e corresponde à alteração gradual das variantes no sistema linguístico e nos contextos sociais; (4) a avaliação, que corresponde ao juízo de valor, à significação social que o falante atribui a uma variante, e constitui uma questão que pode atuar na implementação ou retrocesso de uma mudança; (5) a implementação, que também está relacionada à esfera linguística e à social. Em relação à esfera linguística, tendo em vista que a mudança estaria encaixada no sistema linguístico, a expansão de uma variante inovadora ocorreria gradualmente. Contudo, a estrutura social poderia provocar interferências na expansão de uma variante, na medida em que a significação social da mesma pode atuar em sua implementação ou em seu retrocesso.

Em nossa pesquisa, nosso objetivo foi investigar a variação linguística no âmbito da expressão do objeto acusativo anafórico de 3ª pessoa na variedade carioca do PB, verificando os condicionadores linguísticos e sociais que estariam relacionados ao pronome lexical, enquanto uma das variantes inovadoras, bem como o seu encaixamento e implementação no sistema.

3. Metodologia

Para a realização da pesquisa, analisamos 21 entrevistas da variedade carioca do PB, oriundas do PEUL - UFRJ[10]. Como variável, estudamos a expressão do objeto anafórico acusativo de 3ª pessoa e, como variantes, o objeto nulo, o pronome lexical e o pronome clítico.

Nossa pesquisa parte da hipótese de que haveria restrições ou impossibilidade de realização do pronome lexical com SNs [-animados; -específicos] indefinidos ou quantificados. Essa hipótese se orienta pelos diferentes trabalhos desenvolvidos sobre o PB (Duarte, 1986; Cyrino, 1994; Casagrande,

[10] As entrevistas estão disponíveis em: https://peul.letras.ufrj.br/amostras Acesso em: 05 jan. 2026.

2012; Kato; Martins; Nunes, 2023) e também pelos resultados de nossas pesquisas (Simões, 2015, 2022), na qual comparamos as variedades de espanhol que investigamos ao PB.

Quanto aos contextos linguísticos analisamos: (1) a estrutura do SN antecedente; (2) o traço semântico de animacidade; (3) o traço semântico de especificidade; (4) a semântica verbal; (5) aspecto verbal I: construções [+/-perfectivas]; (6) aspecto verbal II: verbos dinâmicos e de estado; (7) as construções intensionais; (8) o objeto indireto; (9) a predicação secundária; (10) a perífrase verbal; (11) os contextos discursivos.

Em relação aos contextos sociais, investigamos três faixas etárias, sendo falantes de 20 a 34 anos (faixa etária 1), falantes de 35 a 54 anos (faixa etária 2) e falantes a partir dos 55 anos (faixa etária 3). Também investigamos o contexto social escolaridade, cujos fatores são falantes com o Ensino Fundamental, falantes com o Ensino Médio e falantes com Ensino Superior.

Para realizar a análise quantitativa, submetemos os dados ao programa estatístico *Goldvarb X*, que permite a observação da sistematicidade subjacente à variação linguística.

4. Análise dos dados

A partir da análise quantitativa dos dados, encontramos 84,9% de objetos nulos, 13,6% de pronomes lexicais e apenas 1,5% de pronomes clíticos. Observe-se a distribuição das variantes na tabela a seguir:

Tabela 2: Frequência de objetos nulos, pronomes lexicais e clíticos nas entrevistas analisadas do PEUL

Objetos nulos		**Pronomes lexicais**		**Pronomes clíticos**		**Total**	
n.	%	n.	%	n.	%	n.	%
1071	84,9%	172	13,6%	19	1,5%	1262	100%

Fonte: Elaboração da autora

Para a análise multivariada, que nos fornece os pesos relativos e nos permite verificar os contextos que favorecem uma das variantes, optamos por excluir os pronomes clíticos ao invés de incluí-los entre os pronomes lexicais, já que escolhemos essa variante como o valor de aplicação da regra variável com o objetivo de verificar os condicionadores linguísticos e sociais que a favorecem. Nossa hipótese geral é de que haveria restrições ou impossibilidade de realização do pronome lexical com SNs [-animados; -específicos] indefinidos ou quantificados.

Os condicionadores selecionados como significativos para a ocorrência do pronome lexical em ordem de relevância foram a animacidade do antecedente, as construções com contexto intensional, as construções com predicação secundária, a especificidade do antecedente e o nível de escolaridade do falante.

4.1. A animacidade do antecedente

Em relação ao condicionador relativo à animacidade do SN antecedente, a nossa hipótese era de que os SNs [+animados] favoreceriam a retomada do antecedente pelo pronome lexical. Nossa hipótese foi confirmada, já que o peso relativo foi de .87. Em relação ao objeto nulo, observamos 95.4% dessa variante com SNs [-animados]. As tendências encontradas confirmam os trabalhos de Duarte (1986) e Cyrino (2018 [1993], 2020), que revelam que o pronome lexical estaria associado a antecedentes [+animados] e os objetos nulos a antecedentes [-animados]. Observem-se os dados na tabela a seguir:

Tabela 3: Objetos nulos e pronomes lexicais com SNs antecedentes [+/-animados]

	Objetos nulos		**Pronomes lexicais**		
	n.	%	n.	%	Peso relativo
[+animados]	181/310	58.4%	129/310	41.6%	.87
[-animados]	890/933	95.4%	43/933	4.6%	.35

Fonte: Elaboração da autora

Nas construções em (28), apresentamos algumas construções com pronomes lexicais e objetos nulos retomando entidades [+animadas] e [-animadas] encontradas nos dados. Em (28a), temos um antecedente [+animado; +específico] retomado pelo pronome lexical. Quanto a (28b), se trata de um antecedente [+animado; -específico] retomado por objeto nulo. Em (28c), temos um antecedente [-animado; -específico] retomado pelo pronome lexical e, em (28d), um antecedente do mesmo tipo retomado por objeto nulo. Observa-se que em todos os casos os antecedentes aparecem próximos ao objeto anafórico que os retoma, sendo na ocorrência do próprio SN ou em sua retomada por pronome, como em (28a).

(28a) E- E essas pessoas que você morou? Por exemplo em Maricá: Você disse que foi por causa da **sua vizinha**, né? [F- Da vizinha] Como é que você convivia com essas pessoas?

F- Ah, convivia bem.

E- Como é que era, assim? Tava sempre na casa dela?

F- Tava. Ela me botou na escola (est) (hes) e... ela me tratava como filha, saia comigo, levava pros canto, pa igreja, que ela é muito evangélica também. A gente ia na igreja, eu ajudava muito **ela**, eu levava a filha dela pa escola. […] (Entrevista 8)

(28b) E- Mas você acha que no tempo do seus pais era assim também?

F- […] É, minhas tia… [minhas tia] conta, minha... uma tia minha que já morreu conta, que elas não conheciam **o noivo**, é os pais que arrumavam **Ø** ia lá: "olha, tenho uma filha aqui ocê vai casá, tal tal, e as meninas doida pra conhecê os rapazes. […] (Entrevista 8)

(28c) E- Quais são os procedimentos. Como você faz.

F- Eu lavo o arroz, né? soco **o alho** boto **ele** pá podê «morená» no óleo, jogo **ele**, refogo bem **ele**, e jogo água. (Entrevista 7)

(28d) F- […] é muitas palavras muito difíceis de de... de... (riso f) (“progredir” ou “prosseguir”); eh... acertá, entendeu? [(est)] [acertá] **as palavras** pra ela. Porque... (“t-”) (hes) por exemplo: <el-> (hes) ela vai e escreve **Ø** no quadro, então a gente tem que escrevê **Ø**, passá tudo pro caderno. Tá entendendo? (Entrevista 05)

4.2. O contexto intensional

O segundo contexto selecionado como significativo corresponde às construções com elementos intensionais. Conforme Leonetti (1999), nessas construções, os SNs receberiam uma interpretação [-específica] prototípica. Esses contextos compreendem elementos modais com âmbito oracional, como orações interrogativas, tempo futuro, futuro do pretérito, modo subjuntivo, perífrases verbais, nas quais se anula a existência do SN; construções com negação; e verbos intensionais, como *buscar, querer, necessitar*. Embora esse contexto não tenha sido selecionado como significativo em nossas pesquisas anteriores (Simões, 2015, 2022), optamos por mantê-lo em nossa pesquisa sobre o PB, partindo da hipótese de que o pronome lexical seria favorecido nas construções que não apresentam contexto intensional. Encontramos peso relativo de .58, que revela que a nossa hipótese foi confirmada. Na esfera dos objetos nulos, observamos frequências elevadas dessa variante tanto em construções com elementos intensionais quanto sem esses elementos, sendo que nas construções com esses elementos a frequência foi de 90.1%. Observe-se a tabela a seguir:

Tabela 4: Objetos nulos e pronomes lexicais em construções com e sem elementos intensionais

	Objetos nulos		**Pronomes lexicais**		
	n.	%	n.	%	Peso relativo
Construções com contextos intensionais	445/494	90.1%	49/494	9.9%	.38
Construções sem contextos intensionais	626/749	83.6%	123/749	16.4%	.58

Fonte: Elaboração da autora

Nas construções em (29), apresentamos ocorrências de contextos intensionais. Em (29a), temos um caso de perífrase e retomada do antecedente pelo objeto nulo. Em (29b), temos duas construções com negação e a retomada pelo pronome lexical.

(29a) E- E como é que faz a pintura de carro?

F- Ah... [a gent-] [a gent-] a gente tem que prepará **o carro**, tem que vê onde (“tá”) podre, tem que raspá... **Ø** lixá **Ø**, tem dá a massa, tem que dá o praime. (Entrevista 05)

(29b) E- Eh... o vagabundo mata **o marido dela**.

F- Isso. Aí depois ele volta pa podê conversá com ele. Então a alma dele, desse rapaz ("de") que foi morto, [ele-] [ele-] ele vê tudo! então ele fica... ele fica... maluco! Tá entendendo? Ele fica... (hes) fora de série fora de série, então ele fica nervoso, então [ele-] [ele-] ele começa a- a tentá batê nele, mas não conseguia, que passava por ("dentro dele"), era só alma! aí começava a tentá batê nele, chamava ele de "Safado! Traidor!..." não sei que lá, porque... porque ele via realmente tava afim (hes) da namorada dele. Então ele <n-> ele falou que não ia deixá isso em branco. Então começou corrê atrás de (hes), começou a viajá <den-> dentro dos vagões... é que <u-> (hes) [um-] uma alma penada. Começou a andá pelo mundo a fora... aí- <va-> aí ("ele") chegou num certo ponto que ele conheceu um rapaz, uma... [um-] uma alma também que tava dentro <do...> Não! Minto. Ele <com-> e- essa namorada dele foi no enterro dele... aí [ele-] ele tava atrás dela assim, mas ninguém via **ele**, que ele era uma alma... Aí... ele- ele mesmo viu o <tê-> [o-] [o-] o corpo dele ser enterrado. Aí ficou assim... ficou olhando... Aí (hes) chorou pá caramba, sentiu muita pena dela, que ela... [que-] que ele viu que ela tava sofrendo... Aí depois ele viu uma mulher, uma outra alma também que ("fez assim") com ele; [deu-] deu tchau pra ele... aí depois entrou na (inint) [e-] e desceu pu túmulo dela. Aí ele foi tentá falá com ela, mas ela tinha sumido; essa outra pessoa. Então foi dali que ele começou a resolvé os problema deles. Aí ele começou a viajá, começou a andá, andá, andá, começou a tomá conta da casa dela (hes), da vida dela, começou (hes), tentou na- ("ir dentro") da casa dela... sentava do lado dela mas ela não via **ele**. […] (Entrevista 5)

4.3. A predicação secundária

O terceiro contexto selecionado como significativo corresponde às construções com predicação secundária. De acordo com Masullo e Demonte (1999), os predicados verbais podem estar acompanhados de complementos predicativos, que compreendem elementos que modificam tanto o predicado verbal quanto um SN, que, por sua vez, pode ser o sujeito ou o objeto acusativo. Esse predicativo tem a função de atribuir um estado ou propriedade a um SN e mantém com o mesmo uma dependência sintática, que irá se manifestar mediante concordância de gênero e número, nos casos em que o predicativo é um adjetivo. Selecionamos para nossa pesquisa apenas as construções nas quais o predicativo modifica o objeto acusativo.

Tendo em vista os resultados de Duarte (1989), partimos da hipótese de que o pronome lexical seria favorecido nas construções de predicação secundária. Encontramos peso relativo de .80, de modo que confirmamos a nossa hipótese. Com respeito aos objetos nulos, encontramos frequência elevada dessa variante tanto nas construções com predicação secundária quanto nas demais construções, sendo a maior frequência neste último caso, com índice de 86.8%. Observe-se a tabela a seguir:

Tabela 5: Objetos nulos e pronomes lexicais em construções com e sem predicação secundária

	Objetos nulos		**Pronomes lexicais**		
	n.	%	n.	%	Peso relativo
Construções com predicação secundária	45/61	73.8%	16/61	26.2%	.80
Construções sem predicação secundária	1026/1182	86.8%	156/1182	13.2%	0.48

Fonte: Elaboração da autora

Nas construções em (30a) e (30b), apresentamos uma construção de predicação secundária em que o pronome lexical retoma o antecedente e uma em que ocorre objeto nulo, respectivamente. Em (30a), vemos que se trata de um SN antecedente [+animado; +específico]. Em (30b), o SN antecedente é [-animado; +específico]. Conforme Duarte (1986), a predicação secundária seria um contexto em que o pronome lexical é produtivo e isso se confirma em nossos resultados. Na construção em (31), temos um caso de SN antecedente [+animado; +específico], retomado pelo pelo lexical, em que não ocorre predicação secundária.

(30a) E- Eh... **o vagabundo** mata o marido dela.

F- Isso. Aí depois ele volta pa podê conversá com ele. Então a alma dele, desse rapaz ("de") que foi morto, [ele-] [ele-] ele vê tudo! então ele fica... ele fica... maluco! Tá entendendo? Ele fica... (hes) fora de série fora de série, então ele fica nervoso, então [ele-] [ele-] ele começa a- a tentá batê nele, mas não conseguia, que passava por ("dentro dele"), era só alma! aí começava a tentá batê nele, chamava **ele** de "Safado! Traidor!..." […] (Entrevista 5)

(30b) F- Isso... sentaram, <co-> a gente ficou conversando, batendo papo até chegá **o museu**.

E- E... o que <a-> ocorreu com a turma quando chegou no museu?

F- Ué!?... acharam **Ø** legal. (Entrevista 05)

(31) […] **Meu filho** caiu de dois metro. O de três ano [E- Meu Deus!] ficô todo mole, todo se bateno, não reagia. Aí de repente começô a dormí, num parava. Foi atendido sabe aonde? Minha prima correu, meu primo, botô **ele** no carro, e tava numa festa […] (Entrevista 08)

4.4. A especificidade do antecedente

Quanto ao traço de especificidade do antecedente, que compreende o quarto condicionador selecionado como significativo, nossa hipótese era de que o pronome lexical seria favorecido por antecedentes [+específicos]. Observamos que o peso relativo foi de .60, o que revela que nossa hipótese foi confirmada, na medida em que os SNs [+específicos] favoreceriam a retomada pelo pronome. Em relação aos objetos nulos, observamos frequências elevadas, sobretudo com os SNs [-específicos], cujo índice foi de 92.2%. Esses resultados se assemelham aos da pesquisa de Casagrande (2012), na qual a autora detecta que o traço de especificidade do SN seria o mais relevante para determinar a expressão pronominal ou não do antecedente. Observe-se a tabela 6:

Tabela 6: Objetos nulos e pronomes lexicais com SNs antecedentes [+/-específicos]

	Objetos nulos		**Pronomes lexicais**		
	n.	%	n.	%	Peso relativo
[+específicos]	275/380	72.4%	105/380	27.6%	.60
[-específicos]	796/863	92.2%	67/863	7.8%	.46

Fonte: Elaboração da autora

Nas construções em (32), apresentamos alguns casos em que temos pronomes lexicais e objetos nulos retomando SNs [+específicos] e [-específicos]. Em (32a), temos um SN [-animado; +específico] retomado pelo pronome lexical. Em (32b) e (32d), se trata de SNs indefinidos, [-animados; -específicos] e retomados pelo pronome lexical e objeto nulo, respectivamente. Quanto a (32c), temos um SN também indefinido, mas [-animado; +específico] e retomado pelo objeto nulo.

(32a) F- […] Eu gostava muito da que terminou das sete, como é que era o nome, eu não me lembro o nome, ah que terminou. Agora **essa "Uga uga"** eu acho engraçada, que tá dando, eu assisto **ela** alguns dias, porque depende do horário que eu chego em casa […] (Entrevista 22)

(32b) F- Olha, **um bife**... (hes) se tiver cortado, é só... se a pessoa quisé ele mais espalhadinho, dá uma batidazinha ou tivé aquela máquina de Imprenssá, dá uma Imprenssadazinha (hes) nela, como lá no meu trabalho tem. (inint) Ali você abriu **ele**, temperô, jogô na frigidera… […] (Entrevista 25)

(32c) E- Como é que era, assim? Tava sempre na casa dela?
F- […] Que ele tava meio confiado. Que o banheiro, lá, tinha **uma cortina**, que não eraporta, era uma cortina. A filha dele levantô **Ø**. Desde esse dia, que ele começou a me olhá diferente. […] (Entrevista 8)

(32d) F- […] Aí cresceu o bolo, você vai lá despeja **uma lata de leitE-... "Leite Moça" cum leite de coco**, bate no liqüidificador, fura ele todinho, joga **Ø** por cima depois bota coco ralado (inint) por cima. (se dirigindo a E) (inint) Ih! Ela vai querê que eu vô lá na casa dela fazê (risos E) as coisa pra ela. (Entrevista 8)

4.5. A escolaridade

O quinto e último condicionador selecionado como relevante refere-se ao contexto social de escolaridade. Considerando o estudo de Duarte (1986), partimos da hipótese de que os falantes do PB com menos escolaridade seriam os que favoreceriam o pronome lexical. Os dados revelam que o peso relativo foi de .60 para os falantes que têm Ensino Fundamental, de modo que confirmamos a nossa hipótese. Em relação às frequências, há 19.4% de pronome lexical entre os falantes do Ensino Fundamental, 12% entre os falantes do Ensino Médio e 6.8% entre os falantes do Ensino Superior. Quanto aos objetos nulos, é muito elevada a frequência dessa variante entre os falantes de todos os níveis de escolaridade, sobretudo entre os com Ensino Superior, cujo índice foi de 93.2%. Quanto aos falantes com Ensino Médio e Ensino Fundamental, os índices foram de 88% e 80.6%, respectivamente. Como já mencionamos, em Simões (2022), observamos objetos nulos quase categóricos nas entrevistas que analisamos da variedade de São Paulo do PB com informantes de alta escolaridade. Observem-se os dados na tabela abaixo:

Tabela 7: Objetos nulos e pronomes lexicais nas entrevistas dos falantes distribuídos por diferentes níveis de escolaridade

	Objetos nulos		Pronomes lexicais		
	n.	%	n.	%	Peso relativo
Ensino Fundamental	366/454	80.6%	88/454	19.4%	.60
Ensino Médio	513/583	88%	70/583	12%	.46
Ensino Superior	192/206	93.2%	14/206	6.8%	.41

Fonte: Elaboração da autora

Na construção em (33), apresentamos uma ocorrência de pronome lexical com antecedente [+animado; -específico], encontrada na entrevista de um falante com o Ensino Fundamental.

(33) F- É. A MESMA coisa acontece na <polít...>. Cê vê gente morrê aí, assassinada, aquela coisa toda, só porque sabe que vai ser um adversário. É a mesma coisa, entendeu? ter **um adversário**, desejo que ele morra ou mando alguém matar **ele** só porque eu acho «dono da cocada preta" […] (Entrevista 25)

4.6. Outros contextos linguísticos

Esclarecemos que foi necessário excluir da análise o condicionador referente à estrutura do SN, que compreende os diferentes determinantes que introduzem o SN antecedente ou a ausência de determinante. Essa exclusão se justifica devido fato de que, não apenas houve objetos nulos categóricos com SNs sem determinante, como houve apenas um caso de SN quantificado e nove de SNs introduzidos pelo artigo indefinido retomados por um pronome lexical, sendo as frequências de 1,6% e 6,2%, respectivamente. Quanto aos determinantes definidos, o índice de retomada de SNs antecedentes introduzidos por pronome lexical foi de 20,3%. Manter esse condicionador em nossa análise multivariada levava à não convergência dos resultados das rodadas.

Esse condicionador linguístico foi muito relevante nos dois estudos (Simões, 2015, 2022) que realizamos sobre a expressão do objeto acusativo nas variedades de Madri e Montevidéu. Como vimos, em Simões (2015), verificamos que os objetos nulos nessas duas variedades foram favorecidos por SNs sem determinante, quantificados e introduzidos pelo artigo indefinido. Em nossa pesquisa posterior (Simões, 2022), os SNs sem determinante e os quantificados favoreceram os objetos nulos na variedade de Madri cujos falantes tinham escolaridade equivalente ao Ensino Médio. Além disso, conforme vimos, já havíamos observado em Simões (2015, 2022) que, no PB, a possibilidade de retomada pelo pronome lexical estava associada principalmente aos SNs introduzidos por determinantes definidos, havendo restrições para os introduzidos pelo artigo indefinido e, sobretudo, pelos quantificados. Devido a esses resultados das pesquisas anteriores, nossa hipótese para esse contexto era de que haveria restrições ou impossibilidade de realização do pronome lexical com SNs indefinidos ou quantificados. A tendência que observamos na análise das entrevistas do PEUL é compatível com o traço semântico de definitude dos determinantes definidos, pois são esses os que

apresentam retomada pelo pronome lexical, assim como o objeto nulo seria a variante que retoma um SN sem determinante ou introduzido por determinante com o traço semântico de indefinitude (Leonetti, 1999). Essa tendência também se alinha com a hipótese que tínhamos para esse contexto linguístico, embora não possamos afirmar que a mesma foi confirmada, devido à exclusão desse contexto da análise multivariada. Observe-se a tabela abaixo com os resultados:

Tabela 8: Objetos nulos e pronomes lexicais com SNs antecedentes introduzidos por diferentes determinantes ou sem determinante

	Objetos nulos		**Pronomes lexicais**	
	n.	%	n.	%
Determinante definido	617/798	77.3%	162/798	20.3%
Artigo indefinido	136/145	93.8%	9/145	6.2%
Quantificador	62/63	98.4%	1/63	1.6%
SN sem determinante	256/256	100%	0/256	0%

Fonte: Elaboração da autora

Nas construções em (34), apresentamos casos que encontramos de SNs antecedentes introduzidos pelos diferentes tipos de determinante e sem determinante. Em (34a), temos um SN definido, [-animado; -específico], retomado pelo pronome lexical. Nas construções de (34b) a (34d), todos os antecedentes foram retomados por objeto nulo, sendo um SN indefinido [-animado; -específico], um SN quantificado [-animado; -específico] e um SN sem determinante [-animado; -específico], respectivamente.

(34a) F- […] Aí cresceu **o bolo**, você vai lá despeja uma lata de leitE-... "Leite Moça" cum leite de coco, bate no liqüidificador, fura **ele** todinho, joga por cima depois bota coco ralado (inint) por cima. (se dirigindo a E) (inint) Ih! Ela vai querê que eu vô lá na casa dela fazê (risos E) as coisa pra ela. (Entrevista 8)

(34b) E- É comida normal (falando rindo) pode sê. Como você faz, por exemplo, **um arroz**?
F- Como é que eu faço **Ø**? [E- Isso] Ah, eu faço arroz normal! (Entrevista 7)

(34c) E- Você já ouviu **algum debate**?
F- Nunca parei pa podê escutá **Ø** não. (Entrevista 7)

(34d) E- [Han-han.] E o que você acha por exemplo da... desse movimento Gospel...? de **música religiosa** que começou com a... igreja evangélica e agora tá com a igreja católica... com o Padre Marcelo... O que você acha de tudo isso?
F- Cada um tem... tem o seu destino né? realmente, cada um tem o seu destino [o-] [o-] o destino já é cruzado, [(est)] certo? então a pessoa... se a pessoa acha que aquilo ali é bom pra ela, então a pessoa segue <aquel-> <a-> [aquele-] aquele seu destino, tá entendendo? porque [ela-] ela vai adianta, então pra mim... pra mim não tem nada vê isso. Eu escuto **Ø** ("também") de vez em quando eu escuto **Ø** mas... […] (Entrevista 5)

Ao cruzarmos o contexto de animacidade com o de especificidade do antecedente, observamos que entre os antecedentes [-animados; -específicos] e [-animados; +específicos] houve uma frequência elevada de objetos nulos, sendo esta de 94% e 98%, respectivamente. Por outro lado, a frequência do pronome lexical é bastante baixa nesses contextos. Quanto aos antecedentes [+animados; -específicos], a frequência de objetos nulos foi de 75% e a de pronomes lexicais, 22%. Contudo, na esfera dos antecedentes [+animados; +específicos], que seriam os mais referenciais, conforme a Hierarquia Referencial proposta por Cardinaletti e Starke (1994), a frequência de pronome lexical foi de 50% e a de objetos nulos, 43%. Esses resultados revelam que o contexto de SN [+animado] parece ser crucial para uma maior frequência do pronome lexical, tendência que estaria em consonância com o principal contexto de favorecimento dessa variante. Para finalizar a reflexão, detectamos nessa análise 86 ocorrências de objetos nulos com antecedentes [+animados; +específicos], o que revela a possibilidade de omissão do objeto com esses SNs. Nosso resultado diverge da análise de Cyrino (2020), na qual o PB não permitiria essa categoria vazia com esse tipo de antecedente. Observe-se a tabela a seguir com os resultados:

Tabela 9: Objetos nulos e pronomes lexicais com SNs antecedentes [+/-animados; +/-específicos]

	Objetos nulos		**Pronomes lexicais**	
	n.	%	n.	%
[-animados; -específicos]	701/742	94%	39/742	5%
[+animados; -específicos]	95/126	75%	28/126	22%
[-animados; +específicos]	189/193	98%	4/193	2%
[+animados; +específicos]	86/201	**43%**	101/201	**50%**

Fonte: Elaboração da autora

Na construção em (35a), o pronome lexical retoma um antecedente [+animado; -específico], enquanto, em (35b), o pronome lexical retoma um antecedente [+animado; +específico]. Quanto às construções de (36a) a (36g), todas apresentam objetos nulos com antecedentes [+animados; +específicos]. Observamos que a maior parte dessas construções apresenta aspecto imperfectivo, com verbos no presente e pretérito imperfecto do modo indicativo. Além disso, em (36b), o objeto nulo aparece em uma construção final seguida de verbo no infinitivo e, em (36c) e (36f), temos contextos intensionais com a negação. É possível que esses contextos possibilitem os objetos nulos com os antecedentes [+animados; +específicos]. Como vimos, Lopes (2007) encontra evidências de que os objetos nulos estariam relacionados ao aspecto imperfectivo. Duarte (1986) encontra alta frequência dessa variante com verbos no infinitivo. Landa (1993), em seu estudo sobre os objetos nulos na variedade do espanhol basco, observa que, embora a omissão do objeto seja mais frequente com antecedentes [-animados], também é possível com antecedentes [+animados] e, nesses casos, as construções apresentam contextos como aspecto imperfectivo e verbos no infinitivo, assim como em alguns dos exemplos que apresentamos nas construções em (36). Além disso, esses dados de objetos

nulos ocorrem quando o antecedente está no contexto discursivo e há um caso de tópico também, ambos contextos em que Kato, Martins e Nunes (2023) afirmam que os objetos nulos [+animados] poderiam ocorrer. Em outro trabalho, elegiremos a variante objeto nulo para a aplicação da regra variável, de modo que observaremos os contextos que a favorecem e teremos pistas mais precisas sobre a possibilidade de objetos nulos com antecedentes mais referenciais.

(35a) E- Como você educaria... **o seu filho**, se você tivesse um?

F- Uhm... (inint) educaria **ele** tê... tê um pouco de... responsabilidade na vida, certo? […] (Entrevista 5)

(35b) F- […] **Meu filho** caiu de dois metro. O de três ano [E- Meu Deus!] ficô todo mole, todo se bateno, não reagia. Aí de repente começô a dormí, num parava. Foi atendido sabe aonde? Minha prima correu, meu primo, botô **ele** no carro […] (Entrevista 8)

(36a) E- Ahn!... Então me diz! como é que foi esse dia que... (falando rindo) você foi defendê **seu irmão**! (risos) Você tinha quantos anos?!... Ele tinha quantos anos?!...

F- Eu tinha na faixa de três pa quatro ano e ele tinha dois anos. (pausa) Aí ("pa-") brincando, brincando assim, aí os dois os dois começaram (hes) a brigá lá, aí eu começava ("a entrar") no meio ("um") não deixava, porque o outro tava ganhando, tava batendo nele, então eu não deixava. [Eu-] [eu-] eu (hes) defendia **Ø** ao mesmo tempo e não deixava os dois brigá. […] (Entrevista 5)

(36b) E- […] então... ele se interessou por **essa garota**, ("tá entendendo")? Então pra... conseguir **ela**, então ele achou um lugar pra podê... deixá ela sozinha, então [ele-] ele contratou um vagabundo […] (Entrevista 5)

(36c) F- […] [Eu-] [eu-] eu (hes) defendia ao mesmo tempo e não deixava **os dois** brigá. Tá entendendo? Eu separava **Ø**. (pausa) Porque eu não gostava daquilo, não gostava de briga, não gostava [de...] [de...] [de...] de <ban-> é vandalismo, entendeu? Não gostava de ninguém brigá; de vê ninguém brigá. (Entrevista 5)

(36d) E- Todos na mesma escola, né?

F- Não. **Esse mais velho** era da Casa da Criança da escola, mas não era da creche, [E- Ah!] e eu botei **Ø** pro Joaquim Nabuco. […] (Entrevista 8)

(36e) F- […] E **os meus dois pequeno**, eu já vou e pego **Ø**. São menores. E... a violência não acaba não que é... daqui a pior. Vai piorá mais ainda! (Entrevista 8)

(36f) E- E eles se conheceram aonde, seus pais?

F- Eles se conheceram:... em Centenário, Caxias....Centenário. O velho tinha um ciúme dela, minha filha! Até, tadinho, (inint) ele morreu ...ele antes de ele morrê, ele <olh...> abriu o olho, olhô BEM pra ela, olhô bem mesmo, como se tivesse, assim: "Eu vô olhá bem pra ela, porque eu num vô vê **Ø** mais, né?" e morreu. […] (Entrevista 8)

(36g) F- De manhã, até de noite. Eu acordo de manhã, faço o leitinho dos **meus filhos**, chocolate dô pra eles, boto **Ø** pra escola, aí depois, levo **Ø** pra escola, aí volto, espero o velho acordá, que é meu marido, levanto, boto o café dele, aí vai embora trabalhá […] (Entrevista 8)

Faz-se necessário esclarecermos também que, além do condicionador linguístico referente à estrutura do SN, excluímos da análise os condicionadores referentes à semântica verbal e às construções com objeto indireto, devido ao fato de que nesses contextos houve ocorrência categórica de objetos nulos.

Com relação ao condicionador da semântica verbal, houve objetos nulos categóricos nas construções com verbos de comunicação. Kany (1969) demonstra que nesse contexto linguístico ocorreriam objetos nulos em variedades da língua espanhola e os dados desse autor revelam que isso não se restringiria aos objetos oracionais. Esse seria, portanto, um condicionador que apresenta a possibilidade dessa categoria vazia em outras línguas naturais. Observem-se os resultados na tabela a seguir:

Tabela 10: Objetos nulos e pronomes lexicais em construções com verbos cognitivos, verbos de comunicação e demais verbos

	Objetos nulos		**Pronomes lexicais**	
	n.	%	n.	%
Verbos cognitivos	220/252	87.3%	29/252	11.5%
Verbos de comunicação	13/13	100%	0/13	0%
Demais verbos	838/997	84.1%	143/997	14.3%

Fonte: Elaboração da autora

Na construção em (37a), temos um caso de pronome lexical em construção com o verbo cognitico *esperar*. Quanto à (37b), temos um caso de objeto nulo com verbo cognitivo e, em (37c), um caso de objeto nulo com verbo de comunicação.

(37a) F- Que eu gostei muito muito muito dessa última remessa foi Titanic. (est) Entendeu? Titanic. e... e tem um filme também dessa <Ma...>: da **Madonna**, que é Evita. Eu só num gostei porque o filme foi baseado cantando (est), né? É como se eu tivesse falando com você cantando. Entendeu, aí eu num gostei por causa disso. Que eles nunca falavam conforme a gente tá falando, entendeu? Mas tudo era cantado, né? Qué dizê, eu só gostei mesmo por causa da música do final, né? que ela ficou doente, né? e antes de ela morrê ela apareceu na janela, entendeu, aí aquele povo, né? lá esperando **ela** aparecê […] (Entrevista 19)

(37b) F- […] Eu amo mesmo **essa profissão**. (est) Amo **Ø**, eu gosto da minha profissão […] (Entrevista 19)

(37c) F- […] Eles pedem **informação** a mim, eu dou [pela...pela] pelos anos de trabalho (est), eu também peço alguma informação a eles também que eles também que eles tão fazendo faculdade. (est) Então um pedindo **Ø** ao outro. […] (Entrevista 19)

Quanto às construções com objeto indireto, houve objetos nulos categóricos nesse contexto. Como vimos, o estudo de Duarte (1986) revelou que esse é um contexto de alta frequência de omissão do objeto. Além disso, conforme os trabalhos de Suñer e Yépez (1988) e Landa (1993), se trata de um contexto em que ocorrem objetos nulos nas variedades de espanhol de Quito e do País Basco,

respectivamente. Sendo assim, observamos a possibilidade de objetos nulos nesses contextos nas línguas naturais. Observem-se as tabelas a seguir:

Tabela 11: Objetos nulos e pronomes lexicais em construções com ou sem objetos indiretos

	Objetos nulos		**Pronomes lexicais**	
	n.	%	n.	%
Presença de objeto indireto	29/29	100%	0/29	0%
Ausência de objeto indireto	1042/1233	84.5%	172/1233	13.9%

Fonte: Elaboração da autora

Em (38), temos um caso de objeto nulo em uma construção com objeto indireto.

(38) F- De manhã, até de noite. Eu acordo de manhã, faço **o leitinho dos meus filhos**, chocolate dô **Ø** pra eles, boto pra escola, aí depois, levo **Ø** pra escola, aí volto […] (Entrevista 8)

As tendências de nossa análise da variedade carioca do PB revelam que os pronomes lexicais são favorecidos pelos SNs [+animados] e/ou [+específicos], pelas construções que apresentam predicação secundária, pelas construções que não apresentam contextos intensionais e pelos falantes com menor escolaridade. Quanto aos objetos nulos, vemos que essa variante é altamente frequente, sobretudo com os SNs [-animados], além de ter havido ocorrência categórica de objetos nulos em construções com verbos de comunicação e em construções com objetos indiretos. Vemos que essas tendências são as que aparecem nos trabalhos de Duarte (1986), Cyrino (1994) e Casagrande (2012) sobre o PB e nos trabalhos de Suñer e Yépez (1988), Landa (1993) e em Simões (2015, 2022) sobre as variedades de espanhol. Esses resultados nos permitem observar os contextos linguísticos e o contexto social em que os pronomes lexicais estão encaixados, bem como como uma mais alta implementação dessa variante com antecedentes [+animados] e em construções de predicação secundária, que foram as que apresentaram pesos relativos mais elevados.

Além disso, vemos que os pronomes lexicais retomam de maneira majoritária os SNs introduzidos pelos determinantes definidos e são poucos os casos em que temos o artigo indefinido e, sobretudo, um quantificador. Já havíamos observado essas tendências em Simões (2015, 2022), já que haveria uma incompatibilidade entre o pronome lexical do PB e os determinantes com o traço semântico de indefinitude ou mesmo a ausência de determinante. Vemos, portanto, um alto grau de implementação do pronome lexical nesse contexto.

Observamos também a Hierarquia Referencial de Cardinaletti e Starke (1994) refletindo-se em nossos resultados, uma vez que são os antecedentes menos referenciais os que mais tendem a expressar-se mediante um objeto nulo, enquanto os mais referenciais tendem a expressar-se mediante um pronome.

A partir desses resultados, torna-se bastante clara a assimetria existente entre o PB e as variedades de espanhol que estudamos, na medida em que os antecedentes que permitem objetos nulos nessas variedades do espanhol são aqueles que não permitem o pronome lexical no PB.

5. Análise teórica

Em Simões (2015), com base nos desenvolvimentos minimalistas (Chomsky, 1999, 2000, 2004) e em outros trabalhos gerativistas, propomos uma análise para a derivação das construções de objeto acusativo anafórico nas variedades de espanhol de Madri e Montevidéu, assim como para o PB[11]. Nossa compreensão é de que podemos propor essa mesma análise para a variedade carioca do PB, na medida em que as tendências encontradas se assemelham às de nossos estudos anteriores: objetos nulos retomando entidades menos referenciais, pronomes retomando entidades mais referenciais e diferentes contextos linguísticos atuando nessas possibilidades.

Em relação às construções em que o pronome retoma o objeto, como em sentenças como *María lo vio* e *A Maria viu ele*, os pronomes *lo* e *ele* teriam, por um lado, os traços semânticos [+humano], [+masculino], de acordo com o seu antecedente, que correspondem aos traços semânticos intrínsecos. Por outro lado, teriam o traço [+específico], que seria opcional. No que se refere aos traços sintáticos ou formais, esses dois pronomes apresentariam o traço categorial D, de determinante, e também os traços-φ, que seriam de pessoa, número e gênero, e correspondem a [3ª pessoa], [-plural], [+masculino], respectivamente. O traço categorial e os traços-φ seriam [+interpretáveis] e o traço de Caso seria [-interpretável]. Uma diferença entre esses pronomes é a de que o clítico teria o Caso marcado no léxico, sendo um Caso morfológico (*cf.* Groppi, 1997), enquanto o pronome lexical, que teria perdido o Caso morfológico (*cf.* Kato, 2002), teria o traço de Caso atribuído no arranjo lexical. Por fim, o *v* teria Caso acusativo para atribuir ao argumento interno e traços-φ [-interpretáveis].

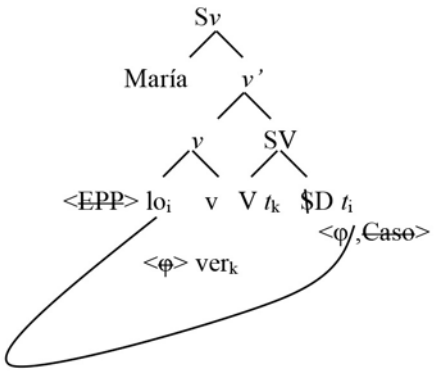
Apoiando-nos nas análises de Groppi (1997) e Galves (2001), sugerimos que seria desencadeada a operação de movimento, pois haveria a seleção do traço-EPP[12] pelo predicado verbal. Portanto, nas variedades do espanhol, o pronome clítico verificaria e apagaria esse traço em adjunção a *v*, enquanto no PB o pronome lexical faria o mesmo procedimento no especificador de S*v*. Em (39a) e (39b), apresentamos as derivações das construções com pronome clítico no espanhol e pronome lexical no PB (Simões, 2015, p. 367) conforme a proposta de Groppi (1997) e Galves (2001), respectivamente.

---

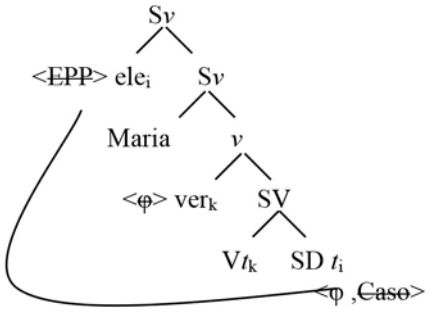
[11]Sugerimos a leitura do capítulo 5 de nossa tese (Simões, 2015) para mais detalhes de nossa proposta teórica.

[12] Conforme Chomsky (2000, p. 102), no âmbito da categoria sintática de sujeito, o traço-EPP de T seria universal. Por outro lado, no âmbito de *v*, que corresponde ao verbo *light* núcleo das orações transitivas, essa categoria poderia ter um traço-EPP análogo ao traço de T, mas que seria opcional. A presença do traço-EPP geraria um especificador extra e, no caso de v, esse especificador teria a função de receber o objeto direto movido da posição na qual foi gerado. Sendo assim, em línguas em que não são possíveis sujeitos nulos, como é o caso do inglês, seria inserido um expletivo, que checaria esse traço. Entretanto, o traço-EPP de *v* não poderia ser checado por um expletivo. Chomsky (2001, p. 26) propõe que o movimento do objeto ao especificador mais externo de Sv, que encontra-se acima do especificador em que o sujeito da sentença é gerado, se denomina *object shift*

(39a)



(39b)



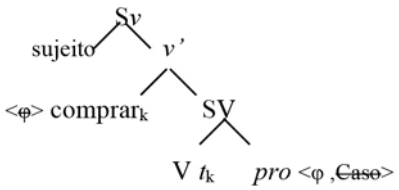
Quanto às construções de objetos nulos, encontramos evidências de que essa categoria vazia constitui um *pro* nas variedades de espanhol de Madri e Montevidéu (Simões, 2015) e, considerando o estudo de Galves (2001), assumimos que se trata dessa mesma categoria vazia no PB.

Nos casos em que temos um SN sem determinante, a proposta de Chomsky (1999) é de que a elevação visível do objeto favorece os SNs definidos e [+específicos]. Por outro lado, quando se trata dos SNs [-específicos], a ideia é de que estes permanecem na posição na qual foram gerados. Nos desenvolvimentos recentes do minimalismo, Chomsky (2000) defende que a elevação visível de um SN corresponderia à operação de movimento, enquanto, se um SN permanece *in situ*, seria desencadeada apenas a operação de concordância.

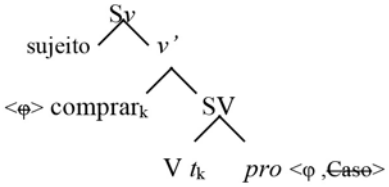
Com base na ideia de Chomsky (1999) de que os SNs definidos e específicos favoreceriam a elevação visível do objeto e nos indícios das línguas escandinavas (Thráinsson, 2008), nossa proposta é de que o *pro* permaneceria na posição em que foi gerado, de modo que ocorreria apenas a operação de concordância.

No arranjo lexical de sentenças que apresentam um SN sem determinante, como *café/café*, ocorreria a seleção de um *pro*. Esse nome teria os traços semânticos [-animado], [+masculino], [-específico] e como traços formais o traço categorial D, os traços-φ de [3ª pessoa], [-plural], [+masculino] e o traço de Caso acusativo. Tendo em vista que não seria desencadeada a operação de movimento, o *v* não selecionaria o traço-EPP. Portanto, apenas a operação de concordância seria desencadeada e seriam verificados os traços-φ de *v* e o Caso do objeto. Observem-se em (40a) e (40b) as derivações dessas construções (Simões, 2015, p. 369):

(40a)



(40b)



Nossa proposta é de que essa seria a derivação de uma sentença de objeto nulo com SN antecedente sem determinante tanto nas variedades de espanhol de Madri e Montevidéu quanto no PB, incluindo a variedade carioca de nosso estudo.

Feita essa análise para os objetos nulos com antecedentes sem determinante, o problema seria como analisar os casos em que os objetos nulos tem como antecedentes SNs introduzidos por determinantes. Para solucioná-lo, elaboramos duas hipóteses de análise. Na primeira, o traço-EPP seria selecionado e a operação de movimento seria desencadeada, enquanto, na segunda, o traço-EPP não seria selecionado e apenas ocorreria a operação de concordância.

Para a primeira hipótese, a ideia seria de que o *v* selecionaria o traço-EPP, de modo que a operação de movimento seria desencadeada. Segundo a análise de Galves (2001), da mesma forma que o pronome lexical se adjunta a *v*P, o objeto nulo de natureza pronominal também seria legitimado nessa posição. Considerando a tipologia de Cardinaletti e Starke (1994), Galves (2001) defende que a categoria pronominal sem manifestação fonética constituiria um pronome deficiente. Conforme a tipologia proposta por esses autores, o *pro* se classificaria como um pronome deficiente, que se caracteriza por não poder permanecer na posição temática e, portanto, deveria mover-se.

Essa primeira hipótese também poderia ser sustentada pela proposta de Chomsky (1999) de que tenderia a haver a elevação visível dos objetos definidos e [+específicos], correspondente à operação de movimento. Contudo, se adotássemos essa primeira hipótese, como explicaríamos os objetos nulos com antecedente introduzido por determinante e interpretado como [-específico]?

Com respeito à segunda hipótese, nossa ideia seria de que *v* não selecionaria o traço-EPP e, portanto, *pro* não se moveria para o especificador de S*v*, mas permaneceria na posição na qual foi gerado.

Como primeiro argumento para sustentar a segunda hipótese, temos a proposta de Chomsky (1999, p. 380):

> (...) os traços-φ de um DP especificador aparecem usualmente tanto no DP como no núcleo verbal, mas o traço Casual do DP não aparece no núcleo verbal. Existe pelo menos uma tendência para os traços-φ se manifestarem visivelmente quando a elevação para o domínio de verificação é visível em vez de não-visível, como se vê na concordância verbal com o sujeito versus o objeto em línguas do tipo nominativo-acusativo com o princípio EPP (...), ou na concordância visível com o particípio em francês, como reflexo da elevação visível.

Considerando-se que a elevação não-visível corresponderia à operação de concordância e a elevação visível à operação de movimento (*cf*. Chomsky, 2000), a proposta de Chomsky sugere que, se não há concordância visível entre verbo e argumento, ocorreria apenas a operação de concordância, na qual os traços [-interpretáveis] são verificados sem que ocorra a operação de movimento. Por outro lado, se a concordância entre verbo e argumento se manifesta de forma visível, seria desencadeada a operação de movimento e esse fato requer a existência do traço-EPP. Os morfemas flexionais e os clíticos seriam elementos pronominais e formas de manifestação da concordância. No âmbito

dessa proposta, a presença do pronome seria um elemento de concordância manifestando-se de maneira visível entre o verbo e o objeto[13]. Por outro lado, nas construções em que temos o *pro* como argumento do verbo, a concordância não seria visível, na medida em que os traços-φ do argumento não se manifestariam foneticamente. Sendo assim, haveria apenas a operação de concordância.

O segundo argumento para sustentar a nossa segunda hipótese constitui a proposta de Chomsky (2004) de que o léxico determinaria a geração de um segundo especificador. Portanto, quando um núcleo seleciona o traço-EPP, um segundo especificador seria gerado e a operação de movimento seria desencadeada. Tendo em vista essa proposta de Chomsky (2004), a nossa pergunta é: o que as tendências de nossos estudos poderiam revelar?

Se o predicado verbal é o elemento responsável pela seleção ou não seleção do traço-EPP, em nossa pesquisa mais recente (Simões, 2022), na variedade de espanhol de Madri, os objetos nulos foram favorecidos pelas construções com perífrase verbal e verbos de estado. Além disso, a omissão do objeto foi mais frequente em construções com aspecto [-perfectivo], que expressa eventos não delimitados, e em construções intensionais, que são interpretadas como contextos não factuais. Em Simões (2015), os objetos nulos foram mais frequentes nas construções [-perfectivas], com clítico dativo e com perífrase verbal, além de terem apresentado elevada frequência nas construções com clíticos dativos. Na variedade do PB de São Paulo, na qual os objetos nulos estão altamente implementados, verificamos que a omissão foi categórica nas construções com perífrase verbal, que tendem a ser contextos de aspecto [-perfectivo]. Sendo assim, haveria evidências de que existiria uma relação entre os predicados que expressam eventos não acabados e/ou não factuais e a não seleção do traço-EPP. Em nossos resultados da variedade carioca, vemos os objetos nulos ocorreram de maneira

---

[13] Na língua espanhola, no PB e em outras línguas românicas, temos a manifestação da concordância entre verbo e objeto mediante um pronome. Por outro lado, existe a concordância entre verbo e objeto mediante flexão verbal. De acordo com Huang (1984), o pashto, língua falada no Afeganistão, exibe sujeitos e objetos nulos quando ocorre concordância verbal. Portanto, nessa língua, quando se expressam eventos no presente, ocorre a concordância do verbo com o sujeito em orações intransitivas e transitivas [exemplos (1a) e (1b)]. Entretanto, quando os verbos expressam eventos no passado, ocorre a concordância do verbo com o sujeito nas construções intransitivas e com o objeto nas transitivas [exemplos (2a) e (2b)]. Assim como as línguas românicas, o Pashto seria também uma língua indo-europeia, pertencente à subfamília indo-iraniana (*cf.* Faraco, 2006).

(1a) Jān ra-z-i.
John DIR-come-3msg
'John comes'.

(1b) zə mana xwr-əm.
I apple eat-1msg.
'I eat the apple'.

(2a) Jān ra-ğ-ay.
John ASP-come-3msg
'John came'.

(2b) ma mana wə-xwar-a.
I apple PRF-eat-3fsg
'I ate the apple'. (Huang, 1984, p. 535)

categórica nas construções com verbos de comunicação e com objetos indiretos. Há uma gradação na possibilidade de expressão pelo objeto nulo ou pelo pronome lexical que está relacionada com a referencialidade do antecedente e com o predicado verbal.

O terceiro argumento que sustenta a nossa segunda hipótese compreende os resultados das pesquisas de Fujino e Sano (2002) e Ezeizabarrena (1997) a respeito da aquisição do espanhol como língua materna. Fujino e Sano (2002) observaram que, durante a aquisição do espanhol, haveria um período em que as crianças omitem o objeto e essa omissão diminui à medida que o clítico se torna mais frequente. Quanto à pesquisa de Ezeizabarrena (1997), esta verificou que as crianças começam a apresentar os pronomes clíticos de maneira produtiva depois dos dois anos de idade. Conforme essa autora, tanto os morfemas flexivos quanto os clíticos correspondem a morfemas de concordância. Portanto, a sua proposta é de que os clíticos constituiriam a realização morfológica de Conc-O e a sua presença na produção das crianças sugeriria que a categoria funcional Conc já estaria disponível em sua gramática. Nossa ideia é de que os resultados dessas pesquisas poderiam revelar que existe uma relação entre a produção dos clíticos e a categoria funcional Conc-O, que, nos desenvolvimentos minimalistas mais recentes (Chomsky, 2000), corresponde a S*v*[14].

Nosso último argumento para sustentar a segunda hipótese é de que a agramaticalidade das construções em (41) e (42) se explicaria pela ausência de um pronome para retomar um antecedente [+animado; +específico], o mais referencial entre todos. Propomos que isso ocorreria devido ao fato de o *v* selecionar o traço-EPP. Contudo, se, ao invés do pronome, é selecionado o *pro*, que não apresenta traços fonéticos e não é capaz de verificar e apagar o traço-EPP, as derivações não seriam convergentes. Além disso, observe-se que nessas construções temos aspecto perfectivo. Vimos em nossa análise da variedade carioca do PB várias construções de objetos nulos [+animados; +específicos] em construções de aspecto imperfectivo e verbo no infinitivo. Nossa ideia é de que esses objetos nulos ocorreriam com os antecedentes mais referenciais devido à natureza de suas construções, que não levariam à seleção do traço-EPP pelo predicado verbal.

(41) **María** llegó hace poco. Juan dijo que **la** / ***Ø** trajo.

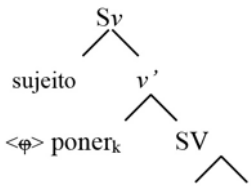
(42) **A Maria** chegou faz pouco tempo. O João disse que trouxe **ela** / ***Ø.** (Adaptado de Simões, 2015, p. 378)

Tendo em vista essa segunda hipótese, consideremos a derivação de uma sentença como *a gente tem que prepará* ***o carro****, tem que vê onde ("tá") podre, tem que raspá...* ***Ø*** *lixá* ***Ø***, extraída da entrevista 05 da amostra. De acordo com o SN antecedente, *pro* apresentaria os traços semânticos [-animado], [+masculino] e [-específico]. Quanto aos traços sintáticos, estes seriam o traço categorial D, os traços-φ de [3ª pessoa], [-plural] e [+masculino] e o traço [-interpretável] de Caso. Na medida em que não haveria a seleção do traço-EPP, apenas a operação de concordância seria desencadeada

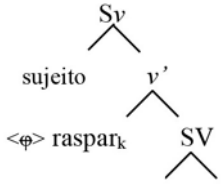
[14] Como vimos anteriormente, adotamos a proposta de análise de Groppi (1997) e Galves (2001) para a derivação de sentenças com clíticos no espanhol e com pronome lexical no PB, respectivamente, que implica a seleção do traço-EPP pelo predicado verbal. Sendo assim, consideramos que esse argumento sobre a aquisição do espanhol como língua materna também se aplicaria aos dados do PB carioca.

e seriam verificados e apagados os traços-φ do verbo e o traço de Caso de *pro*. Os traços-φ de *pro* poderiam corresponder ao morfema zero que Landa (1993, 1995) propõe e que permitiria que os objetos nulos na variedade de espanhol do País Basco fossem licenciados. Os traços-φ de *pro*, que não teriam traços fonéticos, permitiriam identificar o antecedente do objeto argumento do verbo. Esse mecanismo explicaria a possibilidade do objeto nulo. Em (43a) e (43b), apresentamos a derivação para objetos nulos cujo SN antecedente apresenta um determinante. Nas variedades de espanhol estudadas, nesses casos, o SN tende a ser indefinido ou quantificado como vimos.

(43a)



(43b)



Considerando os argumentos qua apresentamos para sustentar nossa segunda hipótese, concluímos que essa seria a hipótese mais plausível e adequada. Portanto, a análise que sugerimos é a de que, nas construções que apresentam objetos nulos com antecedente que constitui um SN introduzido por determinante, o predicado verbal não selecionaria o traço-EPP, de maneira que não ocorreria a operação de movimento, mas apenas a de concordância. Seriam os traços-φ de *pro* que identificariam o antecedente do objeto.

As tendências encontradas em nossas pesquisas anteriores (Simões, 2015, 2022), as tendências encontradas na variedade carioca do PB e nossa análise teórica indicam que haveria uma coexistência de gramáticas (Chomsky, 1999; Lightfoot, 1999) nas variedades de espanhol de Madri e Montevidéu e nas variedades do PB, que se manifesta em forma de variação linguística. Portanto, conforme Chomsky (1999), uma das gramáticas corresponderia à retomada do objeto por um pronome e a outra corresponderia à sua omissão.

6. Considerações finais

Os resultados de nossa pesquisa sobre a variedade carioca do PB revelaram que os pronomes lexicais foram favorecidos pelos SNs [+animados], as construções que não apresentam elementos intensionais, as construções com predicação secundária, os SNs [+específicos] e os falantes com o Ensino Fundamental. Observamos também que são os SNs introduzidos por determinantes definidos que permitem a retomada por essa variante. Por outro lado, os objetos nulos apresentaram elevada frequência, além de terem sido categóricos em construções com verbos de comunicação e objetos indiretos. As tendências que encontramos assemelham-se aos resultados de Duarte (1986), Cyrino (1994) e Casagrande (2012) no âmbito do PB, de Suñer e Yépez (1988) e Landa (1993) no âmbito de variedades do espanhol e de nossas pesquisas anteriores sobre a variedade de espanhol de Madri e Montevidéu e o PB (Simões, 2015, 2022).

Tendo em vista esses resultados, propomos nossa análise gerativista (Simões, 2015), com base nos desenvolvimentos minimalistas (Chomsky, 1999, 2000, 2004), para a derivação de construções com pronomes e objetos nulos. Sendo assim, nas construções com pronomes, seja o clítico para o espanhol ou o pronome lexical para o PB, estes seriam a manifestação visível da concordância de objeto, enquanto nas construções de objeto nulo teríamos uma concordância não visível, possibilitada pela identificação que os traços-φ de *pro* fazem do antecedente do objeto.

Referências

CARDINALETTI, Anne; STARKE, Michael. The typology of structural deficiency: On three grammatical classes. *Working Paper in Linguistics*, University of Venice, v. 4, n. 2, pp. 41-109, 1994.

CASAGRANDE, Sabrina. *A correlação entre aspecto e objeto no PB*: uma análise sintático-aquisicionista. 2010. 301 f. Tese de doutorado, IEL/UNICAMP, Campinas.

CASAGRANDE, Sabrina. Restrições de ocorrência do objeto direto anafórico no português brasileiro: gramática adulta e aquisição da linguagem. *ReVEL*, n. 6, pp. 131-163, 2012. Disponível em: http://www.revel.inf.br/files/caeb5315459fe2b37f2046bdd7d3c53c.pdf Acesso em: 16 jan. 2026.

CHOMSKY, Noam. *Lectures on Governing and Binding*. Dordrecht: Foris, 1981.

CHOMSKY, Noam. *Knowledge of Language:* its nature, origin and use. New York: Praeger, 1986.

CHOMSKY, Noam. *O Programa Minimalista*. Lisboa: Caminho, 1999.

CHOMSKY, Noam. Minimalist inquires. *In*: MARTIN, Roger; MICHAELS, David; URIAGEREKA, Juan (eds.). *Step by step*. Cambridge, MA: MIT Press, 2000. pp. 89-155.

CHOMSKY, Noam. Derivation by phase. *In*: KENSTOWICZ, Michael (ed.). *Ken Hale*: A life in language. Cambridge, MA: MIT Press, 2001. p. 1-52.

CHOMSKY, Noam. Beyond explanatory adequacy. *In*: BELLETTI, Ana (ed.). *Structures and beyond.* Oxford: Oxford University, 2004. pp. 104-131.

CYRINO, Sonia. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. *In*: ROBERTS, Ian; KATO, Mary (org.). *Português brasileiro:* Uma viagem diacrônica. Campinas - SP: Editora da Unicamp, 2018 [1993]. pp. 163-184.

CYRINO, Sonia. *O objeto nulo no português do Brasil*: um estudo sintático diacrônico. 1994. 217 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1994.

CYRINO, Sonia. Objetos nulos em português brasileiro. *Cuadernos de la ALFAL*, [*s. l.*], v. 2,n. 12, pp. 387-410, nov. 2020. Disponível em: https://www.mundoalfal.org/sites/default/files/revista/12_2_cuaderno_017.pdf. Acesso em: 16 jan. 2026.

DEMONTE, Violeta; MASULLO, Pascual José. La predicación: los complementos predicativos. *In*: BOSQUE, I.; DEMONTE, V. (dir.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. pp. 1461-2524.

DUARTE, Maria Eugenia. *Variação e sintaxe: clítico acusativo, pronome lexical e categoria vazia no português do Brasil*. 73 f. Dissertação de Mestrado - PUC-SP, São Paulo, 1986.

DUARTE, Maria Eugenia. Clítico acusativo, pronome lexical e categoria vazia no português do Brasil. *In*: TARALLO, Fernando (org.). *Fotografias sociolingüísticas*. Campinas - SP: Pontes - Ed. da Unicamp, 1989. pp. 19-34.

EZEIZABARRENA, María José. Morfemas de concordancia con el sujeto y con los objetos en el castellano infantil. *In*: PÉREZ-LEROUX, Ana Teresa; GLASS, William (eds.). *Contemporary perspectives on the acquisition of Spanish.* Somerville, MA: Casacadilla, 1997. v 1, pp. 21-36. (Developing Grammars).

FARACO, Carlos Alberto. *Linguística histórica. Uma introdução ao estudo da história das línguas*. São Paulo: Parábola, 2006.

FUJINO, Hanako; SANO, Tetsuya. Aspects of the null object phenomenon in child Spanish. *In*: PÉREZ-LEROUX, Ana Teresa; LICERAS, Juana (eds.). *The acquisition of Spanish morphosyntax. The L1/L2 connection.* Dordrecht, Boston, London: Kluwer, 2002. pp. 67-88.

GALVES, Charlotte. *Ensaios sobre as gramáticas do português*. Campinas: Editora da Unicamp, 2001.

GALVES, Charlotte. O enfraquecimento da concordância no português brasileiro. *In*: ROBERTS, Ian; KATO, Mary (org.). *Português brasileiro*: Uma viagem diacrônica. Campinas - SP: Editora da Unicamp, 2018 [1993]. pp. 387-408.

GROPPI, Mirta. P*ronomes pessoais no português do Brasil e no espanhol do Uruguai*. 1997. 152 f. Tese (Doutorado em Linguística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 1997.

HUANG, James. On the distribution and referente of empty pronouns. *Linguistic Inquiry*, v. 15, n. 4, pp. 531-574, 1984.

KANY, Charles. *Sintaxis hispanoamericana.* Madrid: Gredos, 1969.

KATO, Mary; TARALLO, Fernando. Anything YOU can do in Brazilian Portuguese. *In*: JAEGGLI, Oswaldo; SILVA-CORVALÁN, Carmen (eds.). *Studies in Romance Linguistics*. Dordrecht: Foris, 1986. pp. 346-358.

KATO, Mary. Pronomes fortes e fracos na sintaxe do Português Brasileiro. *Revista Portuguesa de Filologia*, Coimbra, Portugal, v. XX, pp. 101-122, 2002.

KATO, Mary. A gramática do letrado: questões para a teoria gramatical. *In*: MARQUES, Maria Aldina *et al.* (org.). *Ciências da Linguagem:* trinta anos de investigação e ensino. Braga: CEHUM (Universidade do Minho), 2005. pp. 131-145.

KATO, Mary; TARALLO, Fernando. Anything YOU can do in Brazilian Portuguese. *In*: JAEGGLI, Oswald; SILVA-CORVALÁN, Carmen. (eds.). *Studies in Romance Linguistics*. Dordrecht: Foris, 1986. pp. 346-358.

KATO, Mary; MARTINS, Ana Maria; NUNES, Jairo. Objetos nulos e possessivos nulos. *In*: *Português brasileiro e português europeu.* Sintaxe comparada. São Paulo: Contexto, 2023. pp. 265-293.

LABOV, William. *Padrões Sociolingüísticos*. São Paulo: Parábola Editorial, 2008.

LANDA, Alazne. *Conditions on null objects in Basque Spanish and their relation to leísmo and clitic doubling*. 1995. 252 f. Tese (Doutorado em Linguística) - University of Southern California, Los Angeles.

LANDA, Allazne. Los objetos nulos determinados del español del País Vasco. *Lingüística*, n. 5, pp. 131-146, 1993.

LEONETTI, Manuel. El artículo. *In*: BOSQUE, Ignacio; DEMONTE, Violeta (dir.). *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa, 1999. pp. 787-890.

LIGHTFOOT, David. *The development of language. Acquisition, change, and evolution*. Malden, Mass.: Blackwell, 1999.

LIGHTFOOT, David. *How new languages emerge*. New York: Cambridge, 2006.

LOPES, Ruth. Traços semânticos na aquisição da linguagem. *Letras de Hoje*, v. 41, pp. 161-178, 2006.

LOPES, Ruth. Acquiring a null object language. When children sound target-like but really aren't. *In*: BAAW, Sergio; VAN KAMPEN, Jacqueline; PINTO, Manuela (eds.). *The acquisition of Romance Languages. Selected papers from the Romance Turn II*. LOT Occasional Series. Utrecht: Utrecht University, 2007. pp. 113-126.

MAGALHÃES, Telma. *O sistema pronominal sujeito e objeto na aquisição do Português Europeu e do Português Brasileiro*. 175 f. Tese de Doutorado em Linguística - IEL/Unicamp, São Paulo, 2006.

SIMÕES, Adriana Martins. La adquisición/aprendizaje del español por aprendices brasileños: evidencias acerca de la gramática no nativa. *Signo y seña*, n. 22, pp. 87-104, 2012. Disponível em: https://dialnet.unirioja.es/servlet/articulo?codigo=5033399. Acesso em: 19 jan. 2026.

SIMÕES, Adriana Martins. *O objeto pronominal acusativo de 3ª pessoa nas variedades de espanhol de Madri e Montevidéu comparado ao português brasileiro:* clíticos como manifestação visível e objetos nulos como manifestação não visível da concordância de objeto. Tese de Doutorado – DLM/FFLCH/USP, São Paulo, 2015. Disponível em: https://www.teses.usp.br/teses/disponiveis/8/8145/tde-09092015-175408/pt-br.php. Acesso em: 14 abr 2026.

SIMÕES, Adriana Martins. Uma proposta de análise para o objeto acusativo anafórico na variedade de espanhol de Madri e no português brasileiro de São Paulo. *Revista de Estudos da Linguagem*, v. 30, n. 4, pp. 1547-1595, 2022.

SUÑER, Margarita; YÉPEZ, María. Null definite objects in Quiteño. *Linguistic Inquiry*, v. 14, pp. 561-565, 1988.

TARALLO, Fernando. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. *In*: ROBERTS, Ian; KATO, Mary. (org.). *Português brasileiro:* Uma viagem diacrônica. Campinas - SP: Editora da Unicamp, 2018 [1993]. pp. 35-68.

THRÁINSSON, Höskuldur. Object shift and scrambling. *In*: BALTIN, M.; CoLLINS, C. (eds.). *The handbook of contemporary syntactic theory*. Oxford: Blackwell, 2008. pp. 148-202.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança lingüística*. São Paulo: Parábola Editorial, 2009.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72251

A expressão do sujeito pronominal em peças de teatro capixabas: uma análise em tempo real da 2ª pessoa do discurso

The expression of the pronominal subject in capixaba plays: a real-time analysis of the second person of discourse

Iesa Venturin Mulinari de Rodran[1]

Leila Maria Tesch[2]

RESUMO

Neste trabalho, propõe-se um recorte do estudo realizado em Rodran (2025), acerca a evolução, na variedade capixaba, da mudança rumo a maiores frequências de expressão do sujeito pronominal, já reportada por trabalhos como os de Paredes Silva (1988, 2003) e Duarte (1993, 1995). O estudo tem como base a Sociolinguística Variacionista (Labov, 2008 [1972]; Weinreich, Labov, Herzog, 2006 [1968]) e a Sociolinguística Histórica (Romaine, 2009 [1982]; Conde Silvestre, 2007). São apresentados somente os resultados da 2ª pessoa do discurso, a mais proeminente no processo de mudança. Nosso *corpus* é composto por 12 peças teatrais capixabas, divididas em quatro períodos/sincronias, e publicadas entre os anos de 1889 e 2004. As variáveis linguísticas observadas foram: pessoa do discurso (pronome), ambiguidade, ênfase, tipo sintático da oração e conexão discursiva. Na dimensão social, foram observadas as variáveis fase da vida, sexo/gênero, grau de instrução e classe social dos personagens, além do período/sincronia das peças. Ao todo, para todas as pessoas do discurso, foram coletados 4785 dados, sendo 1076 de 2ª pessoa. Os resultados indicam substituição de *tu* por *você* a partir da década de 1930 (período II), com salto tanto no número de dados quanto na expressão. O pronome *tu* apenas aparece nos períodos I (séc. XIX) e II, sempre com baixas frequências de expressão. A presença de ambiguidade contextual, além da morfológica, é grande impulsionadora da mudança, assim como os contextos enfáticos. Apesar da proeminência das variáveis linguísticas, as sociais também mostraram relevância, indicando as classes mais baixas e menos instruídas como *locus* inicial da mudança.

PALAVRAS-CHAVE: Sociolinguística. Sujeito pronominal. Mudança linguística. Peças teatrais. Variedade capixaba.

ABSTRACT

In this paper, we propose a focused analysis of the study conducted in Rodran (2025) on the evolution, in the Espírito Santo (Capixaba) variety, of the change toward higher frequencies of overt pronominal subject expression, as already reported in works such as Paredes Silva (1988, 2003) and Duarte (1993, 1995). The study is grounded in Variationist Sociolinguistics (Labov, 2008 [1972]; Weinreich, Labov: Herzog, 2006 [1968]) and Historical Sociolinguistics (Romaine, 2009 [1982]; Conde Silvestre, 2007). Only the results for the second person are presented here, as this is the most prominent category in the change process. Our *corpus* consists of 12 Capixaba plays, divided into four periods/synchronies and published between 1889 and 2004. The linguistic variables examined were: person of discourse (pronoun), ambiguity, emphasis, syntactic type of clause, and discourse connectivity. On the social dimension, the variables observed were life stage, sex/gender,

[1] Universidade Federal do Espírito Santo (UFES/ES), iesavmr@gmail.com, https://orcid.org/0009-0006-8619-5081.

[2] Universidade Federal do Espírito Santo (UFES/ES), leilatesch@gmail.com, https://orcid.org/0000-0003-3919-1230.

Recebido em 20/01/2026 - Aceito em 10/06/2026

level of education, and social class of the characters, in addition to the period/synchrony of the plays. In total, 4,785 tokens were collected, of which 1,076 correspond to second person. The results indicate a replacement of *tu* by *você* starting in the 1930s (Period II), with a sharp increase both in the number of tokens and in overt expression. The pronoun *tu* appears only in Periods I (19th century) and II, always with low rates of overt expression. The presence of contextual ambiguity, in addition to morphological ambiguity, strongly drives the change, as do emphatic contexts. Despite the prominence of linguistic variables, social variables also proved to be relevant, indicating the lower and less educated social classes as the initial *locus* of the change.
KEYWORDS: Sociolinguistics. Pronominal subject. Linguistic change. Theatrical plays. Espírito Santo Portuguese variety.

1. Introdução

O sujeito pronominal pode ser realizado de duas maneiras:

● Sujeito nulo[3]

(1) *COMANDANTE - Agora meu sobrinho, lembra-te de que Ø és soldado do 30 B. C. Esse Batalhão tem uma bandeira tecida das mesmas glórias que imortalizaram Greenhalgh, Pedro Affonso e os 18 heróes de Copacabana. [...]* (O voluntário do 3º B.C., Ernesto Guimarães, 1938)

● Sujeito pleno[4]

(2) *JAIRO - **Você** está no seu terceiro casamento, não é?*
IZAURA - (Bem humorada) É. É difícil encontrar a nossa cara metade.
*JAIRO - Olhando você, sentada aqui, parece que **você** nunca saiu daqui.* (Uma última vez, Vera Viana, 1985)

A realização plena do sujeito pronominal no português é considerada excepcional pela tradição gramatical. Gramáticos como Rocha Lima (2011 [1972]), Cunha e Cintra (2017 [1985]) e Bechara (2019 [1999]) recomendam o uso do sujeito pronominal expresso apenas para dar ênfase ao sujeito ou para desfazer ambiguidade entre formas verbais comuns a mais de uma pessoa do discurso. Essa perspectiva se configura devido à compreensão de que as desinências distintivas dos verbos ainda são as principais responsáveis pela identificação do sujeito, pois são exclusivas para cada pessoa do discurso, assim, o pronome correspondente não precisa ser explicitado. Essa afirmação, por sua vez, está baseada em uma visão do quadro de pronomes pessoais do português brasileiro (PB) muito próxima ainda do português de Portugal (PP) e que não considera o amplo uso das formas *você*, *vocês* e *a gente* como pronomes pessoais.

Conforme Tarallo (2018 [1993]) e Castilho (2010), é no século XIX, em meio à busca por uma identidade nacional, que mudanças linguísticas começam a se intensificar no Brasil, afastando o PB

[3] Também chamado de sujeito não expresso ou sujeito não preenchido.

[4] Também chamado de sujeito expresso ou sujeito preenchido.

do PP. Castilho (2010) aponta, no entanto, que a gramaticalização dos pronomes *você*, *vocês* e *a gente* só se implementa e se difunde no século XX.

Ao serem incorporadas, no uso linguístico, ao quadro de pronomes pessoais do PB, essas novas formas pronominais provocaram uma redução no paradigma flexional dos verbos. O quadro 1, elaborado por Duarte (1993) a partir de peças teatrais cariocas, ilustra essa evolução por meio de verbo regular no presente do indicativo. No paradigma 1, ainda existem seis formas verbais distintivas, uma para cada pessoa, mas o pronome *vocês* já está presente compartilhando a desinência com a 3ª pessoa do plural, *eles*.

Quadro 1: Evolução nos paradigmas flexionais no português do Brasil

	Pronomes nominativos	**Paradigma 1 Século XIX**	**Paradigma 2 Século XX/1**	**Paradigma 3 Século XX/2**
1PS	eu	cant*o*	cant*o*	cant*o*
1PP	nós *a gente*	canta*mos* -	canta*mos* cantaØ	canta*mos* cantaØ
2PS	tu *você*	cantas -	cantas cantaØ	canta(*s)* cantaØ
2PP	vós *vocês*	canta*is* canta*m*	- canta*m*	- canta(*m)*
3PS	ele, ela	cantaØ	cantaØ	cantaØ
3PP	eles, elas	canta*m*	canta*m*	canta(*m)*

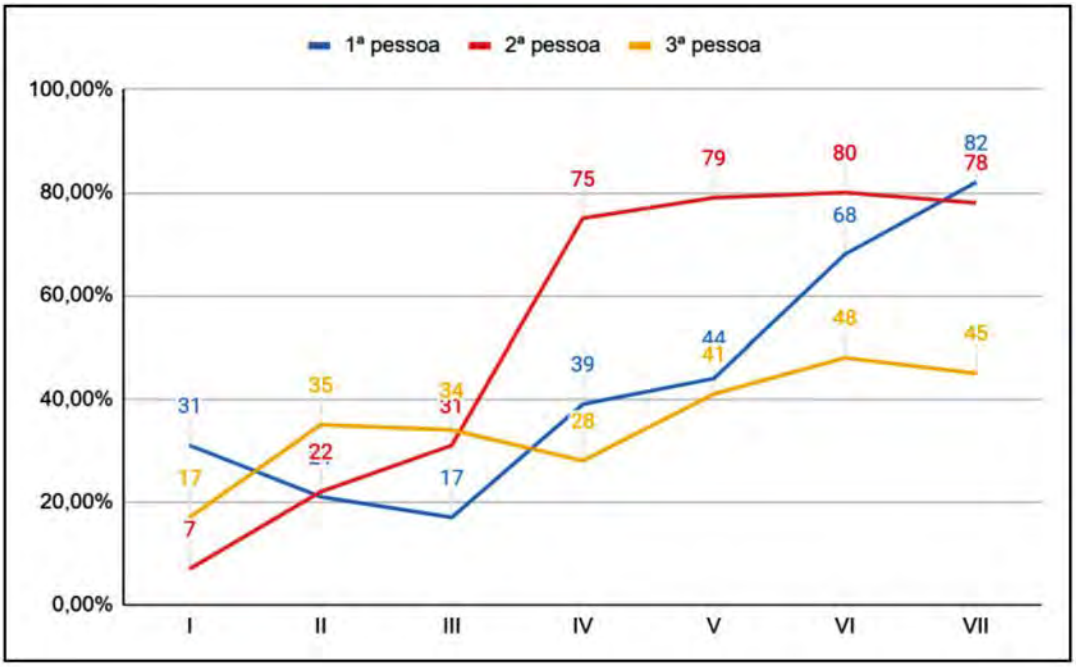
Fonte: Duarte (2012), adaptado de Duarte (1993, p. 85)

No paradigma 2, primeira metade do século XX, *você* está em variação com *tu*, mas utilizando a mesma desinência da 3ª pessoa do plural, assim como ocorre com *a gente* em variação com *nós*, ou seja, o número de desinências distintivas cai para quatro. A forma *vós* já não aparece mais nos paradigmas 2 e 3. No paradigma 3, o que muda é a possibilidade de não realização da concordância verbal, o que reduz o número de desinências distintivas para o mínimo de três. Em suma, uma vez que são derivadas de sintagmas nominais, as novas formas pronominais fazem concordância com a 3ª pessoa do discurso, de modo que o compartilhamento de formas verbais se tornou muito mais comum, aumentando as possibilidades de ambiguidade, e, consequentemente, a necessidade de expressão do pronome sujeito.

Duarte (1993), a partir de sete peças teatrais cariocas correspondentes a sete sincronias que vão de 1845 a 1992, encontra a distribuição de frequências de sujeitos plenos representada no gráfico 1 para cada uma das três pessoas do discurso. Liderando a mudança está a 2ª pessoa, que apresenta um salto na expressão do sujeito pronominal entre as sincronias III e IV (1918 e 1937, respectivamente) e mantém altas frequências até a última sincronia. Em segundo lugar, está a 1ª pessoa, cujas frequências de expressão apresentam tendência de aumento gradual a partir da sincronia III. A 3ª pessoa, por sua

vez, mantém-se em último lugar a partir do período IV, apesar de também ter mostrado aumento na expressão do sujeito. Além disso, indicando uma correlação entre esses processos, observa-se uma correspondência entre os resultados de Duarte (1993) e os períodos de consolidação dos novos pronomes: décadas de 1920/30 para *você*, segundo Lopes (2008, p. 2), e fins do século XX para *a gente*, conforme Zilles (2005, *apud* Zilles, 2007, p. 32).

Gráfico 1: A trajetória das três pessoas do discurso ao longo do tempo (frequência de sujeitos plenos)[5]



Fonte: Adaptado de Duarte, 1993, p. 117 (retirado de Duarte, 2012, p. 22)

Podemos atribuir essa hierarquia entre as três pessoas do discurso ao fato de que a necessidade de expressão dos pronomes sujeito de 2ª pessoa foi maior, por causa da incorporação de novos pronomes tanto no singular quanto no plural. A 1ª pessoa evolui mais gradualmente, pois a consolidação do uso de *a gente* em variação com *nós* acontece mais tarde, enquanto a 3ª pessoa manteve certa estabilidade, apesar de também mostrar aumento na frequência de sujeitos plenos, uma vez que não sofreu alteração em seus pronomes.

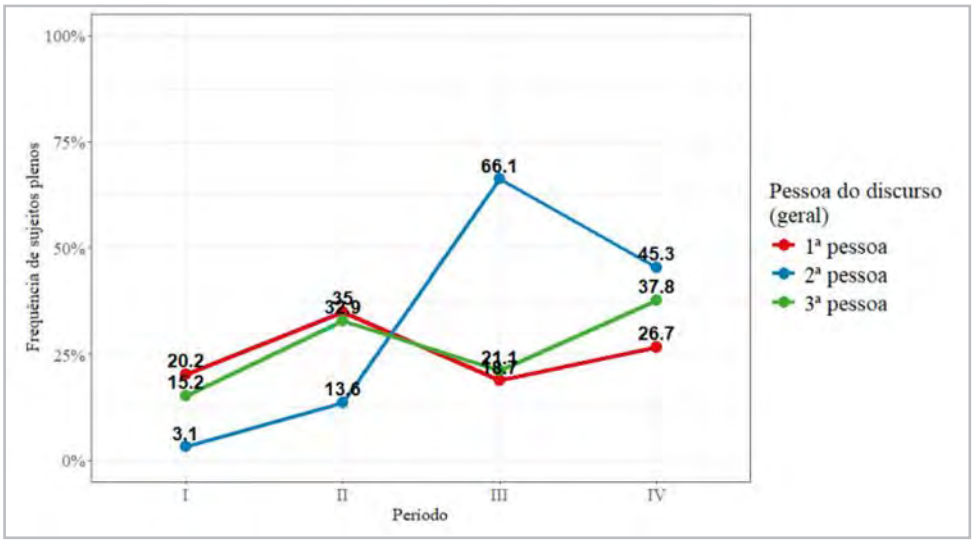
A comparação entre os estudos de Genuíno (2017) e Massariol (2018) nos permite afirmar que a mudança também vem ocorrendo na variedade capixaba. Massariol (2018) encontra preferência por sujeitos nulos (apenas 13,9% de sujeitos plenos) em cartas e postais da primeira metade do século XX. Já na fala da comunidade de fala de Vitória[6] do início dos anos 2000, Genuíno (2017) confirma a preferência por sujeitos expressos, com 70,9% de frequência. Além disso, atesta as maiores frequências de expressão com as novas formas pronominais *você*, *vocês* e *a gente*.

[5] Períodos das peças de Duarte (2018 [1993]): I = 1845, II = 1882, III = 1918, IV = 1937, V = 1955, VI = 1975, VII = 1992.

[6] Análise do *corpus* PortVix 2000.

Tomando como referência os estudos citados, em Rodran (2025), realizou-se um estudo em tempo real acerca do fenômeno da expressão do sujeito pronominal utilizando-se peças teatrais capixabas publicadas entre 1889 e 2004, a fim de traçar o percurso da mudança na variedade capixaba do português brasileiro. O resultado geral, ilustrado no gráfico 2, indica a 2ª pessoa como líder dessa mudança, sendo a passagem da década de 1930 para as décadas de 1970 e 1980 (sincronias II e III, respectivamente) o período de maior contraste. Por outro lado, a 1ª e a 3ª pessoas seguiram uma trajetória muito parecida, com tendências de queda e aumento nas mesmas sincronias.

Gráfico 2: Frequência relativa de sujeitos plenos nas três pessoas do discurso em tempo real[7]



Fonte: Rodran (2025, p. 136).

Assim, considerando a proeminência da 2ª pessoa no resultado geral e sua conexão com os resultados de Duarte (2018 [1993]) e com as afirmações de Lopes (2008), propõe-se, no presente artigo, um recorte do estudo realizado em Rodran (2025), em que serão apresentados somente os resultados referentes à 2ª pessoa do discurso. Todas as variáveis observadas em Rodran (2025) serão retomadas, contemplando tanto a dimensão linguística quanto a dimensão social da mudança.

2. Fundamentação teórico-metodológica

Neste trabalho, a principal base teórico-metodológica é a Sociolinguística Variacionista (Labov, 2008 [1972]; Weinreich, Labov, Herzog, 2006 [1968]), cujo principal pressuposto é o de que todas as línguas naturais apresentam variação, ou seja, "a opção de dizer a mesma coisa de várias maneiras diferentes, isto é, as variantes são idênticas em valor de verdade ou referencial, mas se opõem em sua significação social e estilística." (Labov, 2008 [1972], p. 313). A variação linguística é o que permite

[7] Sincronias estabelecidas em Rodran (2025): I = décadas de 1880/90; II = década de 1930; III = décadas de 1970/80; IV = décadas de 1990/2000.

que uma língua mude ao longo do tempo; entretanto, não é aleatória, mas condicionada tanto por aspectos linguísticos quanto por aspectos sociais, ou seja, todas as línguas naturais são dotadas de heterogeneidade ordenada.

Uma mudança linguística em progresso pode ser identificada quando se observa que o uso de uma variante mais recente, ou inovadora, tende a superar uma variante que já estava sendo usada há mais tempo, ou variante conservadora. Para o estudo desse processo, Weinreich, Labov e Herzog (2006 [1968], p. 125) estabelecem cinco questões a serem investigadas empiricamente: o problema dos fatores condicionantes, o problema da transição, o problema do encaixamento, o problema da avaliação e o problema da implementação. O problema dos fatores condicionantes diz respeito aos contextos linguísticos e sociais que atuam favorecendo ou desfavorecendo a evolução da mudança. O problema da transição refere-se à maneira que se dá a transmissão da variante inovadora entre gerações. O encaixamento é observado pela distribuição das variantes tanto na matriz linguística (encaixamento linguístico), quanto na matriz social (encaixamento social), por meio da classificação dos falantes por categorias sociais como escolaridade, sexo/gênero, idade, etc. A avaliação interfere no processo de mudança à medida que a associação de um significado social neutro ou positivo à variante emergente pode impulsionar sua difusão, enquanto um significado social negativo pode inibir seu uso. O problema da implementação, por fim, consiste em procurar explicar as motivações para a ocorrência de uma mudança linguística em determinada língua, em determinada comunidade de fala, em determinado momento, e não em outros.

Em geral, busca-se, no estudo de uma mudança linguística, observar a integração da nova variante tanto no sistema linguístico quanto na estrutura social da comunidade de fala investigada. Conforme Labov (2008, [1972], p. 320), uma mudança linguística "só se torna parte da língua quando é adotada pelos outros, ou seja, quando se propaga. Por conseguinte, a origem de uma mudança é sua propagação ou aceitação pelos outros [...]", processo que ocorre de maneira gradual. De acordo com o percurso da variante inovadora pela matriz social, Labov (1994) estabelece dois tipos de mudança linguística: a mudança de cima (*change from above*) e a mudança de baixo (*change from below*), que "referem-se simultaneamente a níveis de consciência social e posições na hierarquia socioeconômica" (Labov, 1994, p. 78). A mudança de cima envolve uma variante que apresenta prestígio social, ou seja, é avaliada positivamente pela comunidade de fala, e geralmente está associada ao falar de classes sociais mais abastadas. No entanto, conforme Labov (1994), a classe média alta é que costuma liderar esse tipo de mudança, pois tem a fala da classe alta como referência, de maneira que as formas inovadoras envolvidas nesse tipo de mudança costumam aparecer inicialmente na fala formal, atingindo o vernáculo apenas posteriormente. Enquanto isso, as mudanças de baixo costumam surgir no vernáculo, pois geralmente estão abaixo do nível de consciência dos falantes, não sendo percebidas nem avaliadas. De acordo com Labov (1994, p. 78), "podem ser introduzidas por qualquer classe social, embora não tenham sido registrados casos em que o grupo social de status mais elevado atue como grupo inovador".

Neste trabalho, propomos investigar a mudança em direção à preferência pelo sujeito expresso por meio da comparação de resultados em diferentes sincronias, caracterizando, assim, uma análise em tempo real. Contudo, uma vez que o material linguístico analisado é constituído por peças teatrais, não lidamos diretamente com falantes empíricos, mas com personagens construídos pelos autores. Diante disso, recorremos também à Sociolinguística Histórica (Romaine, 2009 [1982]; Conde Silvestre, 2007), que oferece suporte teórico-metodológico para a investigação sociolinguística em sincronias passadas, especialmente quando se fazem necessárias adaptações como o uso de textos escritos, a reconstrução histórica das categorias sociais e a interpretação adequada dos resultados à luz do contexto sócio-histórico de cada período, de modo a evitar anacronismos. Nesse sentido, as peças teatrais constituem uma fonte privilegiada, uma vez que são textos escritos para posterior oralização e apresentam estrutura dialogal, buscando simular conversações espontâneas. Essas características possibilitam o estudo histórico da língua com relativa proximidade da oralidade (Duarte, 1993; Berlinck; Brandão, 2021).

3. Procedimentos metodológicos

3.1. Constituição do *corpus*: definição das peças teatrais

O *corpus* em análise é composto por 12 peças teatrais escritas por autores capixabas ou que viveram a maior parte da vida no estado do Espírito Santo escrevendo sobre e para o público capixaba. Foram selecionadas somente "comédias de costumes", gênero de cunho popular em que as histórias são ambientadas na vida cotidiana, "a fim de obter uma amostra o mais representativa possível da linguagem do dia a dia, e também uma caracterização dos personagens o mais próxima possível da realidade de cada época." (Rodran, 2025, p. 87). Além disso, buscou-se peças com personagens de perfis sociais diversos e histórias que se passam no Espírito Santo e na mesma época em que foram escritas, "a fim de garantir ao máximo uma representatividade satisfatória da comunidade capixaba e das categorias sociais" (Rodran, 2025, p. 87).

As peças abarcam um período de mais de 100 anos, da década de 1880 à década de 2000. Para a execução do estudo em tempo real, as sincronias foram estabelecidas levando em consideração momentos de inflexão na história do PB, principalmente no que se refere ao sujeito pronominal. Assim, o período I representa o final do século XIX, momento em que, segundo Tarallo (2018 [1993]) e Castilho (2010), houve a intensificação de mudanças linguísticas no PB, tornando-o cada vez mais distante do PP. O período II, representando a década de 1930, é apontado por Lopes (2008), como o momento em que o uso de *você* supera o uso de *tu*. Conforme Zilles (2005, *apud* Zilles, 2007), nas décadas de 1970 e 1980, correspondentes ao período III, é observada a consolidação do uso de *a gente* em variação com *nós*. O período IV, por fim, corresponde às décadas de 1990 e 2000, cerca de 20 anos à frente do período III, momento em que se espera que a mudança já tenha avançado o suficiente para atingir até mesmo os contextos que eram resistentes. No quadro 2 são apresentadas algumas informações básicas sobre as peças selecionadas, assim como a distribuição dos dados entre elas e entre as sincronias.

Quadro 2: Informações básicas sobre as peças teatrais capixabas selecionadas para a análise

Período/ Sincronia	Nº	Ano de publicação	Autor	Título	Quantidade de dados	
Período I (Décadas de 1880 e 1890)	1	1889	Aristides Freire	A república na roça	211	1457
	2	1890	Amâncio Pereira	O tio Mendes	390	
	3	1894	Amâncio Pereira	Virou-se o feitiço	592	
	4	~1897	Aristides Freire	Amigo íntimo	264	
Período II (Década de 1930)	5	1938	Ernesto Guimarães	O voluntário do 3º B.C.	232	1027
	6	1939	Ernesto Guimarães	A princesinha	795	
Período III (Décadas de 1970 e 1980)	7	1975	Vera Viana	Conversa de bar	189	811
	8	1978	Vera Viana	Fila eterna	256	
	9	1985	Vera Viana	Uma última vez	366	
Período IV (Décadas de 1990 e 2000)	10	1998	Edilson Pedrini	Por favor, matem minha empregada!	364	1490
	11	2000	Jovany Sales Rey	A volta da vedete	594	
	12	2004	Edilson Pedrini	Por favor, matem meu patrão!	532	

Fonte: Rodran (2025, p. 89).

No processo de seleção de peças, foi levado em conta o número de páginas de cada peça para assegurar ao máximo uma uniformidade na quantidade de dados. Ao todo, foram coletados 4785 dados, dos quais 1076 são de 2ª pessoa, sobre a qual daremos enfoque neste artigo.

3.2. Seleção dos dados

Em Rodran (2025), foram selecionados dados apenas de sentenças finitas, com sujeitos de referência definida, como exemplificado em (3) e (4).

(3) *JAIRO - É ridículo...* ***Você*** *tem tanta coisa pra fazer... Tantos compromissos… Aqui nessa praça... Desculpe!* (Uma última vez, Vera Viana, 1985)

(4) *CATARINA - Ora, deixe-se ficar. É que a senhora, D. Faustina, nessa idade, já perdeu o gosto para viagens... Mas vai gostar daqui... Eu farei tudo para agradá-las...* ***Vocês*** *me vão fazer excelente companhia... Vivo tão só…* (A princesinha, Ernesto Guimarães, 1939)

Não foram incluídos na análise da 2ª pessoa, pois consideramos que esses casos demandam uma análise própria: sujeitos com referência genérica; verbos nas formas nominais e no imperativo; sujeitos pospostos ao verbo; ocorrências com ausência de concordância verbal; sentenças existenciais com os verbos *ter* e *haver*.

3.3. Tratamento estatístico dos dados

A codificação dos dados de acordo com cada variável independente e seus fatores foi realizada no *Google* Planilhas. Posteriormente, o tratamento estatístico foi feito por meio do *software* R (R Core Team, 2019), que permite análises descritivas, gerando tabelas de proporções e gráficos, além de análises inferenciais, ao possibilitar a realização de testes estatísticos univariados e multivariados.

Neste trabalho, apresentamos somente a análise descritiva, a partir da observação das frequências de sujeitos plenos. Os resultados dos testes uni e multivariados[8] não serão discutidos. No entanto, é importante notar que a ausência da variável "modo e tempo verbal" na discussão dos resultados se explica pelo resultado do teste de qui-quadrado de Pearson, teste univariado, ou seja, "que correlaciona uma variável dependente a uma variável independente" (Oushiro, 2022, p. 19), que apontou a referida variável como não significativa para a 2ª pessoa.

3.4. Variáveis independentes

Com base nos estudos de referência sobre o fenômeno, foram estabelecidas 12 variáveis independentes, ou grupos de fatores, das quais cinco são linguísticas, a saber: pessoa do discurso (com separação entre singular e plural), ambiguidade, ênfase, tipo sintático da oração e conexão discursiva; e quatro são extralinguísticos, a saber: fase da vida (dos personagens), sexo/gênero (dos personagens), instrução (dos personagens) e classe social (dos personagens). É importante notar que a variável "período da peça" permeia a análise de todas as outras variáveis, e a divisão das sincronias já foi descrita e justificada no início desta seção. As variáveis cujos resultados são discutidos neste artigo estão descritas abaixo:

Variáveis e fatores linguísticos:

i) Pessoa do discurso: indica ao mesmo tempo a pessoa do discurso - se primeira, segunda ou terceira (neste trabalho, apenas trataremos da 2ª pessoa), se singular ou plural, e a forma verbal utilizada para fazer referência a essa pessoa discursiva, se conservadora ou inovadora.
Fatores: 2ª pessoa do singular conservadora (2PSC) = tu; 2ª pessoa do singular inovadora (2PSI) = você; 2ª pessoa do plural conservadora (2PPC) = vós; 2ª pessoa do plural inovadora = vocês (2PPI).

ii) Ambiguidade: refere-se ao compartilhamento de formas verbais por mais de uma pessoa do discurso. A tradição gramatical já previa casos em que isso ocorre, nos tempos do subjuntivo na referência à 1ª e à 3ª pessoas do singular, por exemplo. Entretanto, a entrada das novas formas pronominais *você*, *vocês* e *a gente* ampliou as possibilidades de ambiguidade, e consequentemente, ampliou também a necessidade de expressão do pronome sujeito.
Fatores: sem ambiguidade morfológica; ambiguidade apenas morfológica; ambiguidade morfológica e contextual.

[8] Levando em consideração a longa extensão da análise descritiva, decorrente da quantidade de variáveis e fatores, optamos por abordar os resultados dos testes uni e multivariados em outra oportunidade.

iii) Ênfase: indicados pela tradição gramatical como alguns dos casos excepcionais em que a expressão do sujeito pronominal é recomendada, os contextos enfáticos são aqueles em que se busca dar destaque ao sujeito. Devido à dificuldade de se identificar com precisão a ocorrência de ênfase, Paredes Silva (1988) adota a noção de contrastividade, e, a partir de Chafe (1976, *apud* Paredes Silva, 1988), estabelece algumas situações de contrastividade em que ocorreria também a ênfase, a saber: presença de conectivos contrastivos, verbos de sentido oposto (antônimos), sujeitos diferentes para um mesmo item verbal, complementos distintos para duas ocorrências do mesmo verbo, uma afirmativa e uma negativa atribuídas a núcleos predicados idênticos ou semelhantes, oposição no espaço ou no tempo.

Fatores: contexto enfático; contexto não enfático.

iv) Tipo sintático da oração: conforme considerações de Paredes Silva (1988), a expressão do sujeito pronominal é influenciada pelo tipo de oração na medida em que pode ser revelada uma centralidade conferida ao sujeito, se ele for expresso em uma oração principal, por exemplo; na medida em que há ou não correferência, pois o sujeito já expresso na primeira oração tende a ser omitido na segunda, o que ocorre principalmente entre orações coordenadas; e de acordo com a posição das orações, uma vez que as antepostas tendem a apresentar mais sujeitos plenos que as pospostas.

Fatores: oração absoluta; oração coordenada inicial; oração coordenada não inicial com correferência; oração coordenada não inicial sem correferência; oração principal anteposta; oração principal posposta com correferência; oração principal posposta sem correferência; oração subordinada substantiva com correferência; oração subordinada substantiva sem correferência; oração relativa com correferência; oração relativa sem correferência; oração subordinada adverbial anteposta; oração subordinada adverbial posposta com correferência; oração subordinada adverbial posposta sem correferência; oração interrogativa.

v) Conexão discursiva[9]: esse grupo de fatores foi elaborado por Paredes Silva (1988) com o objetivo de refinar a observação da influência da dimensão discursivo-funcional no fenômeno em cartas pessoais, abarcando questões de referenciação. Em Paredes Silva (2003), a autora realiza uma adaptação dos graus de conexão para estudo com entrevistas sociolinguísticas. Em Rodran (2025), foi utilizada essa adaptação, pois consideramos que as entrevistas se aproximam mais das peças em termos de estrutura textual que as cartas. A variável consiste em sete graus de conexão do discurso, que são inversamente proporcionais à necessidade de preenchimento do sujeito: quanto maior o grau, mais fraca é a conexão do discurso, e maior se torna a necessidade de retomar o referente explicitamente.

Fatores: Grau 1: mesmo referente, tempo e modo verbal em relação à oração adjacente anterior (exemplo 5); Grau 2: mesmo referente, mas o tempo e o modo verbal mudam em relação à oração adjacente anterior (exemplo 6); Grau 3: mesmo referente, mas há mudança de turno de fala (exemplo 7); Grau 4: segundo Paredes Silva (2003, p. 105), há "a interferência de orações com sujeito impessoal

[9] Consideramos importante exemplificar todos os fatores dessa variável devido a sua especificidade. Os exemplos referentes à outras variáveis serão apresentados na seção de discussão dos resultados de acordo com a pertinência da exemplificação, levando em conta a limitação de espaço deste artigo.

entre o sujeito em causa e sua última menção" (exemplo 8); Grau 5: há interferência de orações com sujeitos de outras pessoas gramaticais entre a menção em causa do referente e a sua última menção (exemplo 9); Grau 6: o referente é retomado como sujeito após ter sido mencionado pela última vez em outra função sintática (exemplo 10) e Grau 7: há mudança de tópico ou subtópico discursivo entre a última menção do referente e a menção em causa (exemplo 11).

(5) *PIERRE - [...] Você não quer ficar popular, conhecida??? Então querida, vá ao encontro do destino. Vá fazer uma visitinha a esse tal de Flexal 2001... **você** comparece, **Ø** levanta os problemas, **Ø** cava uma entrevista na imprensa, **Ø** diz que **Ø** vai fazer um jantar beneficente... **Ø** reúne uns convidados, aí, meu amor, todo mundo vai saber quem é Melani Marques.* (Por favor, matem minha empregada!, Edilson Pedrini, 1998)

(6) *ODETE - Estás resmungando, Faustina? Não estás contente?*
FAUSTINA - Tou satisfeita sim.... Então nós não vai pra pensão da D. Maroca?
CATARINA - Só se vocês acharem que a minha casa não presta.
*ODETE Ora, que absurdo, titia. **Ø** Estás vendo, Faustina? **Ø** Magoaste a titiazinha.* (A princesinha, Ernesto Guimarães, 1939)

(7) *HEITOR - (Chega ao F. e grita) Bravos, camaradas ! Vou já abraçá-los. (Volta à boca de cena e exclama) Soldado do 3° B. C. **Ø** Sinto-me orgulhoso.*
*DR. LEAL - E **Ø** tens justa razão. Soldado valente.* (O voluntário do 3º B.C., Ernesto Guimarães, 1938)

(8) *DÉBORA - Rouba. Como já sei disso quando vou ao supermercado compro umas coisinhas a mais, de menor valor e deixo a mostra...é que nem rato. No outro dia não tem mais nada.*
*ROSÂNGELA - **Você** compra e deixa ela roubar?* (Por favor, matem minha empregada!, Edilson Pedrini, 1998)

(9) *ZEBEDEU (entrando) - Estava à vossa procura para dizer-vos que acabei por despachar da melhor maneira, uma sucia de beocios que viera pedir a Isaura em casamento. E isso fiz, por saber que também não desejaes que ella se case, conforme, a pouquinho, **Ø** me garantisses, em vista das minhas suspeitas.* (Virou-se o feitiço, Amâncio Pereira, 1894)

(10) *DELEGADA - (para Mimi) A vigarista estava **te** enganando, falsificou os resultados dos seus exames. **Você** não tem doença nenhuma.* (A volta da vedete, Jovany Sales Rey, 2000)

(11) *MELANI - Você tem razão. Mas **eu** já sei quem vai poder me acompanhar nesta visitinha aonde se esconde a classe operária. (Pierre termina o cabelo de Melani, dá uns últimos retoques. Ela fica muito pior do que antes)*
*PIERRE- Olha querida, seu cabelo está um deslumbre. Uma coisa de louco. **Você** está linda, maravilhosa, absoluta. Com esse visual, não tem prá ninguém, você vai arrasar. Dez a zero, uhhh! [...]* (Por favor, matem minha empregada!, Edilson Pedrini, 1998)

Fatores extralinguísticos:

i) Fase da vida: a observação da faixa etária dos falantes possibilita a análise da distribuição geracional das variantes, na perspectiva do tempo aparente[10]. A definição das categorias internas das variáveis sociais foi baseada em Brandão (2022), que segue, por sua vez, os critérios do Projeto Língua EmCena[11]. No caso da variável "faixa etária", adota-se a nomenclatura "fase da vida", que representa uma noção mais ampla que a classificação apenas numérica. Para estabelecer as categorias dessa variável, Brandão (2022) leva em conta as transformações históricas ocorridas no Brasil: a mudança na expectativa de vida, a alteração nas noções de juventude e velhice, a qualidade dos serviços de saúde em cada época, as relações familiares dos personagens e as diferenças étnicas. Com essas reflexões, Brandão (2022) estabelece os rótulos e critérios para a categorização da fase da vida dos personagens apresentados no quadro 3.

Quadro 3: Rótulo e critérios gerais para análise da fase da vida baseados em Brandão (2022)

Fase da Vida	Idade (Aproximada)	Critério Social
Fase 1 criança e jovem	0 - 15	Crianças ou adolescentes estudando, sem filhos.
Fase 2a jovem adulto	16 - 25	Jovens que estão estudando, sem filhos, ou, se tem filho, o filho deve ser uma criança recém nascida; solteiro, recém-casado, socioeconomicamente dependente de alguém, personagens em protestos, jovens sonhadores.
Fase 2b adulto	26 - 40	Pai ou mãe de personagem da fase 2a ou da fase 1 que não seja recém-nascido; mais experientes que personagens da fase 2a; possuem mais vivência ou que são donos do próprio negócio; possuem cargos que exigem experiência: gerência, por exemplo.
Fase 3a meia idade	40 - 55	Pai ou mãe de pessoa da fase 2, caminhando para a velhice
Fase 3b	56 ou mais	Pai ou mãe de pessoa da fase 3, pessoas ditas velhas, de bengalas com idade mais avançada

Fonte: Rodran (2025, p. 123); adaptado, Brandão (2022, p. 134).

[10] Duas perspectivas são propostas por Labov (1994) para o estudo da mudança linguística: o tempo aparente e o tempo real. Na perspectiva do tempo aparente, observa-se o movimento da mudança entre as gerações de falantes: um uso mais frequente de uma variante inovadora pelas gerações mais jovens e menos frequente pelas gerações mais velhas pode indicar uma substituição da variante conservadora ao longo do tempo. Já um estudo em tempo real consiste em comparar o estado da variação em diferentes sincronias, utilizando para isso os dados dos mesmos falantes (estudo de painel) ou de falantes diferentes em cada sincronia, mas pertencentes à mesma comunidade de fala e com o mesmo perfil social dos participantes anteriores.

[11] O projeto Língua EmCena insere-se no contexto do grupo SoLAr (Núcleo de Pesquisas em Sociolinguística de Araraquara), da Unesp/Araraquara, e tem como objetivo a constituição de um *corpus* de peças teatrais brasileiras para uso em pesquisas linguísticas, além do constante refinamento teórico-metodológico para o estudo desse material. (Berlinck; Brandão; Barbosa; Biazolli; Sene, 2023).

ii) Sexo/gênero: Labov (2001, p. 293) afirma que "as mulheres se adaptam mais do que os homens às normas sociolinguísticas que são explicitamente prescritas, mas se adaptam menos do que os homens quando elas não são." Como discutido em Rodran (2025, p. 124), "isso faz com que as mulheres estejam à frente de mudanças que não são percebidas e não apresentam marcação social negativa, conforme o autor". A partir do postulado de Labov (2001) e considerando que a expressão do sujeito pronominal é um fenômeno que costuma ser reportado como situado abaixo do nível de consciência dos falantes, espera-se que as mulheres estejam à frente dessa mudança.
Fatores: feminino; masculino.

iii) Instrução: o olhar sobre a escolaridade costuma revelar o resultado, no uso linguístico do falante, da influência inibidora exercida pela escola e pelo contato com a língua escrita no que se refere a inovações linguísticas. Quanto maior for o grau de escolaridade, menor tende a ser a aderência às mudanças linguísticas, e vice-versa. Adaptando a noção de escolaridade ao considerar também a ocupação profissional dos personagens e levar em conta as transformações históricas do nível educacional da população brasileira, Brandão (2022) estabelece os quatro graus de instrução apresentados no quadro 4 junto aos critérios para cada um.

Quadro 4: Rótulo e critérios para análise da instrução

Instrução	Característica
NA	Não há qualquer indício acerca da instrução
Sem instrução	Personagens claramente analfabetas, que não sabem ler
Básica	O único indício é de que o personagem sabe ler ou sabe ler e escrever
Técnica	A personagem exerce algum ofício e/ou tem algum tipo de conhecimento especializado sobre algo. (ex: caixeiro/negociante; soldados; policiais; ministros; musicista; pianista; poeta; flautista etc.)
Avançada	A personagem, além de saber ler e escrever, está se preparando para ingressar ou está cursando o ensino superior ou já exerce uma profissão que demanda formação superior (ex: médicos, advogados, engenheiros, etc.)

Fonte: Brandão (2022, p. 130)

Conforme Paiva e Scherre (1999, p. 217), o fato de um fenômeno estar abaixo do nível de consciência não enfraquece o efeito normativo da escola, pois ele ocorre indiretamente por meio da inserção do indivíduo "em uma modalidade mais rígida e mais homogênea como a escrita". Assim, tem-se como hipótese a representação da mudança, inicialmente, por meio da fala dos personagens menos instruídos.

iv) Classe social: conforme Weinreich, Labov, Herzog (2006 [1968], p. 144), uma mudança linguística ocorre quando se propaga atingindo grupos sociais para além do seu grupo de origem. Por isso, o olhar sobre a distribuição da variação pelas classes sociais em cada sincronia pode revelar o

trajeto de difusão social da mudança e indicar o tipo de mudança, se de baixo (*change from below*) ou de cima (*change from above*), conforme estabelecido por Labov (1994). Utilizamos, mais uma vez, a classificação proposta por Brandão (2022), que se baseia principalmente na ocupação profissional dos personagens, levando sempre em conta as principais transformações sociais na história do Brasil desde o século XIX. Os critérios e rótulos para a variável classe social são apresentados no quadro 5.

Quadro 5: Rótulos e critérios para a análise da classe social

Classe Social	Critério
NA	sem qualquer indício de classe
Classe 1	alta sociedade/ alta burguesia/ dominantes
Classe 2	classe emergente/ intermediárias e/ou "classes médias" - junção de duas "classes" - média e pequena burguesia
Classe 3	proletariado/ camponeses/ assalariados (classe subalterna)
Classe 4	oprimidos/ marginalizados

Fonte: Brandão (2022, p. 125)

Como tratamos de um fenômeno que está abaixo do nível de consciência, e que não sofre, portanto, avaliação social, supomos estar lidando com uma mudança de baixo. Além disso, a origem social do uso das formas *você* e *vocês* como pronomes, também não parece indicar uma mudança de cima. Conforme Castilho (2010, p. 479),

> Com o desenvolvimento da burguesia, os novos-ricos quiseram esse tratamento *[vossa mercê]* para eles também. Indignado, o rei passou a reclamar Vossa Majestade para ele, lembrando decerto aos burgueses que uma forca tinha sido convenientemente erigida defronte ao paço, caso eles resolvessem repetir a gracinha. De todo modo, Vossa Mercê e derivados eram um tratamento cerimonioso, dado "pelos de baixo" "aos de cima". Veja como é a roda da fortuna: pois não é que o derivado você passou a ser no PB um tratamento de igual para igual? Para o tratamento cerimonioso, inventou-se o senhor. Em regiões brasileiras em que o tratamento tu continua vigente, o uso de você traz de volta o antigo distanciamento. E onde o tu bateu com as botas, ele e seus derivados ressuscitam, quando se quer afetar distanciamento [...] (Castilho, 2010, p. 479, inserção nossa)

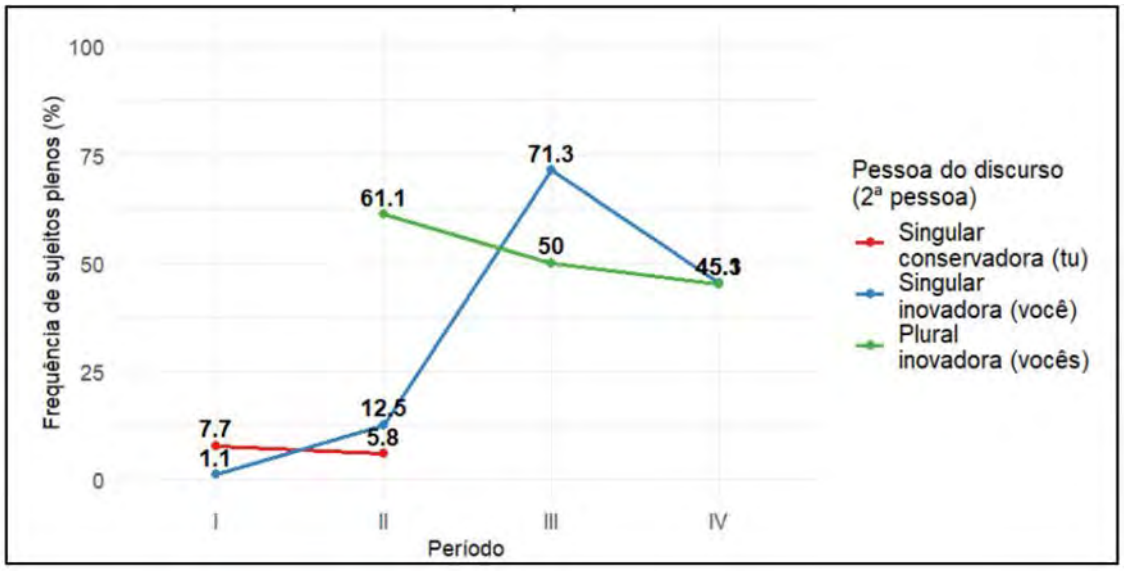
Isso nos leva a esperar que a mudança seja representada nas peças pelas falas dos personagens das classes subalternas, ao menos inicialmente, atingindo também classes mais altas ao longo das décadas. Vejamos, a seguir, o que nos dizem os resultados da análise em tempo real da 2ª pessoa do discurso.

4. Apresentação e discussão dos resultados da variação da expressão pronominal na segunda pessoa

4.1. Resultados das variáveis linguísticas

No que se refere à pessoa do discurso, as ocorrências da desinência de 2ª pessoa do singular, correspondente ao pronome *tu*, apenas ocorrem nas sincronias I e II - 104 e 103 dados, respectivamente - e sempre com índices de expressão pronominal muito baixos, como consta no gráfico 3. Enquanto isso, verifica-se um aumento no número de dados com *você* na análise longitudinal, com 86 no período I, 56 no período II, 143 no período III e 309 no período IV. É no intervalo do período II para o período III, portanto, que o pronome *tu* deixa de ser usado nas peças e a referência à 2ª pessoa do singular passa a ocorrer exclusivamente por meio da desinência de 3ª pessoa do singular, correspondente à forma *você*. Assim, conforme o gráfico 3, observa-se um salto na expressão do pronome *você*, de cerca de 58 pontos percentuais, nesse intervalo, indicando o espaço entre o final da década de 1930 e as décadas de 1970 e 1980 como um período crítico para a manifestação da mudança nas peças teatrais capixabas.

Gráfico 3: Frequência relativa de sujeitos plenos de 2ª pessoa em tempo real de acordo com a variável "pessoa do discurso"[12]



Fonte: Rodran (2025, p. 181)

Na 2ª pessoa do plural, apenas há ocorrências com o pronome *vocês*, que só não aparece expresso na sincronia I. Nas demais, a frequência de expressão desse pronome mantém-se acima de 45%, apesar da queda gradual do período II ao IV. Ocorrências com o pronome *vós* apenas aparecem

[12] Legenda: Período I = décadas de 1880 e 1890; Período II = década de 1930; Período III = décadas de 1970 e 1980; Período IV = décadas de 1990 e 2000.

nas peças da sincronia I, sempre com sujeito nulo[13], assim como observado por Massariol (2018), e fazendo referência à 2ª pessoa do singular e não do plural, como exemplificado em (12) e (13).

(12) *AMBRÓSIO - (Cansado) Capitão, não encontrei nem mais uma pessoa para assinar. (Abrindo o papel) Só* ***eu****, Ambrósio Pereira do Espírito Santo.*
GENOVEVA - Já ***Ø*** *sabeis ler?!...* (A república na roça, Aristides Freire, 1889)

(13) *TOMÁS - (Para Flávio) Ide conversar com ela, ide.* ***Ø*** *Quereis que ela vos venha buscar pela mão? Sempre és marinheiro de primeira viagem (ri-se).* (O tio Mendes, Amâncio Pereira, 1890)

Conforme Menon (1995), no século XVII já começa a aparecer, na escrita, a forma *vossês*, substituindo *vós*, indício de que a mudança na 2ª pessoa do plural se iniciou primeiro, o que pode explicar a predominância completa de *vocês* sobre *vós* já nas peças do século XIX. Como sempre apresenta ambiguidade ao menos morfológica com a 3ª pessoa do plural, é natural que a expressão do pronome *vocês* mantenha-se elevada. Os casos em que o pronome *vocês* foi omitido são aqueles em que o sujeito pode ser contextualmente identificado, como em (14).

(14) *GENOVEVA - E pensa muito bem, Seu Capitão.*
AMBROSIO - Cala a boca, Genoveva, o capitão não fala convosco.
GENOVEVA Ambrósio, Ambrósio! Olha, no reculutamento[14] eu me livro de ti!
ANTONIO - Não briguem, não me contrariem mais do que já estou. ***Ø*** *Podem retirar-se.*
(A república na roça, Aristides Freire, 1889)

Em relação à fala capixaba do início dos anos 2000, as peças do período IV apresentam frequências bem menores de expressão dos pronomes *você* e *vocês*, para os quais Genuíno (2017) encontra 80,2% e 82,2% de sujeitos plenos, respectivamente, enquanto nas peças ambos os pronomes ficam dentro dos 45%. É semelhante, no entanto, a proximidade entre as frequências das duas formas em ambos os estudos. Essas constatações revelam que, no início dos anos 2000, a mudança na 2ª pessoa se encontrava em estágio mais avançado na oralidade que na escrita, e que não havia assimetria entre singular e plural. É importante notar ainda que, diferentemente do observado por Duarte (2018 [1993]) em peça carioca do ano de 1992 e em consonância com Genuíno (2017), nas peças capixabas do período IV não houve quaisquer ocorrências do pronome *tu*.

De maneira geral, é a 2ª pessoa do singular inovadora, correspondente ao pronome *você*, a que conduz a mudança na 2ª pessoa dentro do recorte temporal estudado, uma vez que apresenta o mesmo percurso que o apresentado no gráfico 2 para a 2ª pessoa geral.

Os resultados da variável "ambiguidade" mostram de maneira mais clara o efeito do compartilhamento de formas verbais na referência a mais de uma pessoa discursiva. Conforme

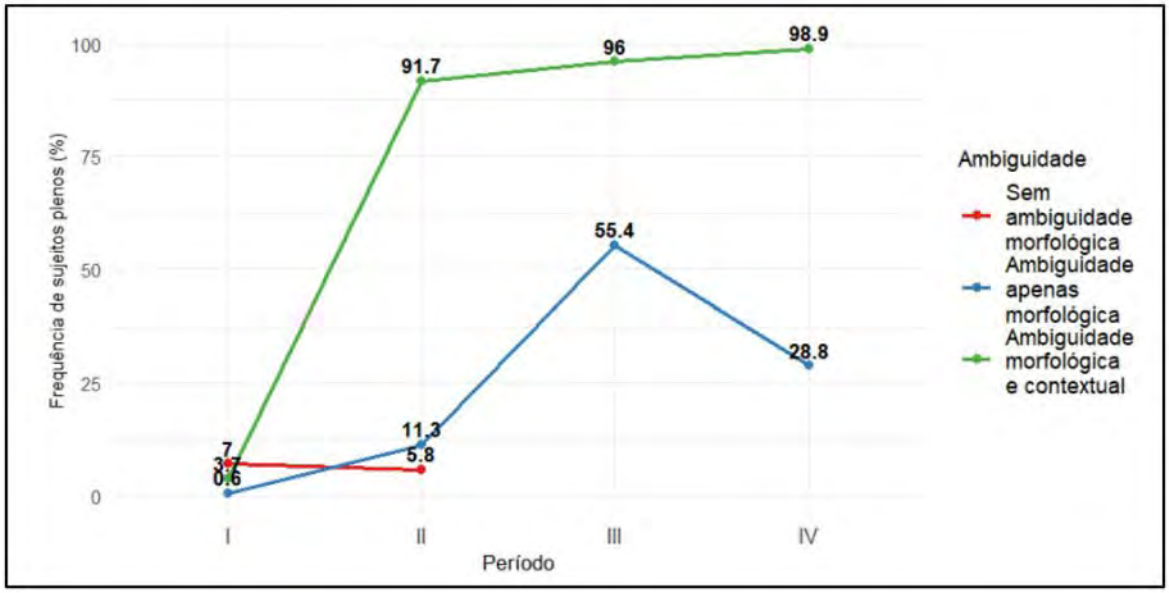
[13] Por isso, essas ocorrências não aparecem no gráfico.

[14] Recrutamento.

indicado no gráfico 4, as formas verbais de 2ª pessoa sem ambiguidade morfológica apenas aparecem nas sincronias I e II (exemplo 15), com frequências baixas de sujeitos plenos:

(15) *HEITOR - Incorporei-me hoje. Vou tirar o meu tempo no 3.° B. C.*
*DR. LEAL - E **Ø** perdes o ano na Faculdade de Medicina? Já no quarto ano? Que loucura!*
*Nem **Ø** esperaste o sorteio!* (O voluntário do 3o B.C., Ernesto Guimarães, 1938)

Gráfico 4: Frequência relativa de sujeitos plenos de 2ª pessoa (geral) em tempo real de acordo com a variável "ambiguidade"[15]



Fonte: Rodran (2025, p. 188)

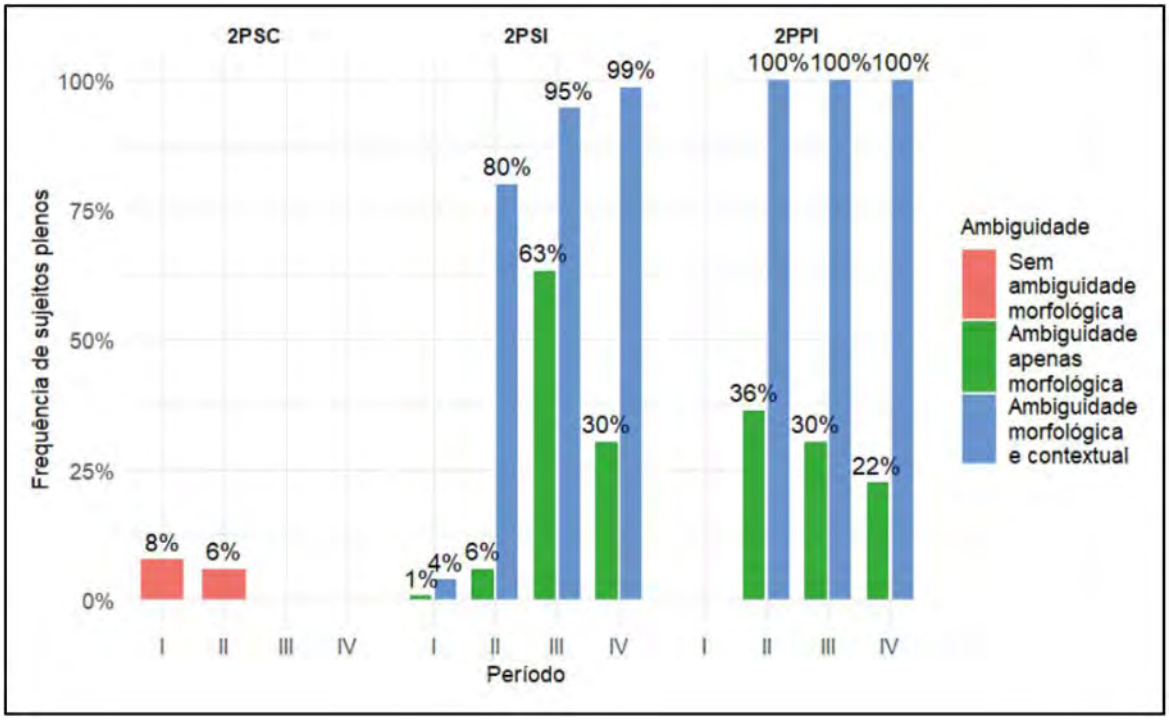
Assim, a referência à 2ª pessoa passa a ser feita exclusivamente por meio das novas formas pronominais, o que faz com que a ocorrência de ambiguidade morfológica aumente muito nos períodos seguintes, e que os pronomes, consequentemente sejam mais expressos:

(16) *PIERRE - É mona, **você** precisa agitar o mais rápido possível esse babado. **Você** não tem idéia da quantidade de festas, coquetéis, casamentos e recepções que **você** anda perdendo...* (Por favor, matem minha empregada!, Edilson Pedrini, 1998)

A ambiguidade morfológica e contextual só não se destaca na expressão de pronomes no período I, mantendo frequências acima de 90% em todos os outros. Quando a ambiguidade é apenas morfológica, a trajetória da mudança acompanha a tendência geral da 2ª pessoa. Como se observa na análise por pronome mostrada no gráfico 5, isso indica uma relação com a 2ª pessoa do singular inovadora, *você*, que também apresenta esse percurso.

[15] Legenda: Período I = décadas de 1880 e 1890; Período II = década de 1930; Período III = décadas de 1970 e 1980; Período IV = décadas de 1990 e 2000.

Gráfico 5: Frequência relativa de sujeitos plenos de 2ª pessoa em tempo real de acordo com a pessoa do discurso e a variável "ambiguidade"[16]



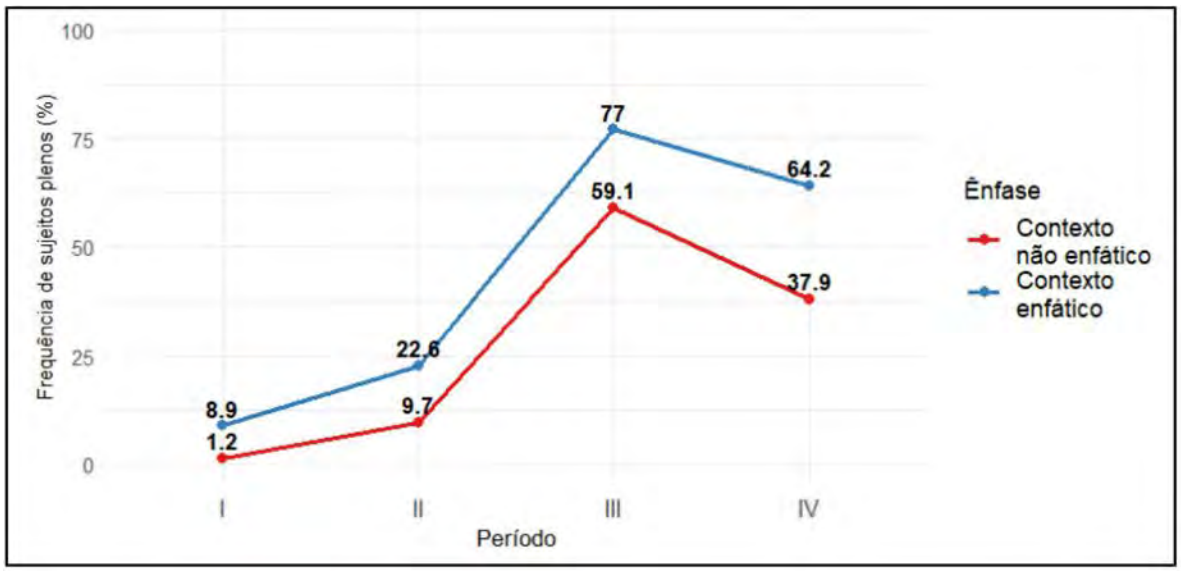
Fonte: Rodran (2025, p. 189)

No geral, a presença de ambiguidade contextual mostra ser grande impulsionadora da mudança, enquanto a ambiguidade apenas morfológica tem uma atuação secundária. Já a ausência de ambiguidade morfológica representa forte resistência à mudança. Além disso, nossos resultados coincidem com os de Paredes Silva (1988), Genuíno (2017) e Massariol (2018) no que se refere à hierarquia entre os fatores.

A variável "ênfase" mostra resultados condizentes com nossa hipótese de que os contextos enfáticos seriam mais favoráveis à implementação da mudança, assim como observado por Paredes Silva (1988) e Genuíno (2017). Como consta no gráfico 6, a ordenação entre os fatores não se altera ao longo das sincronias, e ambos acompanham a tendência geral da evolução da mudança na 2ª pessoa.

[16] Legenda: 2PSC = 2ª pessoa do singular conservadora (tu); 2PSI = 2ª pessoa do singular inovadora (você); 2PPI = 2ª pessoa do plural inovadora (vocês).

Gráfico 6: Frequência relativa de sujeitos plenos de 2ª pessoa (geral) em tempo real de acordo com a variável "ênfase"[17]



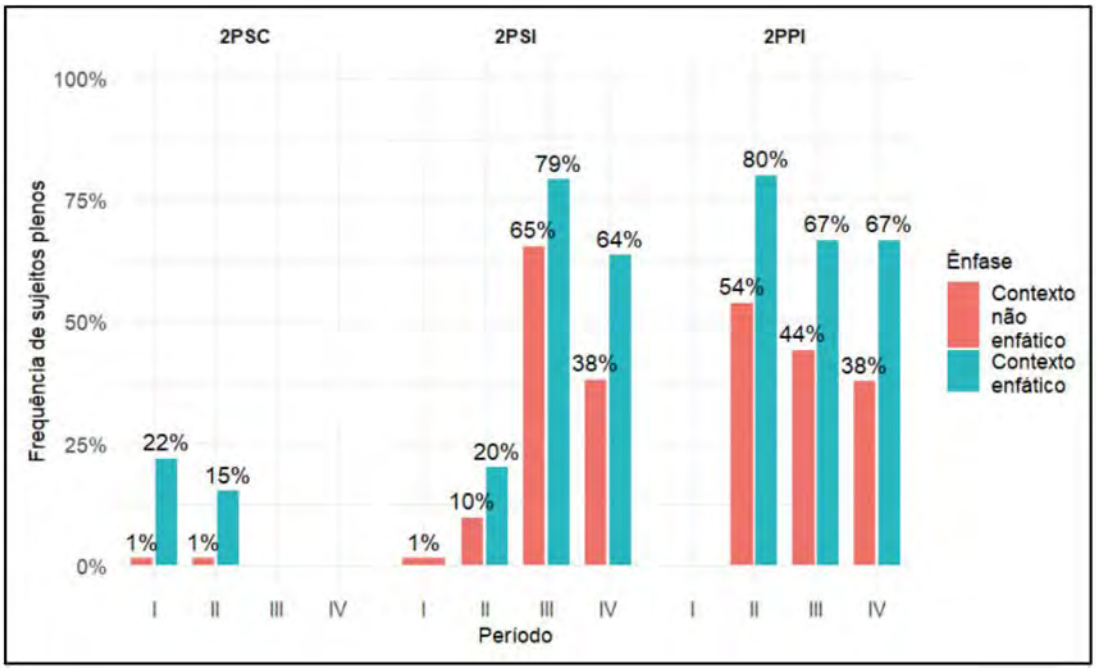
Fonte: Rodran (2025, p. 190)

Conforme o gráfico 7, os poucos dados de pronome *tu* expresso praticamente só ocorrem em contextos enfáticos, "o que indica uma motivação estritamente discursiva, e não gramatical para essas ocorrências" (Rodran, 2025, p. 191). Nos dois dados exemplificados em (17), não há necessidade de expressão do pronome, pois não há comprometimento da compreensão da mensagem. Mesmo assim, no entanto, o pronome *tu* é expresso na segunda ocorrência, com o objetivo de destacar o sujeito, caso que se enquadra nos critérios de Paredes Silva (1988) devido à presença do conectivo contrastivo "entretanto".

(17) *JULIA - E por que não* ***Ø*** *fazes o mesmo? O filho do Dr. Jorge vive quase sempre licenciado, em caçadas e parece mais morar na roça que na cidade; entretanto que* ***tu*** *há um ano não te dispões a passar com tua família nem ao menos 8 dias em nosso sítio.* (Amigo íntimo, Aristides Freire, ~1897)

[17] Legenda: Período I = décadas de 1880 e 1890; Período II = década de 1930; Período III = décadas de 1970 e 1980; Período IV = décadas de 1990 e 2000.

Gráfico 7: Frequência relativa de sujeitos plenos de 2ª pessoa em tempo real de acordo com a pessoa do discurso e com a variável "ênfase"[18]



Fonte: Rodran (2025, p. 191).

A expressão do pronome *você*, por sua vez, parece ser pouco dependente do emprego enfático para acontecer. No período I, há ocorrência de *você* expresso apenas em contextos não enfáticos. Nos períodos II e III, a diferença entre sujeitos plenos enfáticos e não enfáticos fica entre dez e quinze pontos percentuais apenas, sendo que a frequência de não enfáticos atinge 65% na sincronia III. Já a 2ª do plural apresenta distâncias maiores entre os dois fatores, mantendo certa estabilidade entre de 23 a 29 pontos percentuais ao longo das sincronias.

Na variável "tipo sintático de oração", os resultados gerais apontam para a ausência de correferência entre o sujeito da oração analisada em relação à oração anterior dentro do mesmo período como o contexto sintático mais favorável ao desenvolvimento da mudança. Os tipos de oração com maiores frequências de sujeitos plenos, em relação à frequência média de 31,23%, foram as orações relativas sem correferência (54,84%), exemplificada em (18), as subordinadas adverbiais pospostas sem correferência (47,62%), as coordenadas não iniciais sem correferência (46,8%7), as coordenadas iniciais (45,28%), as absolutas (42,57%), as principais pospostas sem correferência (40%), as principais antepostas (38,13%).

(18) *LORENA - Agora estou entendendo...todas aquelas noites que* ***você*** *chegava em casa exausto do trabalho é porque o senhor tava trabalhando DURO!!!* (Por favor, matem minha empregada!, Edilson Pedrini, 1998)

[18] Legenda: 2PSC = 2ª pessoa do singular conservadora (tu); 2PSI = 2ª pessoa do singular inovadora (você); 2PPI = 2ª pessoa do plural inovadora (vocês).

Enquanto isso, representando contextos sintáticos de resistência à mudança estão as orações subordinadas substantivas com correferência (2,94%), exemplificada em (19), as coordenadas não iniciais com correferência (6,25%), as interrogativas (23%) e as relativas com correferência (25%). Os demais tipos de oração mostraram uma influência neutra, com uma diferença de cinco pontos percentuais ou menos em relação à frequência média.

(19) *COMANDANTE - Agora meu sobrinho, lembra-te de que* **Ø** *és soldado do 30 B. C. Esse Batalhão tem uma bandeira tecida das mesmas glórias que imortalizaram Greenhalgh, Pedro Affonso e os 18 heróes de Copacabana. [...]* (O voluntário do 3º B.C., Ernesto Guimarães, 1938)

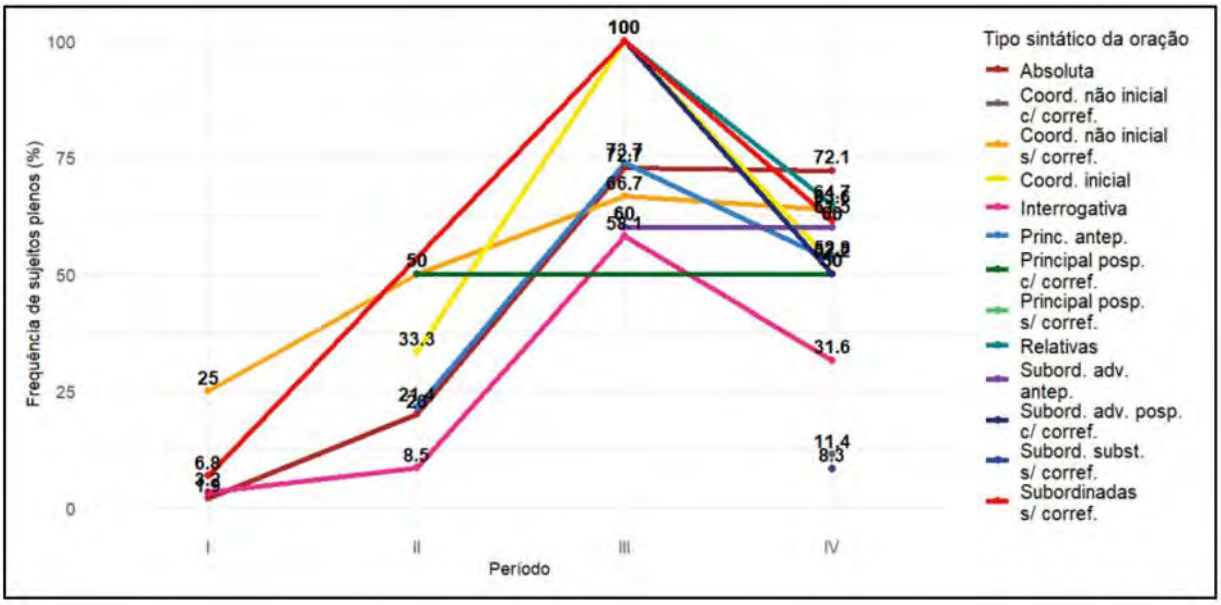
Conforme Duarte (2018 [1993]), especificamente no caso das orações interrogativas, que geralmente são direcionadas a uma 2ª pessoa que pode ser pragmaticamente identificada, os pronomes sujeito tendem a ser menos expressos, mesmo que não haja correferência dentro do período, como consta nos exemplos (20) e (21). Assim como os de Paredes Silva (1988), nossos resultados confirmam essa hipótese. Nas peças teatrais capixabas apenas 23,75% dos sujeitos das interrogativas são plenos, a menor frequência entre as orações sem correferência.

(20) *ANIBAL - (De botas em viagem) Então assim se recebe um amigo ausente há 10 anos?*
PAULO - (Correndo a ele) Anibal, como sou feliz neste momento. Donde **Ø** *vens? Quando* **Ø** *chegaste?* Como tem corrido a sorte? (Amigo íntimo, Aristides Freire, ~1897)

(21) *TELMA - (Para Sandra) Você está em que série?*
SANDRA - Sou universitária...
PEDRO - Essa fila tá muito importante. **Ø** *Vai ser médica, gente fina?* (Fila eterna, Vera Viana, 1978)

Na análise em tempo real, apresentada no gráfico 8, as orações interrogativas parecem ser as mais resistentes, ocupando as posições mais baixas ao longo de todo o período considerado. Apesar disso, seguem a mesma oscilação geral da 2ª pessoa, o que revela conexão com o desenvolvimento da mudança. A maioria dos tipos de oração mostra baixos índices nos períodos I e II e atinge seu ápice no período III - as subordinadas adverbiais antepostas, relativas sem correferência, e coordenadas iniciais aparecem com frequências categóricas de 100% de sujeitos plenos - com queda no período IV. Não seguem esse padrão as absolutas, as subordinadas adverbiais antepostas e as principais pospostas com correferência.

Gráfico 8: Frequência relativa de sujeitos plenos de 2ª pessoa (geral) em tempo real de acordo com a variável "tipo sintático da oração"[19]



Fonte: Rodran (2025, p. 195).

Vale notar o aparecimento de dados de sujeitos plenos de 2ª pessoa em orações subordinadas substantivas com correferência (1 dado) e em coordenadas não iniciais com correferência (5 dados), na última sincronia, contextos apontados por Duarte (1995) como os mais resistentes ao sujeito expresso. Como discutido em Rodran (2025, p. 196), "isso mostra como a mudança na 2ª pessoa avançou nas peças capixabas ao longo das décadas, conquistando um lugar em contextos sintáticos que inicialmente eram resistentes a ela". Os exemplos em (22) e (23) mostram ocorrências desses dois tipos sintáticos de oração.

(22) *BOTELHO - A herança da Mimi pode ser nossa tábua de salvação!*
ALICE - (acalmando-se) Sabe que ***você*** *tem razão? Não tinha pensado nisso.* (A volta da vedete, Jovany Sales Rey, 2000)

(23) *DÉBORA - Você queria que não fosse, mas no fundo, no fundo,* ***você*** *não pode negar, você ainda o ama, e ele não tá nem aí prá você.* (Por favor, matem minha empregada!, Edilson Pedrini, 1998)

Quanto à variável conexão discursiva, os resultados gerais para a 2ª pessoa, apresentados no gráfico 9, indicam que, em relação à frequência média de 31,23%, o sujeito pleno ocorre mais em contextos com graus 6, 7 e 8 de conexão discursiva, com frequências de 44,55%, 41,62% e 41,67%, respectivamente. Os graus 1, exemplificado em (24), e 2 mostraram ser os mais resistentes, com frequências de 14,68% e 20,53%, respectivamente, em consonância com Paredes Silva (1988).

[19] Legenda: Período I = décadas de 1880 e 1890; Período II = década de 1930; Período III = décadas de 1970 e 1980; Período IV = décadas de 1990 e 2000.

(24) *ANTÔNIO - E o Sr. Vigário, por que não faz uma prática ao povo, explicando as coisas?* ***Ø*** *Prestaria um serviço enorme e* ***Ø*** *tranquilizaria o espírito da população. Aquilo que for mais reservado, mande que se entendam comigo.* (A república na roça, Aristides Freire, 1889)

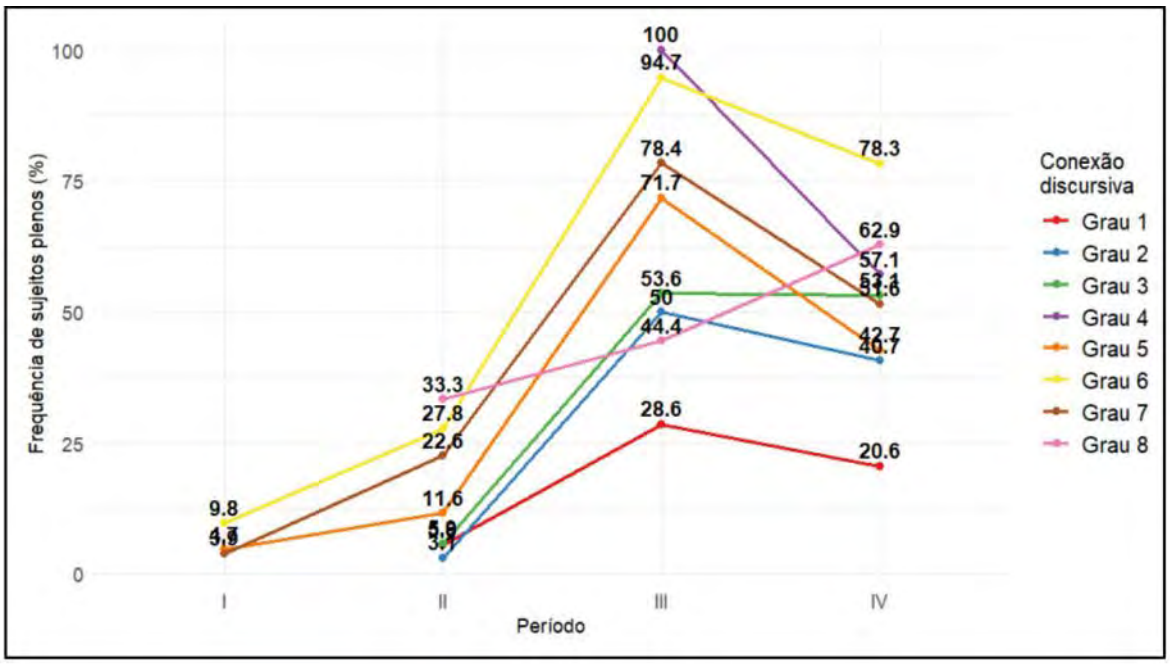
Vemos, portanto, que, textual-discursivamente, a manutenção do referente na fala de um mesmo personagem é inibidora do uso de pronome sujeito. A mudança de turno (grau 3), porém, apresenta influência neutra. Por outro lado, a mudança de função sintática do referente (grau 6), exemplificada em (25), mostra ser o contexto mais favorável à expressão dos pronomes sujeito de 2ª pessoa, seguido do início de cena e primeiras interações entre personagens (grau 8) e da mudança de tópico discursivo (grau 7).

(25) *CATARINA - Ora, deixe-se ficar. É que a senhora, D. Faustina, nessa idade, já perdeu o gosto para viagens... Mas vai gostar daqui... Eu farei tudo para agradá-las...* ***Vocês*** *me vão fazer excelente companhia... Vivo tão só…* (A princesinha, Ernesto Guimarães, 1939)

Na análise em tempo real, conforme o gráfico 9, a hierarquia entre os graus se mantém de maneira geral. Os graus 1, 2 e 3 apenas aparecem a partir do período II e permanecem com as frequências mais baixas a partir daí. No período IV, porém, o grau 3 não apresenta queda, indicando, talvez, uma diminuição da restrição. Os graus 5 (interferência de outros sujeitos), 6 e 7 aparecem desde a primeira sincronia e mostram favorecimento à mudança em relação aos outros graus, principalmente no período III, em que atingem seu ápice. O grau 6, contudo, permanece sendo o mais favorecedor entre os três. Apenas é superado pelo grau 4 no período III, que atinge 100% de frequência, mostrando a relevância da interferência de orações com sujeito impessoal. O grau 8, que julgamos ser o mais propenso à expressão pronominal, na verdade, destoou do percurso da mudança, evoluindo de maneira mais linear.

No período IV, observa-se uma reordenação dos fatores, mas a maioria deles apresenta queda em relação ao período III. No conjunto, a maioria dos graus segue a tendência geral, indicando correlação entre a variável conexão do discurso e a mudança na 2ª pessoa.

Gráfico 9: Frequência relativa de sujeitos plenos de 2ª pessoa (geral) em tempo real de acordo com a variável "conexão discursiva"[20]



Fonte: Rodran (2025, p. 199).

Vejamos agora o que revelam os resultados das variáveis extralinguísticas.

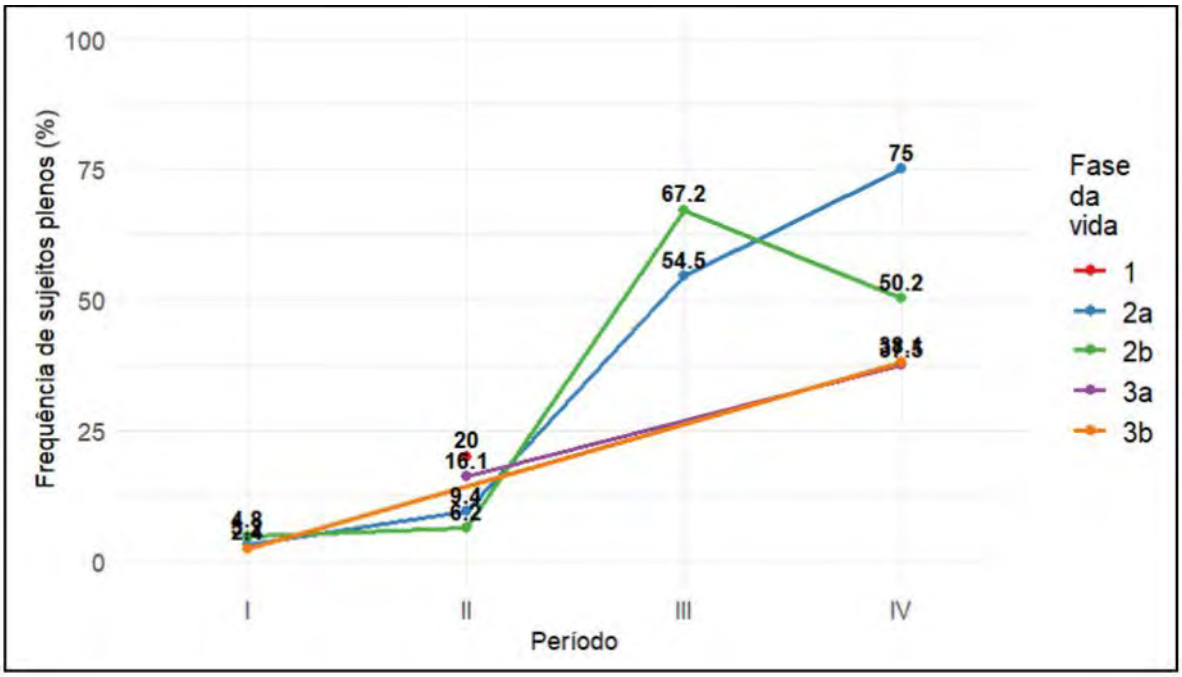
4.2. Resultados das variáveis extralinguísticas

Conforme se observa no gráfico 10, ocorre aumento de sujeitos plenos em todas fases da vida da primeira para a última sincronia. As fases 2a e 2b, contudo, parecem estar mais conectadas com a mudança, uma vez que acompanham as oscilações gerais, principalmente a fase 2b, e mantêm altas frequências a partir do período III. No período IV, a fase 2a é a que eleva a frequência média da sincronia, apresentando aumento em vez de queda. Esses resultados convergem de certa maneira com os de Genuíno (2017), que também observa que a variante inovadora ocorre mais na fala de capixabas jovens adultos, adultos e de meia idade.

As fases 3a e 3b, por sua vez, não ocorrem em todas as sincronias: a 3a apenas ocorre nos períodos I e IV, e a 3b nos períodos I e IV, apresentando, assim, um percurso pouco preciso. Apesar disso, ambas evoluem com aumento de sujeitos plenos chegando ao período IV com diferença de pouco mais de 1% entre suas frequências. Assim, como discutido em Rodran (2025, p. 200), vemos que "a gradação da mudança pelas fases da vida nas peças não diferencia adultos mais jovens e mais velhos, e nem pessoas de meia idade e idosos com idade mais avançada".

[20] Legenda: Período I = décadas de 1880 e 1890; Período II = década de 1930; Período III = décadas de 1970 e 1980; Período IV = décadas de 1990 e 2000.

Gráfico 10: Frequência relativa de sujeitos plenos de 2a pessoa (geral) em tempo real de acordo com a variável "fase da vida"[21]



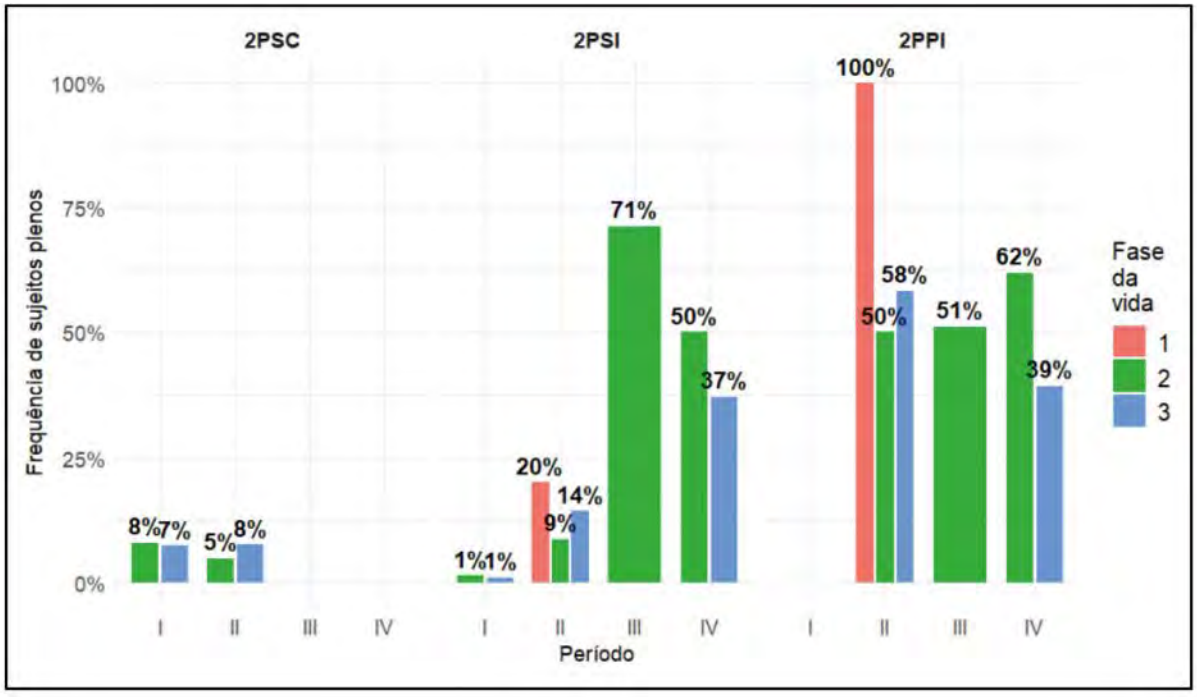
Fonte: Rodran (2025, p. 201)

É interessante observar ainda que personagens da fase 1 somente aparecem nas sincronias II e IV. A presença de dados de sujeito pleno de personagens da fase 1 que ocorre apenas no período II, com a maior frequência da sincronia, pode ser indício de representação da mudança em progresso por meio da fase da vida dos personagens. Essa é uma suposição que a análise por pronome ajuda a esclarecer. Conforme o gráfico 11, quando aparecem, as personagens da fase 1 são as que mais expressam as formas pronominais inovadoras: 20% de *você* e 100% de *vocês* no período II, o que não ocorre com o pronome *tu*.

As formas inovadoras tanto do singular quanto do plural seguem praticamente as mesmas tendências: no período II são mais expressas por personagens da fase 1, seguida da fase 3, e depois da 2; no período III a fase 2 é a única e apresenta altas frequências; no período IV a fase 2 continua predominando, mas a fase 3 também mostra taxas consideráveis de sujeitos plenos. Já o pronome *tu* é muito pouco expresso por ambas as fases 2 e 3.

[21] Fase 1: criança/jovem (0-15 anos); Fase 2a: jovem/adulto (16-25 anos); Fase 2b: adulto (26-40); Fase 3a: meia idade (41-55 anos); Fase 3b: idoso (56 anos ou mais)

Gráfico 11: Frequência relativa de sujeitos plenos de 2ª pessoa em tempo real de acordo com a pessoa do discurso e com a variável "fase da vida"

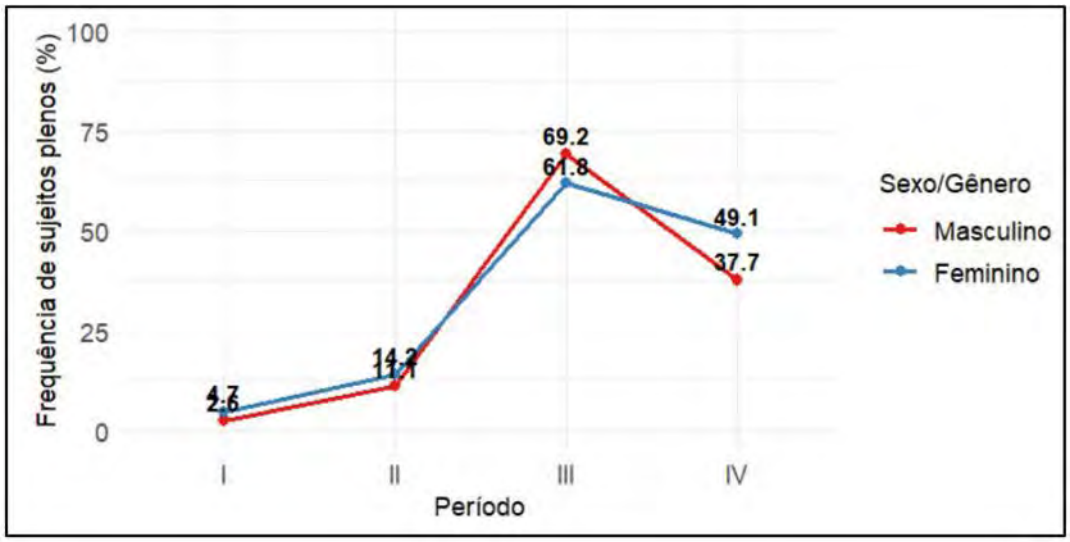


Fonte: Rodran (2025, p. 202)

Quanto à variável sexo/gênero, os resultados gerais indicam mais expressão de sujeitos pronominais nas falas das personagens mulheres, mas com pouca diferença para a frequência dos homens: 34,1% contra 27,46%, respectivamente. Esses resultados estão de acordo com os de Duarte (1995) e Genuíno (2017). Entretanto, a análise em tempo real, apresentada no gráfico 12, mostra que não há proeminência de nenhum dos fatores no desenvolvimento da mudança, o que pode explicar o fato de a variável ter sido a menos significativa no teste de Pearson.

Nos períodos I e II, o fator "feminino" mantém-se mais alto, mas a diferença entre os fatores é estável e muito baixa: 2% e 3% em cada sincronia. No período III, os personagens homens assumem a liderança, mas com frequência apenas 7,4% maior que as mulheres. No período IV, as posições se invertem novamente, com diferença de 11,4% entre homens e mulheres. Apesar do baixo contraste, a trajetória de ambos os fatores coincide com o percurso geral da 2ª pessoa, mostrando conexão com a mudança.

Gráfico 12: Frequência relativa de sujeitos plenos de 2ª pessoa (geral) em tempo real de acordo com a variável "sexo/gênero"



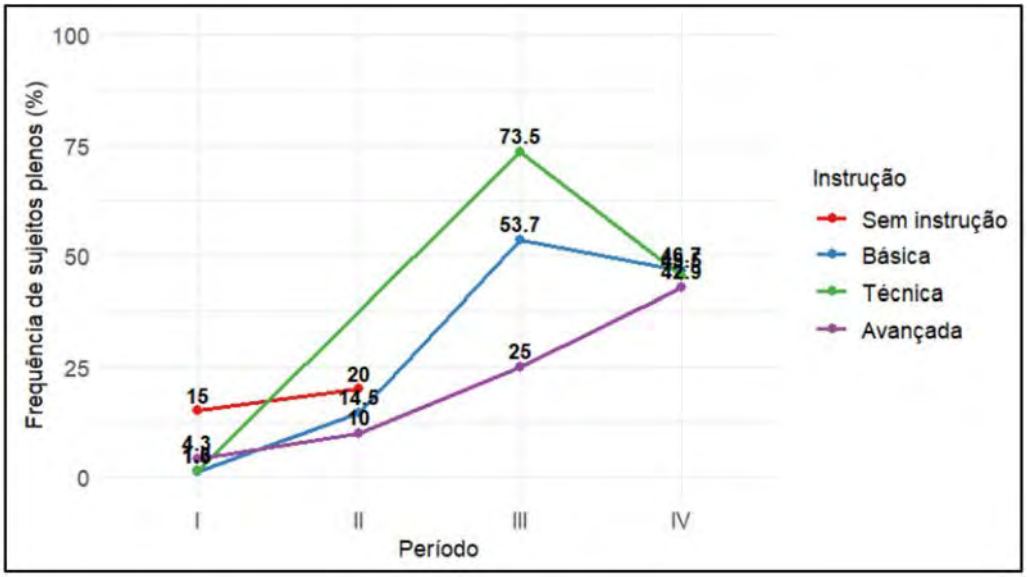
Fonte: Rodran (2025, p. 203)

Dessa maneira, não consideramos seguro afirmar que houve uma utilização das variantes ausência e presença de pronome sujeito na caracterização de sexo/gênero dos personagens no que se refere à 2ª pessoa do discurso.

Quanto à variável "instrução", há, nas sincronias I e II, conforme o gráfico 13, uma liderança dos personagens sem instrução, que estão presentes somente nas peças dessas duas primeiras sincronias. Os demais graus de instrução começam a sobressair-se no período III, em que se destacam os personagens de grau técnico, seguido pelo básico, e depois pelo avançado. Nesse contexto, o trajeto do nível básico de instrução é o mais próximo do percorrido pela 2ª pessoa no geral, caminho também muito parecido com o do grau técnico. No período IV, o grau avançado mantém a tendência de aumento, equiparando sua frequência às dos outros graus de instrução.

Vemos, portanto, que todos os graus de instrução sofreram mudança na 2ª pessoa, mas parece haver uma conexão maior dos níveis intermediários de instrução com esse processo. Entretanto, o nível avançado mostra uma trajetória linear de crescimento, convergindo com os outros graus na última sincronia, o que, como discutido em Rodran (2025, p. 205), "pode ser sinal da perda de uma possível avaliação negativa, ainda que pouco marcada, da variante expressa".

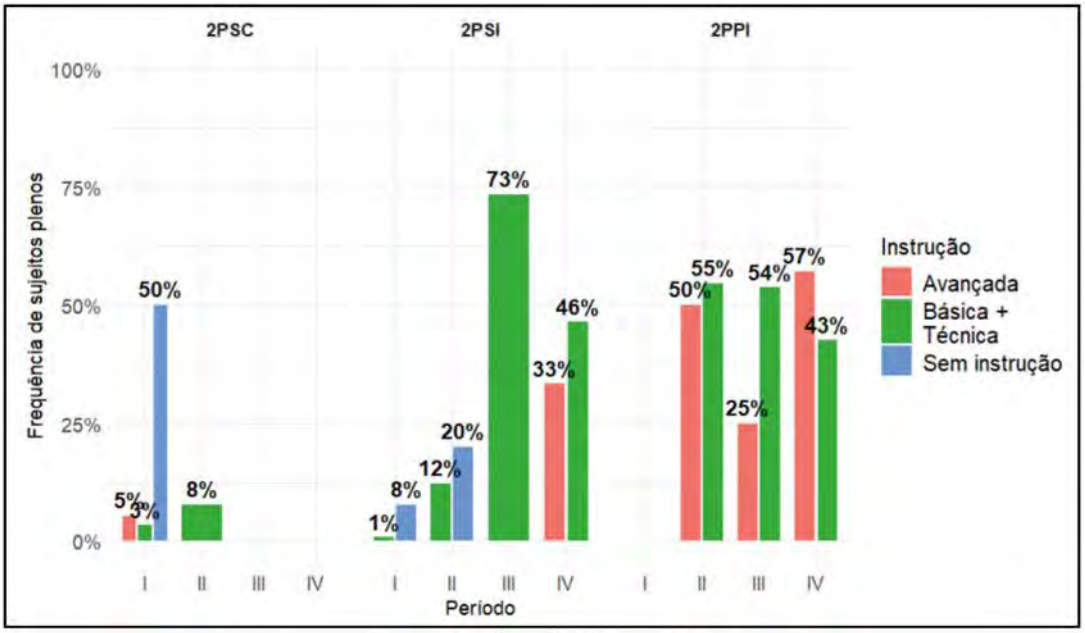
Gráfico 13: Frequência relativa de sujeitos plenos de 2ª pessoa (geral) em tempo real de acordo com a variável "instrução"



Fonte: Rodran (2025, p. 206)

A distribuição das frequências no tempo real por pronome confirma uma clara proeminência dos personagens sem instrução nos períodos I e II, principalmente para o pronome *tu* e exceto para o pronome *vocês*. A amálgama entre os níveis básico e técnico se mantém à frente da mudança de maneira geral, mas é superada no período quadro na expressão do plural, o que pode ser indício da falta de marcação social da variante expressa para o pronome *vocês*.

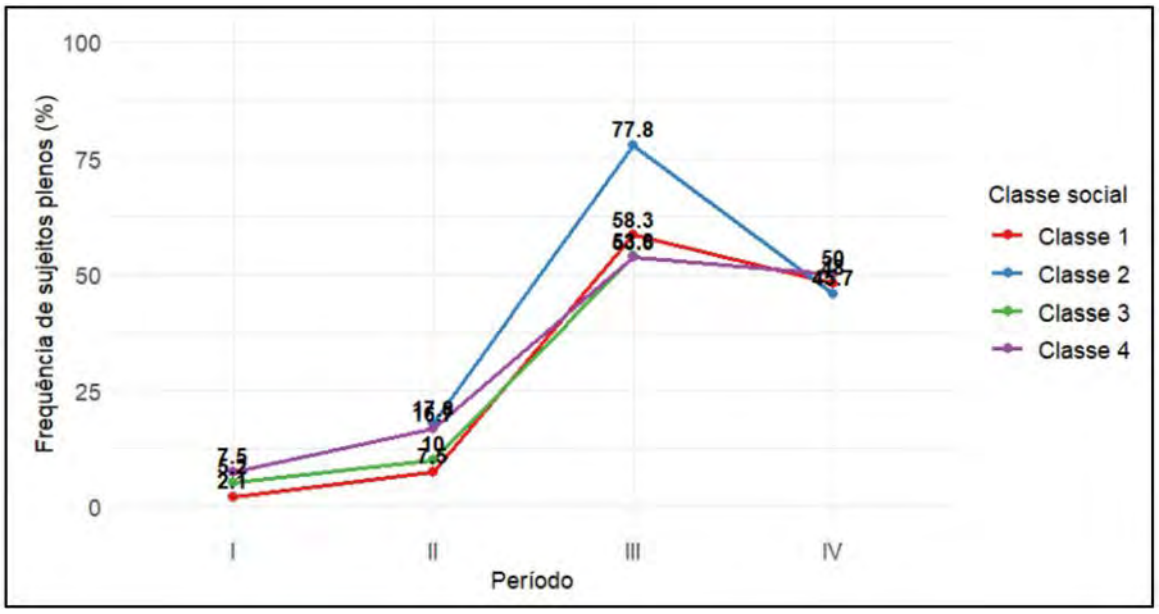
Gráfico 14: Frequência relativa de sujeitos plenos de 2ª pessoa em tempo real de acordo com a pessoa do discurso e com a variável "instrução"



Fonte: Rodran (2025, p. 206)

Por fim, os resultados da variável "classe social" mostram que há uma correlação entre cada uma das quatro classes sociais e a mudança, uma vez que percorrem trajetórias quase iguais entre si e em relação ao geral da 2ª pessoa, conforme o gráfico 15. A ordenação entre os fatores, contudo, apontam para uma liderança inicial dos personagens das classes 4 e 3 nos períodos I e II. No período III, a classe 2 toma o primeiro lugar, sendo seguida ainda pela classe 1, enquanto na última sincronia, há uma convergência entre todas as classes.

Gráfico 15: Frequência relativa de sujeitos plenos de 2ª pessoa (geral) em tempo real de acordo com a variável "classe social"[22]



Fonte: Rodran (2025, p. 208)

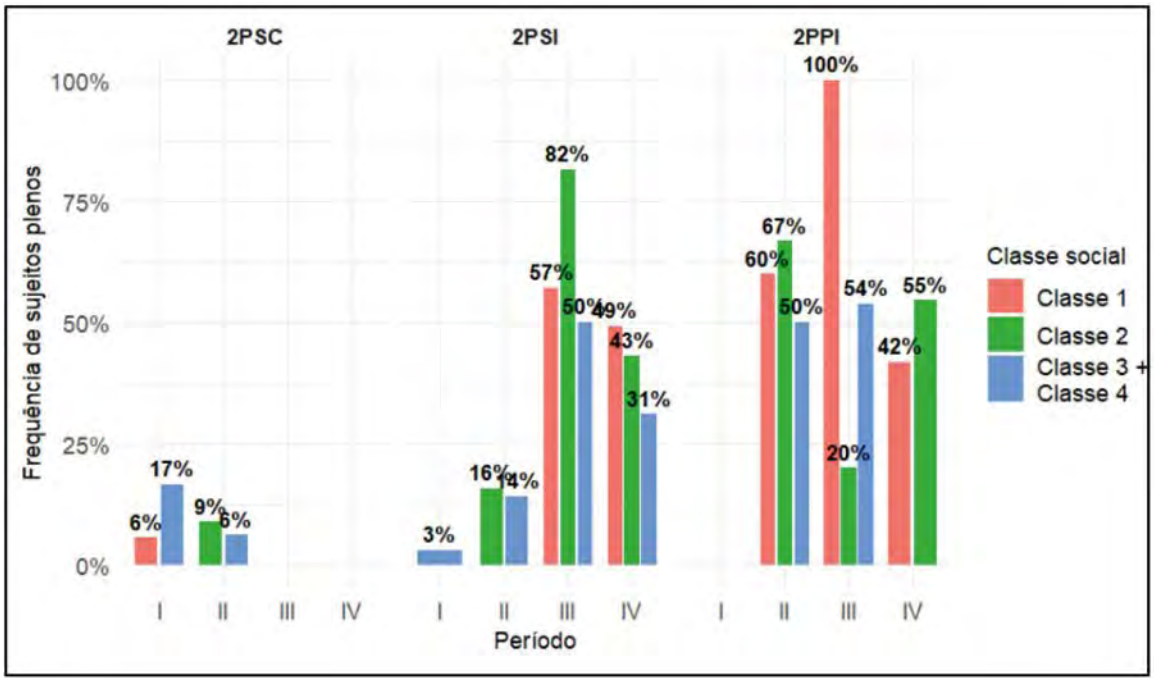
É confirmada, portanto, a hipótese de que a mudança na 2ª pessoa seria representada nas peças mais antigas principalmente por meio dos personagens das classes subalternas. No curso do tempo, a proeminência adquirida pelas classes mais altas no período III, conforme o gráfico 15, mostra o avanço da mudança por meio da difusão social da variante plena. Nessa perspectiva, a convergência de todas as classes no período IV pode ser indício de perda de estigmatização, assim como vimos na análise do grau de instrução.

Os resultados por desinência mostram que a expressão do pronome *tu* ocorre principalmente nas falas dos personagens das classes 3 e 4. O pronome *você*, por sua vez, inicialmente é expresso apenas pelas classes 3 e 4, depois entra em equilíbrio com a classe 2, no período II, e no período III, a classe 2 já está totalmente à frente do processo. Também é notável no período III a presença da classe 1, já

[22] Conforme Brandão (2022): Classe 1: alta sociedade/ alta burguesia/ dominantes; Classe 2: classe emergente/ intermediárias e/ou "classes médias" - junção de duas "classes" - média e pequena burguesia; Classe 3: classe emergente/ intermediárias e/ou "classes médias" - junção de duas "classes" - média e pequena burguesia; Classe 4: oprimidos/ marginalizados.

superando as frequências das classes mais baixas. No período IV, a classe 1 é que está à frente, seguida pela classe 2, e depois pela amálgama das classes 3 e 4. Como discute Rodran (2025, p. 209), "mais uma vez, essas constatações podem estar evidenciando a perda gradativa da provável marcação social comentada anteriormente, mais forte principalmente nas peças do século XIX."

Gráfico 16: Frequência relativa de sujeitos plenos de 2ª pessoa em tempo real de acordo com a pessoa do discurso e com a variável "classe social"



Fonte: Rodran (2025, p. 209)

Já o pronome *vocês* não apresenta um padrão na distribuição pelas classes sociais. O que se observa são frequências altas em todas as classes e sincronias, com destaque para os 100% de sujeitos plenos na classe 1 no período III. Isso pode ser indício de que, por estar há mais tempo em amplo uso, esse pronome não guarda significado social que fosse percebido pelos autores das peças a ponto de poder ser utilizado para caracterizar os personagens.

Como já comentado, os resultados referentes à classe social são muito semelhantes aos da variável "instrução". Ambos revelam as falas de personagens menos instruídos e de classes mais baixas como *locus* inicial de representação da mudança nas peças, com uma posterior difusão para classes e níveis de instrução intermediários, atingindo, por fim, classes mais altas e com instrução avançada. À medida que a escolarização pode significar ascensão social, conforme comenta Freitag (2016), essas semelhanças podem indicar uma correlação entre a classe social e o nível de instrução, confirmando nossa hipótese de que tratamos de uma mudança de baixo (Labov, 1994).

5. Considerações finais

Este estudo teve como objetivo investigar a variação e a mudança na expressão pronominal da 2ª pessoa do discurso em peças teatrais capixabas, a partir de uma análise em tempo real, contemplando tanto variáveis linguísticas quanto extralinguísticas. Os resultados obtidos permitem confirmar a hipótese central de que a implementação do sujeito pronominal expresso na 2ª pessoa não se dá de forma aleatória, mas é fortemente condicionada por fatores estruturais, discursivos e sociais, refletindo um processo sistemático de mudança linguística ao longo dos períodos analisados.

No plano linguístico, os dados evidenciam a substituição progressiva do pronome *tu* pelo pronome *você* como forma predominante de referência à 2ª pessoa do singular, com um ponto crítico de inflexão entre as sincronias II e III. Esse intervalo, correspondente ao período entre o final da década de 1930 e as décadas de 1970 e 1980, mostrou-se decisivo para o avanço da mudança, culminando na exclusividade do paradigma de 3ª pessoa para a referência à 2ª pessoa do singular nas peças mais recentes. A 2ª pessoa do plural, por sua vez, já se apresenta estabilizada desde as sincronias mais antigas, com predominância categórica de *vocês*, o que confirma que a mudança no plural antecede a do singular.

Entre as variáveis linguísticas analisadas, destacam-se a ambiguidade, a ênfase, o tipo sintático de oração e a conexão discursiva como fatores decisivos para a expressão do sujeito pronominal. A presença de ambiguidade contextual mostrou-se o principal motor da mudança, enquanto a ausência de ambiguidade morfológica funcionou como forte elemento de resistência. De modo convergente com estudos anteriores, os contextos enfáticos favoreceram a expressão pronominal, sobretudo nos poucos casos remanescentes de *tu* expresso, cuja motivação se revelou predominantemente discursiva. No âmbito sintático, orações sem correferência e com menor continuidade referencial favoreceram a expressão do sujeito, ao passo que estruturas tradicionalmente resistentes, como subordinadas substantivas com correferência e interrogativas, mantiveram índices mais baixos, ainda que tenham passado a admitir sujeitos plenos nas sincronias mais recentes, sinalizando o avanço da mudança para contextos antes interditados.

No que se refere às variáveis extralinguísticas, os resultados indicam que a mudança é sensível às fases da vida, ao grau de instrução e à classe social dos personagens, ainda que essas categorias devam ser interpretadas com cautela, dada a natureza ficcional do *corpus*. Observou-se que personagens jovens e adultos concentram as maiores frequências de formas inovadoras, sugerindo a representação da mudança em progresso. Do ponto de vista social, as classes mais baixas e os personagens com menor instrução figuram como *locus* inicial da inovação nas peças mais antigas, com posterior difusão para níveis intermediários e, finalmente, para as classes mais altas e personagens com instrução avançada. Esse percurso é compatível com a noção de mudança "de baixo", conforme proposta por Labov, marcada inicialmente por possível avaliação negativa e, posteriormente, por perda de estigmatização e generalização do uso.

A variável sexo/gênero, por sua vez, mostrou-se menos decisiva para o desenvolvimento da mudança, uma vez que as diferenças entre personagens masculinos e femininos foram reduzidas e instáveis ao longo das sincronias, não permitindo afirmar que a expressão do sujeito pronominal tenha sido empregada de forma sistemática como recurso de caracterização social nesse eixo.

De maneira geral, os resultados corroboram a adequação das peças teatrais como fonte para o estudo histórico da variação linguística, na medida em que refletem, ainda que mediadas pela escrita e pela construção autoral, padrões sistemáticos de uso compatíveis com os observados em dados de oralidade. Ao evidenciar o avanço gradual e condicionado da expressão pronominal da 2ª pessoa no português brasileiro, este trabalho contribui para a compreensão dos mecanismos de mudança sintática e discursiva da língua, reforçando a importância de análises que articulem fatores linguísticos e sociais em perspectiva diacrônica. Estudos futuros podem aprofundar essa investigação a partir da comparação com outros gêneros textuais e variedades regionais, bem como explorar de forma mais detalhada a interface entre representação social e escolhas linguísticas na escrita dramatúrgica.

Referências

BECHARA, Evanildo. *Moderna gramática portuguesa*. 37. ed. rev., ampl. e atual. Rio de Janeiro: Nova Fronteira, 2019 [1999].

BERLINCK, Rosane de Andrade; BRANDÃO, Silvia Maria. Por uma sociolinguística histórica: análise multidimensional de cartas pessoais e peças teatrais brasileiras. *In*: BIAZOLLI, Caroline Carnielli; BERLINCK, Rosane de Andrade (org.). *Gêneros textuais-discursivos no estudo de processos de variação e mudança*. 1. ed. Campinas, SP: Pontes Editores, 2021.

BERLINCK, Rosane de Andrade; BRANDÃO, Sílvia Maria; BARBOSA, Juliana Bertucci; BIAZOLLI, Caroline Carnielli; SENE, Marcus Garcia. Projeto Língua EmCena: peças teatrais como fonte para o estudo sócio-histórico da língua. *LaborHistórico: Diálogos entre a Sócio-História do Português e a História Social da Cultura Escrita*. v. 9, n. 2, 2023.

BRANDÃO, Sílvia Maria. *Mudança do quadro verbal brasileiro em sentenças condicionais: contribuições à Sociolinguística Histórica*. 2022. 323 f. Tese (Doutorado em Linguística e Língua Portuguesa) — Faculdade de Ciências e Letras, Universidade Estadual Paulista "Júlio de Mesquita Filho", Araraquara, 2022.

CASTILHO, Ataliba Teixeira de. *Nova Gramática do Português Brasileiro*. São Paulo: Editora Contexto, 2010.

CONDE SILVESTRE, Juan Camilo. *Sociolinguística Histórica*. Madrid: Gredos. 2007.

CUNHA, Celso; CINTRA, Lindley. *Nova Gramática do Português*. Contemporâneo. Rio de Janeiro: Nova Fronteira, 7. ed. 2017 [1985].

DUARTE, Maria Eugênia Lamoglia. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. *In*: ROBERTS, Ian; KATO, Mary A. (org.). *Português brasileiro*: uma viagem diacrônica – Homenagem a Fernando Tarallo. Campinas: Editora da UNICAMP. p. 107-128. 1993.

DUARTE, Maria Eugênia Lammoglia. *A perda do princípio "evite pronome" no português brasileiro*. 1995. 151f. Tese (doutorado) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, SP. Disponível em: https://hdl.handle.net/20.500.12733/1583003. Acesso em: 25 abr. 2024.

DUARTE, Maria Eugênia Lammoglia. (org.). *O sujeito em peças de teatro (1983-1992):* estudos diacrônicos. São Paulo: Parábola, 2012.

DUARTE, Maria Eugênia Lammoglia. Do pronome nulo ao pronome pleno: a trajetória no português do Brasil. *In*: ROBERTS, Ian; KATO, Mary A. (org.). *Português brasileiro*: uma viagem diacrônica. São Paulo: Contexto, 2018 [1993].

FREITAG, Raquel Meister Ko. Sociolinguística no/do Brasil. *Cadernos de Estudos linguísticos*. Campinas, pp. 445-460 - set./dez. 2016.

GENUÍNO, Wladimir Ricardi. *A expressão do sujeito pronominal no sujeito falado em Vitória/ES*. Dissertação (Mestrado em Linguística), Vitória, 2017. Disponível em: https://sappg.ufes.br/tese_drupal//tese_11342_Disserta%E7%E3o%20P%F3s-Qualifica%E7%E3o-Wladimir%20Genuino-FINAL.pdf. Acesso em: 11 abr. 2024.

LABOV, William. *Padrões sociolinguísticos*. São Paulo, Parábola, 2008 [1972]. 389p.

LABOV, William. *Principles of linguistic change* (Volume 1: Internal Factors). Massachusetts: Blackwell. 1994.

LABOV, William. *Principles of linguistic change* (Volume 2: Social Factors). Massachusetts: Blackwell. 2001.

LOPES, Célia Regina dos Santos. Retratos da variação entre "você" e "tu" no português do Brasil: sincronia e diacronia. *In*: RONCARATI, Cláudia; ABRAÇADO, Jussara. (org.). *Português Brasileiro II* - contato linguístico, heterogeneidade e história. 1. ed. Niterói: EDUFF, 2008, v. 2, pp. 55-71.

MASSARIOL, Caroliny Batista. *A expressão do sujeito pronominal em cartas e postais capixabas do século XX*. 2018. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Federal do Espírito Santo, Vitória, 2018. Disponível em: https://sappg.ufes.br/tese_drupal//tese_11882_disserta%E7%E3o%20caroliny.pdf. Acesso: 11 abr. 2024.

MENON, Odete Pereira da Silva. O sistema pronominal do português do Brasil. *Letras*, Curitiba, n. 44, pp. 91-106, 1995.

OUSHIRO, Lívia. *Introdução à Estatística para Linguistas*. Campinas, SP : Editora da Abralin, 2022.

PAIVA, Maria da Conceição de; SCHERRE, Maria Marta Pereira. Retrospectiva sociolinguística: contribuições do PEUL. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, São Paulo, v. 15, n. especial, pp. 201-232, 1999.

PAREDES SILVA, Vera Lúcia. *Cartas cariocas*: a variação do sujeito na escrita informal. 1988. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1988.

PAREDES SILVA, Vera Lúcia. Motivações funcionais no uso do sujeito pronominal: uma análise em tempo real. *In*: PAIVA, Maria Conceição de; DUARTE, Maria Eugênia Lamoglia (org.). *Mudança linguística em tempo real*. Rio de Janeiro: FAPERJ/Contra Capa, 2003. pp. 97-114.

ROCHA LIMA, Carlos Henrique. *Gramática normativa da língua portuguesa*. Rio de Janeiro: José Olympio, 2011 [1972].

RODRAN, Iesa Venturin Mulinari de. *A expressão do sujeito pronominal em peças de teatro capixabas: uma análise em tempo real.* 2025. Dissertação (Mestrado em Estudos Linguísticos) – Programa de Pós-Graduação em Estudos Linguísticos, Universidade Federal do Espírito Santo, Vitória, 2025. Disponível em: https://repositorio.ufes.br/items/1296f69e-5a4e-400e-9c84-2841c8e555d9. Acesso em: 26 abr. 2026.

ROMAINE, Suzanne. *Socio-historical linguistics:* its status and methodology. Cambridge: Cambridge University Press, 2009. (Original publicado em 1982).

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX. *In*: ROBERTS, Ian; KATO, Mary A. (org.). *Português brasileiro: uma viagem diacrônica*. São Paulo: Contexto, 2018. pp. 55-82. (Original publicado em 1993)

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução de Marcos Bagno. São Paulo: Parábola, 2006 (Original de 1968).

ZILLES, Ana Maria Stahl. O que a fala e a escrita nos dizem sobre a avaliação social do uso de a gente? *Letras de Hoje*. Porto Alegre, v. 42, n. 2, pp. 27-44, junho, 2007.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72254

A alternância *tu/você* na escrita piauiense novecentista: os condicionamentos linguísticos e sociais

The alternation between tu/você in 20th-century Piauí writing: linguistic and social conditioning

Luiz Egito de Souza Barros[1]

Márcia Cristina de Brito Rumeu[2]

RESUMO

Este estudo investiga a variação entre as formas de referência ao sujeito de 2ª pessoa do singular (*tu/você*) com base na produção escrita de piauienses ilustres e não ilustres do século XX, mais especificamente entre as décadas de 20 e 90 (1920 e 1990), à luz dos princípios da Sociolinguística Histórica (Hernández-Campoy; Schilling, 2012), da Teoria da Variação (Labov, 1994) e do processo de submissão dos dados históricos ao RBrul (RStudio). A hipótese principal é a de que o *você* já se deixaria entrever como mais produtivo do que o *tu* na era novecentista do PB, como já observado por Scherre *et al.* (2021) a partir de Deus (2009) e Cardoso (2008) para os dados do ALiB. De um modo geral, constatamos a prevalência do *você* (60,03%), ainda que em intensa alternância com o *tu* (39,97%), nas cartas piauienses. Constatamos que a forma *você* é favorecida pelas variáveis linguísticas *pessoa verbal*, *paralelismo formal e semântico*, *expressão do sujeito* e pelas variáveis sociais *faixa etária* e *escolaridade*, acompanhando os resultados de Souza (2021) e de Souza (2012) cujas análises também estão embasadas em amostras dos séculos XIX e XX no âmbito da sociolinguística histórica.
PALAVRAS-CHAVE: Variação *tu/você*. Cartas piauienses novecentistas. Sociolinguística Histórica.

ABSTRACT

This study investigates the variation in forms of reference to the second-person singular subject (*tu/você*) based on the written production of illustrious and not illustrious people from Piauí in the 20th century, more specifically between the 1920s and 1990s, in light of the principles of historical sociolinguistics (Hernández-Campoy; Schilling, 2012), Variation Theory (Labov, 1994), and the process of submitting historical data to RBrul (RStudio). The main hypothesis is that *você* was already emerging as more productive than *tu* in the 20th century era of Brazilian Portuguese, as observed by Scherre *et al.* (2021) in light of Deus (2009) and Cardoso (2008) for the ALiB data. In general, we observed the prevalence of *você* (60.03%), although with intense alternation with *tu* (39.97%) in letters from Piauí. We found that the form *você* is favored by the linguistic variables verbal person, formal and semantic parallelism, subject expression, and by the social variables age group and schooling, following the results of Souza (2021) and Souza (2012), also conducted through the analysis of historical samples within the scope of historical sociolinguistics.
KEYWORDS: Variation *tu/você*. Letters from Piauí from the 20th century. Historical sociolinguistics.

[1] Universidade Federal do Piauí (UFPI), luzegi@yahoo.com.br, https://orcid.org/0000-0003-0656-6817.

[2] Universidade Federal de Minas Gerais (UFMG), mrumeu@ufmg.br, https://orcid.org/0000-0001-9254-976X.

Recebido em 20/01/2026 - Aceito em 02/07/2026

Considerações iniciais

A forma *você* se consolidou no quadro pronominal do Português Brasileiro (doravante PB) em contexto nominativo (Souza, 2021; Rumeu, 2013; Souza, 2012; Lopes; Cavalcante, 2011). Assim sendo, voltamo-nos, neste estudo, à alternância entre as formas *tu* e *você* na produção escrita piauiense novecentista. A hipótese principal é a de que o *você* já se deixaria entrever como uma forma pronominal mais produtiva do que o *tu*, na era novecentista do PB, como constatado por Scherre *et al.* (2021), à luz de Deus (2009) e de Cardoso (2008), para os dados de fala registrados nos inquéritos do Atlas Linguístico do Brasil (ALiB). Assume-se, neste trabalho, o presente como ponto de partida para uma incursão pelo português escrito piauiense na sua expressão novecentista. Ilustramos, em (1) e (2), evidências das formas *tu* e *você*, na fala de Teresina, discutidas por Deus (2009).

(1) "Ói *tu* quiser ir embora, *tu* pega, *tu* vai na pirua, ô então vai no alternativo, no alternativo. [...]"

(Teresina/01, *cf.* Deus, 2009, p.108.)

(2) "Fulano, para onde *você* está indo [...] ô vai?"

(Teresina/06, *cf.* Deus, 2009, p.119.)

Considerando o diversificado panorama dos [3]subsistemas tratamentais do PB falado (Scherre *et al.*, 2015 [2009]) e a variação entre as formas *tu/você* na fala piauiense (Scherre *et al.*, 2021, p. 176) a partir dos dados do ALiB (92% *você* versus 8% *tu*), concentramo-nos, neste estudo, na dinâmica *tu/você* da escrita piauiense novecentista, tendo em vista as seguintes questões:

(1ª) As formas pronominais *tu* e *você* também se mostrariam profícuas na escrita novecentista dos piauienses, considerando que o *você* prevalece na fala atual do Piauí, ainda que em alternância com o *tu* (com concordância média entre 10% e 39%), *cf.* Scherre *et al.* (2021, p. 165)?

(2ª) Quais seriam os contextos linguísticos e extralinguísticos que promoveriam o *você* na produção novecentista piauiense?

Para responder essas questões, prevemos as seguintes hipóteses: (1ª) Assumimos a hipótese de que o *você* já se deixaria entrever como mais produtivo do que o *tu* na era novecentista do PB, *cf.* já levantado por Scherre *et al.* (2021) em relação aos dados da fala piauiense; (2ª) Tendo em vista o fato de o *você* ter sido impulsionado por *formas verbais* de 3SG, pelo *você-sujeito*, *clíticos* e *possessivos* de 3SG, por *formas de você-não-sujeito* em estruturas precedentes (variável *paralelismo formal e semântico*) e pela *faixa etária* do informante nas cartas mineiras dos séculos XIX e XX, conforme a

[3] Trata-se do subsistema de *Você* (I), o subsistema de *Tu* com baixo nível de concordância (II), o subsistema de *Tu* com alto nível de concordância (III), o subsistema *Tu/Você* com concordância baixa (IV), o subsistema *Tu/Você* com concordância média (V), o subsistema *Você/Tu* sem concordância (VI), distribuídos pelas regiões brasileiras (Scherre *et al.* (2021 [2009]). Partimos, pois, desses seis subsistemas tratamentais que foram, por sua vez, sintetizados por Lopes e Cavalcante (2011) em três subsistemas (o subsistema do *você*, o subsistema do *tu* e o subsistema da alternância *tu/você*).

discussão de Souza (2021), conjecturamos que o *você*, na escrita piauiense, também seria conduzido por alguns desses fatores internos (linguísticos) e externos (sociais) ao sistema linguístico.

Trazemos como ponto de partida o "presente para entender o passado", buscando "o melhor uso possível dos dados linguísticos" (Labov, 1994) que sobreviveram à ação do tempo nos âmbitos dos arquivos públicos e privados, tendo sempre em vista as *cartas pessoais* confeccionadas por punhos legitimamente piauienses no decorrer da era novecentista do PB. Em termos teórico-metodológicos, temos um estudo orientado pelos parâmetros da Sociolinguística Histórica (Hernández-Campoy; Conde Silvestre, 2012) não só em relação aos parâmetros de composição de amostras históricas, mas também no que se refere à análise de uma regra variável (*tu/você*). Nesta análise, voltamo-nos à produção escrita de redatores piauienses, com distintos níveis de escolarização (alfabetizados, nível médio e nível superior), entre as décadas de 20 e 90 do século XX. Para o tratamento quantitativo dos dados linguísticos, utilizamo-nos das potencialidades do RStudio (interface gráfica do RBrul) para a geração dos índices probabilísticos em relação à regra variável *tu/você*, assumindo o *você* como valor de aplicação (*application value*). Neste artigo, mantemos a anonimidade dos informantes através das referências aos seus nomes através de siglas.

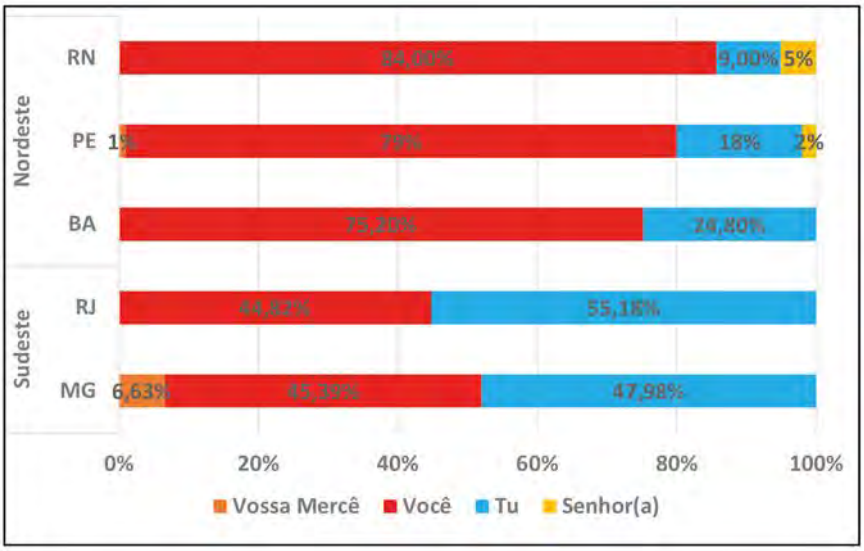
Este texto está estruturado em 5 seções. Nas considerações iniciais, apresentamos o objetivo geral, a hipótese principal e as questões que conduziram esta análise no âmbito da Sociolinguística Histórica. Na 1ª seção, temos uma brevíssima revisão do tema, passando pelos resultados de análises sobre a dinâmica *tu/você* em amostras de cartas pessoais oitocentistas e novecentistas representativas das regiões sudeste e nordeste do Brasil. Na 2ª seção, apresentamos os encaminhamentos teórico-metodológicos desta análise em relação aos parâmetros da Sociolinguística Histórica e às especificidades das amostras históricas. Na 3ª seção, descrevemos as amostras de cartas pessoais piauienses a partir de um quadro geral dos acervos representativos da produção escrita de famílias piauienses. Na 4ª seção, expomos os resultados gerais da alternância *tu/você* no eixo do tempo (1920-1990). Na 5ª seção, trazemos à cena os condicionamentos da regra variável *tu/você* a partir do *você* como valor de aplicação. Nas considerações finais, trazemos breves sistematizações em relação aos contextos linguísticos e extralinguísticos que condicionaram a produtividade da forma *você* na escrita piauiense novecentista, tendo em vista a regra variável *tu/você*.

1. A dinâmica *tu*/você em amostras históricas: uma breve revisão do tema

Considerando a dinâmica *tu/você* no século XX, voltamo-nos aos resultados das análises embasadas em amostras da produção escrita de brasileiros das regiões sudeste (Rio de Janeiro e Minas Gerais) e nordeste (Bahia, Pernambuco, [4]Rio Grande do Norte), *cf.* ilustramos no gráfico 1.

[4] Para as localidades de Pernambuco e do Rio Grande do Norte, os resultados expostos estão discutidos em Lopes *et al.* (2018).

[5]**Gráfico 1:** As formas *vossa mercê*, *você*, *tu*, *senhor*(*a*) distribuídas pela produção escrita de brasileiros do nordeste e do sudeste entre os séculos XIX e XX



Fonte: Barros (2024, p. 24)

Voltando o foco especificamente aos estudos sobre a variação *tu/você*, passamos à descrição de resultados de pesquisas embasadas em cartas pessoais do Nordeste (Batista, 2017; Gomes; Lopes, 2016; Lopes *et al.*, 2018) e do Sudeste (Souza, 2012; Souza, 2021) do Brasil. Passamos a uma síntese dos resultados de análises linguísticas conduzidas pelos encaminhamentos teóricos da Sociolinguística Histórica e da Teoria do Poder e da Solidariedade e pela produção escrita de brasileiros nascidos em tais regiões brasileiras. De um modo geral, observamos que temos, nos estados do sudeste (Rio de Janeiro e Minas Gerais), a dinâmica variável mais intensa entre as formas *tu* e *você*. Por outro lado, no âmbito do nordeste do Brasil (Rio Grande do Norte, Pernambuco e Bahia, respectivamente), observamos que o *você* prevalece, ainda que em convivência com o *tu* e com as formas nominais de tratamento *senhor* e *vossa mercê*, resistindo esta ao momento histórico do século XIX e àquela, até os dias atuais.

Com base na análise de trezentos e cinquenta e quatro cartas pessoais produzidas no intervalo de cem anos (entre 1870 e 1970) por escreventes nascidos e/ou residentes no espaço geográfico do Rio de Janeiro, Souza (2012) volta o foco da sua análise à alternância *tu/você*. Conduzida pelos princípios teórico-metodológicos da Sociolinguística Variacionista e da Teoria do Poder e da Solidariedade, Souza conjecturou que a forma *tu* seria a forma pronominal mais produtiva no século XIX e no início do século XX, sendo, posteriormente, suplantada pelo inovador *você*, a partir dos anos de 1930 do século XX (Duarte 1993). A partir da análise das mil trezentos e oitenta e uma (1381) ocorrências das

[5] Legenda (gráfico 1): RN = análise de Moura e Martins (2018) embasada em amostras de cartas pessoais representativas do Rio Grande do Norte; PE = análise de Gomes e Lopes (2016) embasada em amostras de cartas pessoais representativas de Pernambuco (séculos XIX e XX); BA = análise de Batista (2017) embasada em amostras de cartas pessoais representativas da Bahia (séc. XX); MG = análise de Souza (2021) embasada em amostras de cartas pessoais representativas de Minas Gerais (séculos XIX e XX); RJ = análise de Souza (2012) embasada em amostras de cartas pessoais representativas do Rio de Janeiro (séculos XIX e XX).

formas *tu* e *você*, a autora detectou o *você*, em 44,82% (619 oco), e o uso do *tu*, em 55,18% (762 oco) entre os séculos XIX e XX, confirmando a hipótese de a década de 30 do século XX ter se mostrado como um marco da implementação crescente do *você* (Duarte, 1993). De um modo geral, a autora admite que o *você* é menos produtivo nas dinâmicas sociais marcadas pelo maior nível de intimidade, como é o caso das relações sociais entre casais (relações amorosas).

Ao voltar-se às fontes históricas mineiras, Souza (2021) descreve o nível de alternância entre as formas *tu/você* correlacionando-as às relações sociais simétricas e assimétricas que as conduzem, tendo em vista os princípios da Teoria do Poder e da Solidariedade (Brown; Gilman, 1960), da Sociolinguística Histórica (Hernández-Campo; Conde Silvestre, 2012; Hernández-Campoy; Shilling, 2012) e da Sociolinguística Laboviana (Weinreich, Labov e Herzog, 2008 [1968]; Labov, 1994). Partindo da hipótese já testada por Rumeu, Cruz e Cardoso (2018) de que o *você* seria mais produtivo do que o *tu*, Souza (2021) volta-se ao nível da alternância *tu/você* em função das relações de Poder e de Solidariedade estabelecidas por redatores nascidos e/ou residentes no espaço geográfico mineiro entre os séculos XIX e XX. Conduzido pela análise de 332 cartas pessoais, a partir das quais levantou 1194 ocorrências, chegou à identificação de 573 oco de *tu* (47,98%), 542 oco de *você* (45,39%) e 79 oco de *vossa mercê* (6,63%). Considerando que a proposta de Souza é a de submissão do dados ao RBrul em interface com o RStudio para a geração dos pesos relativos, chegou o autor à conclusão de que as formas verbais de 3SG (concordância), o *você-sujeito*, as formas clíticas de 3SG, as formas possessivas de 3SG e as formas de *você não-sujeito* mostram-se como contextos que tendem a disparar o *você* na produção escrita mineira. No âmbito das relações sociais, Souza identificou o *vossa mercê* circunscrito às relações assimétricas ascendentes (de inferior para superior), enquanto as formas *tu* e *você* mostram-se prevalentes nas relações de simetria social, o que conduz o autor a interpretar a sociedade mineira com uma dinâmica social orientada pelos caminhos da Solidariedade, Brown e Gilman (1960 *apud* Brown; Gilman, 2003).

Voltando o foco especificamente a um estudo sobre a variação *tu/você* plenos embasado em cartas pessoais nordestinas, passamos à descrição dos resultados de Batista (2017), conduzidos pela análise de amostras representativas da escrita baiana do século XX (1930 e 1980). Trata-se de uma análise voltada à produção escrita de escreventes baianos pouco e medianamente escolarizados no âmbito das dinâmicas sociais mais e menos íntimas, o que está consubstanciado nas amostras de cartas familiares (entre informantes unidos por laços consanguíneos), amorosas (entre casais) e de amizade (entre amigos íntimos), no decorrer das décadas de 30 e 80 do século XX. A análise linguística das formas *tu/você* está orientada pelos princípios teóricos e metodológicos da Sociolinguística Quantitativa (Labov, 1994) e da Teoria da Polidez de Brown e Levinson (1987). De um modo geral, os resultados evidenciaram à autora o *você* como a estratégia preferida (75,2%) pelos redatores baianos nas relações simétricas mais solidárias e nas simétricas menos solidárias, *cf.* gráfico 1. Constatou a autora que as relações simétricas menos solidárias se mostram como propulsoras do uso do *você* amplo e generalizado nesse contexto de uso. Por outro lado, o *tu* mostra-se com um uso exíguo (24,8%), restringindo-se à intimidade das relações mais solidárias, *cf.* Batista (2017, p. 109).

Embasadas nas amostras representativas da produção de escreventes cultos do estado brasileiro de Pernambuco, Gomes e Lopes (2016) conduzem-se pelo levantamento de cento e vinte e três (123) cartas pessoais produzidas entre os anos de 1869 e 1969, configurando uma amostra composta por 47.500 palavras. Em termos gerais, as autoras chegam ao levantamento de 354 ocorrências de formas de referência ao sujeito de 2SG, sendo 279 oco de *você* (79%), 64 oco de *tu* (18%), 5 oco de *vossa mercê* (1%) e 6 oco de *senhor*/*senhora* (2%), *cf.* descrito no gráfico 1. A produtividade da forma *você* evidencia a sua alta frequência de uso nas distintas relações sociais, comprovando o seu traço polifuncional já evidenciado em dados de amostras oitocentistas do PB escrito. Isso posto, as autoras admitem que o *você* parece ter se firmado, já no início da era novecentista, como uma forma de referência ao interlocutor *neutra*, mostrando-se propícia às distintas relações sociais e às suas alterações no desenrolar do tempo.

A análise de Moura e Martins (2018), voltada à dinâmica *tu/você* em amostras de cartas pessoais representativas do estado nordestino do Rio Grande do Norte, conduziram-nos pelo levantamento de dados de *tu/você* em trezentas e quatro (304) cartas pessoais novecentistas (1916-1994). Dentre as oitocentas e noventa e duas evidências pronominais de 2SG (892 oco), os autores identificam o *você*, 756 oco (84 %), acompanhado pelo *tu*, em 86 oco (9%), e pela forma *senhora*, em 50 oco (5%). Em termos de relações sociais, os autores constatam o seguinte: a) o uso quase categórico do *você* nas relações familiares simétricas (entre irmãs) e relações simétricas entre amigos; b) o uso da forma *senhor*/*senhora* no âmbito das relações assimétricas ascendentes (de filho para mãe) como evidência de um tratamento mais cortês e respeitoso na produção escrita de 1940; c) o tímido uso do *tu*, que, com frequência de uso de 5%, figura como uma forma pronominal de maior intimidade entre os interlocutores.

2. Orientação teórica e metodológica: critérios e amostras

Os movimentos da variação e da mudança linguísticas tão visíveis em sincronias recentes repercutem o fato de também, no passado, as línguas humanas terem sido conduzidas por tais potenciais como expressão do *Princípio do Uniformitarismo da Mudança Linguística – The Uniformitarian Principle* (Labov, 1994). Isso quer dizer que um dado fenômeno linguístico, atualmente em variação, pode também ser observado em sincronias passadas (Labov, 1994). Com o intuito de detectarmos concretamente as evidências históricas da variação e da mudança linguísticas, é necessário voltarmo-nos às fontes que, por sua vez, devem estar, o máximo possível, aproximadas da sua expressão vernacular em sincronias passadas. Para isso, investimos esforços, também à luz de [6]Schneider (2002 *apud* Conde Silvestre, 2007, p. 46), na construção das nossas próprias amostras de textos não oficiais e, a partir delas, temos produzido análises linguísticas consubstanciadas principalmente em cartas pessoais.

[6] Registros escritos cuya relación con las variedades vernáculas orales del pasado se puede intuir, como determinadas expresiones que aparecen en cartas privadas o en los diarios personales, especialmente si sus autores son hablantes semialfabetizados en cuyo caso la influencia del estándar es menor. (Schneider, 2002 *apud* Conde Silvestre, 2007, p. 46)

Acreditamos que os textos não oficiais estejam mais propensos a nos aproximar, em alguma medida, da expressão da língua em uso. Acreditamos que é tarefa do linguista-pesquisador levantar e selecionar fontes históricas, atentando ao "problema dos filtros" (Romaine, 1982), que é o que nos permite entrever o vernáculo, distinguindo-o também do que é formulaico, ou seja, do que é peculiar ao gênero textual em análise. Uma outra questão igualmente importante é o fato de termos de fazê-lo, ao voltarmo-nos à história da língua, a partir de seus resquícios escritos, o que exige do linguista-pesquisador a habilidade de lidar com amostras textuais que podem ser [7]"[...] fragmentárias, escassas e dificilmente vinculáveis à produção real de seus falantes" (Conde Silvestre, 2007, p. 35). Assim, esse panorama nos conduz aos traços que se deixam entrever nas fontes históricas como evidências que, por sua vez, resistiram à ação do tempo no interior dos arquivos públicos e privados e que, portanto, são passíveis de análises sociolinguísticas.

No âmbito da Sociolinguística Histórica, os estudos se dão a partir da construção de "[...] seus próprios objetivos, metodologias e teorias [...]" (Bergs, 2005, p. 21). Para tal, não podemos meramente transpor a metodologia variacionista laboviana às amostras linguísticas de sincronias passadas. Por outro lado, temos de construir nossas próprias amostras históricas conduzidos por criteriosos parâmetros de organização dos textos tais como a *autenticidade,* a *autoria* e a *validade social e histórica* das fontes (Hernández-Campoy: Schilling, 2012), a fim de alcançarmos confiáveis resultados nas nossas análises, consubstanciadas, pois, em fontes linguísticas fidedignas a um dado momento histórico e social.

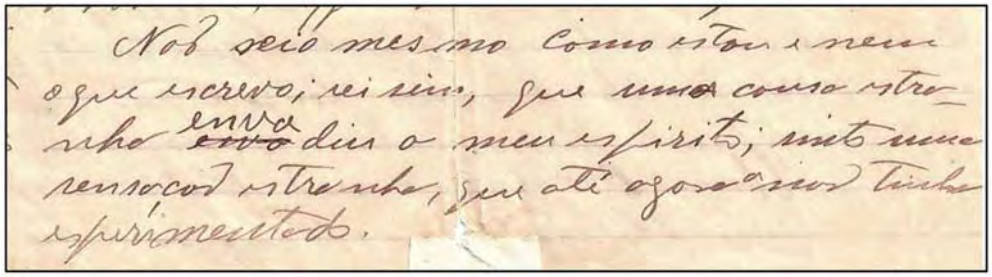
O parâmetro da *autenticidade,* segundo Hernández-Campoy e Schilling (2012, p. 68), está relacionado à "pureza dos textos", tendo em vista os possíveis indícios de *erro*, *mistura dialetal* e *hipercorreção* (Labov, 1994). É importante que tenhamos consciência sobre o fato de que as cartas novecentistas analisadas não refletem a fala "real" (*ipsis litteris*) dos seus autores, ou seja, não são transcrições da fala. Acrescentemos ainda o fato de que a expressão linguística em sincrônicas passadas somente se corporificar a partir de fontes escritas. Assim sendo, as fontes históricas tendem a se mostrar mais sujeitas às pressões da norma-padrão, funcionando, pois, como um dinamizador das suas próprias convenções vinculadas ao gênero textual "carta pessoal".

Por outro lado, como são cartas pessoais, refletindo a intimidade do escrevente, tendem, em essência, a minimizar essa forte pressão da norma-padrão. Desse modo, as cartas pessoais constituem um contexto propício às oscilações entre o padrão e o não padrão, quer por descuido, quer por hipercorreção. Em (3), temos evidências que parecem apontar para a *autenticidade* da amostra. Nele, o escrevente JMF assumiu o item lexical "*seio*" no lugar da forma verbal "sei" e, na linha seguinte, segue a forma verbal "*sei*", evidenciando a oscilação entre a grafia orientada pela noma-padrão e a sua expressão vernacular na língua falada. O mesmo escrevente, no excerto da casa em análise (imagem 1), preferiu "*esvadiu*" no lugar da forma "*invadiu*" e, ao perceber o equívoco, corrigiu para "*envadiu*",

[7] "En comparación con la *diversidad*, *cantidad* y *autenticidad* de los datos a disposición del investigador en sociolingüística sincrónica o en lingüística descriptiva, la información de que dispone quien intenta desarrollar su investigación en el ámbito de la lingüística o la sociolingüística histórica es fragmentaria, escasa y difícilmente vinculable con la producción real de sus hablantes." (Conde Silvestre, 2007, p. 35)

mas preserva a sílaba inicial "*en*" no lugar de "*in*" como um possível indício de *hipercorreção*. Além disso, temos o item lexical "esperimentado" grafado com "s" no lugar da letra "x" ("experimentado").

Imagem 1: Evidência de aspectos ortográficos e de hipercorreção
JMF. Tatu [Oeiras-PI], 19.08.1920

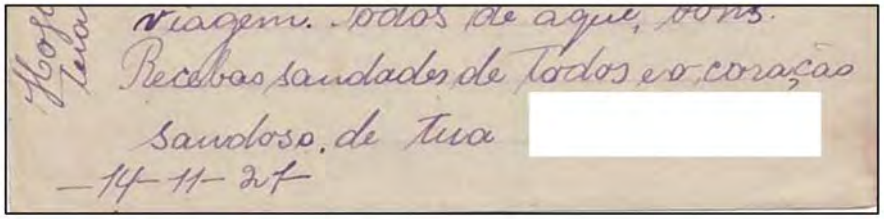


Fonte: Barros (2024, p. 38)

(3) "[...] Não seio mesmo como estou e nem o que escrevo; sei sim, que uma cousa estranha ~~esva~~ <enva↑>diu o meu espirito; sinto uma sensação estranha, que até agora não tinha esperimentado. [...]" JMF. Tatu [Oeiras-PI], 19.08.1920.

Quanto à *autoria* das missivas piauienses em análise, devemos destacar que, se um dos objetivos da SH é identificar traços da língua em sincronias passadas e estabelecer uma correlação entre esses traços e a estrutura social através da qual as mudanças linguísticas também se implementam, é importante que o pesquisador certifique-se sobre a origem de quem escreveu os textos históricos. Essa certeza pressupõe conhecimentos no âmbito da paleografia, de modo que o linguista-pesquisador certifique-se sobre o texto ter sido escrito pela própria mão de quem o assinou (testemunho autógrafo), por um punho distinto daquele de quem o assinou (testemunho apógrafo) ou ainda produzido por outra pessoa (testemunho idiógrafo), mesmo que sob a supervisão do autor (autoria intelectual). Nas cartas piauienses novecentistas, temos também as cartas produzidas e assinadas pelo próprio punho dos redatores (autógrafas) que as redigiram. Nas cartas manuscritas, passamos pela análise comparativa entre a letra do redator no corpo da sua carta e a da sua assinatura, o que nos permitiu também a comparação em relação à assinatura dos autores nas cartas datiloscritas. Em (4), temos evidências de carta amorosa, autógrafa e manuscrita produzida por uma mulher que escreve a seu noivo (imagem 2), ainda no início do século XX.

Imagem 2: Excerto da carta de CRJM ao noivo. S.l, 14.11.1921

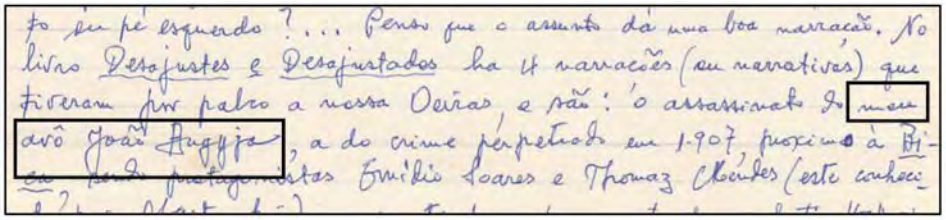


Fonte: Barros (2024, p. 41)

(4) "[...] *Recebas* saudades de todos e o coração saudoso. de tua [...]14-11-21"

O rigor metodológico em relação não só à *autenticidade*, mas também à *autoria*, exige do linguista-pesquisador também a busca pelo levantamento dos perfis biográficos dos redatores, a fim de comprovar que se trata de escreventes brasileiros, nascidos, no caso específico desta análise, no contexto histórico-social novecentista do Piauí. Ainda que a reconstituição dos perfis sociais dos redatores seja uma tarefa trabalhosa, é necessária em virtude do fato de que "geralmente sabemos muito pouco sobre a posição social dos informantes, e não mais acerca da estrutura social da comunidade" (Labov, 1994). Além das próprias cartas terem servido com fontes que nos permitiram confirmar informações biográficas dos redatores (imagens 3 e 4), valemo-nos não só de outras fontes documentais (fontes secundárias) tais como publicações em jornais, índices genealógicos, mas também de testemunhos orais de familiares e amigos nesse processo de reconstituição do perfil biográfico de alguns dos redatores piauienses (todos nascidos no espaço geográfico do Piauí) em análise. Em (5), temos o testemunho do redator ABSB acerca da referência ao assassinato do seu avô "JB" (imagem 3), além de, nesta mesma carta (imagem 4), o redator referir-se ao seu interlocutor (PQ) como "confrade, amigo e conterrâneo", confirmando, em (6), se tratar de carta genuinamente produzida por escrevente nascido no espaço geográfico do Piauí, mais precisamente, em Oeiras.

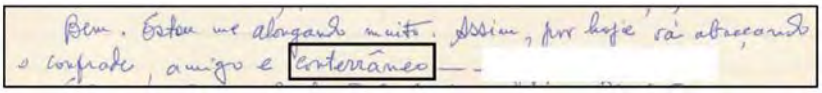
Imagem 3: Excerto da carta de ABSB
Piauí-PI, 31.01.1981



Fonte: Barros (2024, p. 43)

(5) "[...] No livro Desajustes e Desajustados ha 4 narrações (ou narrativas) que tiveram por palco a nossa Oeiras, e são: 'o assassinato do meu avô João Bugyja [...]" ABSB. Fazenda Riacho Fundo, São Pedro do Piauí-PI, 31.01.1981.

Imagem 4: Fechamento da carta de ABSB
Piauí-PI, 31.01.1981



Fonte: Barros (2024, p. 43)

(6) "[...] Bem. Estou me alongando muito. Assim, por hoje vá abraçando o confrade, amigo e conterrâneo [...]" ABSB. Fazenda Riacho Fundo, São Pedro do Piauí-PI, 31.01.1981.

Em se tratando da *validade social e histórica*, segundo Hernández-Campoy e Schilling (2012, p. 70), temos em análise um parâmetro também voltado ao conhecimento que se tem acerca da posição social dos escreventes e da estrutura social da comunidade linguística da qual faziam parte, o que legitima o trabalho de reconstrução dos perfis sociais dos informantes para que se consolidem acertadamente as interpretações relacionadas aos padrões de variação e mudança em sincronias passadas (Milroy; Gordon, 2003 *apud* Hernández-Campoy; Schilling, 2012). Assim sendo, a identificação da estrutura social será tão mais próxima da realidade pretérita, na proporção direta da reconstituição feita em relação à *autoria* e à *autenticidade* das fontes históricas.

Além de empenharmo-nos por implementar os princípios da sociolinguística histórica (*autoria*, *autenticidade* e *validade social e histórica*) no processo de seleção das cartas piauienses, conduzimo-nos também pelos princípios teórico-metodológicos da Teoria da Variação e Mudança (Labov, 1994; Weinreich; Labov; Herzog, 2006 [1968]). Ao passarmos ao levantamento das formas de referência ao sujeito de 2SG enfocando especificamente as formas pronominais *tu* e *você*, consideramos tal dinâmica de alternância como expressão da *variável dependente* (*tu*/*você*) que foram analisadas em relação às variáveis independentes (internas e externas). São elas: (1) a *expressão do sujeito* de 2SG (nulo ou pleno); (2) a *pessoa verbal*; (3) o *paralelismo formal e semântico*; (4) os *subgêneros* das *cartas pessoais* (*cartas de amizade, amorosas* e *familiares*); (5) *As temáticas prevalentes das cartas pessoais*; (6) o *sexo do informante*; (7) a *faixa etária* dos missivistas; (8) os níveis *escolaridazação dos informantes*; (9) os níveis de *prática de escrita*; (10) as *relações sociais* (simétrica e assimétrica); (11) os *períodos das cartas* e (12) os *acervos históricos*. A partir do cruzamento desse doze grupos de fatores alcançamos os resultados estatísticos (via excel) e probabilísticos (via RBrul).

3. As amostras de cartas pessoais piauienses novecentistas: quadro geral

Neste texto, embasamos a nossa análise nas missivas pessoais produzidas por treze (13) redatores históricos com níveis de escolarização que variam do tão somente nível alfabetizado ao nível superior completo, assim distribuídos: três redatores (3) com nível superior; três (3) com ensino médio; um (1) autodidata (músico, historiador, professor, escritor, advogado rábula) e seis (6) alfabetizados. Temos uma amostra bem diversificada no que tange à formação escolar dos redatores e às suas respectivas habilidades de leitura e escrita. Expomos, no quadro 1, a distribuição das amostras de cartas pessoais, em seus subgêneros (*amorosas, amizade e familiares*), em relação aos sete acervos (conjuntos documentais) e ao eixo do tempo (1920-1990).

Quadro 1: Amostras de cartas pessoais piauienses: o século XX em cena

Acervos	Cartas		Períodos	Amor	Amizade	Familiar
	h (5)	m (2)				
Autores piauienses	42	-	1946 - 90	-	42	-
Batista	04	-	1978	-	04	-
Cartas pessoais variadas	08	01	1953 - 76	-	07	02
Moura Fé	26	10	1920 - 22	36	-	-
Queiroz	27	-	1975 - 90	04	02	21
Santos Rego	-	07	1978 - 81	-	07	-
Vieira do Vale	11	-	1981 - 84	11	-	-
Total	118	18	1920 - 90	51	62	23
				136		

Fonte: Barros (2024, p. 50)

Para esta análise, munimo-nos de amostras linguísticas compostas por 136 cartas pessoais (*amorosa* (51), *familiar* (62) e de *amizade* (23)) produzidas por missivistas piauienses (homens e mulheres), entre os anos de 1920 e 1990. Como temos em evidência a produção escrita de redatores com distintos graus de habilidade escrita, voltamo-nos à dinâmica *tu/você* e à sua distribuição em relação aos níveis de escolarização e, consequentemente, de habilidade dos redatores piauienses (mais ou menos hábeis).

Em análises sociolinguísticas voltadas às sincronias passadas, voltamo-nos não só aos níveis de contato dos redatores com os modelos de escrita (práticas de escrita nos níveis baixo, mediano e alto), mas também ao seus graus de escolarização formal. Nesse sentido, passamos a um quadro resumitivo (quadro 2) a partir do qual correlacionamos os níveis de escolarização formal dos redatores piauienses aos seus papeis sociais.

Quadro 2: Quadro-síntese da correlação entre os níveis de escolarização e os perfis sociais dos redatores

Acervos (1920-1990)	Redatores	Níveis de escolarização	Papeis sociais (Profissão principal)
1. Autores piauienses (1946-87)	1. ABSB (1946, 1976-90) 2. JATF (1972-87)	Ensino Superior Ensino Superior	Advogado Jornalista
2. Batista (1978)	1. VB (1978)	Ensino Médio	Estudante
3. Cartas pessoais variadas (1951-78)	1. JRC (1976) 2. JSL (1964) 3. LJL (1951) 4. JPB (1978) 5. VVS (1953-54)	Alfabetizado Alfabetizado Alfabetizado Alfabetizado Alfabetizado	Comerciante Do lar Agricultor Operário de fábrica Funcionário público municipal
4. Moura Fé (1920-21)	1. JMF (1920-22) 2. CRJM (1920-21)	Ensino Superior Alfabetizada	Farmacêutico Do lar
5. Queiroz (1975-90)	1. PNQ (1975-90)	Autodidata	Funcionário Público Municipal
6. Santos Rêgo (1978-81)	1. MNRSR (1978-81)	Ensino Médio	Funcionária Pública do Distrito Federal
7. Vieira do Vale (1981-84)	1. AVV (1981-84)	Ensino Médio	Estudante

Fonte: Barros (2024, p. 62)

De um modo geral, está em análise, neste estudo, a produção escrita de treze (13) redatores piauienses que produziram suas missivas novecentistas entre os anos de 1920 e 1990. Trata-se de redatores alfabetizados (JRC, JSL, LJL, JPB, VVS), com o ensino médio concluído (CRJM, MNRSR), inconcluso (VB e AVV), com o ensino superior (JMF, ABSB, JATF) e um redator autodidata (PNQ). Dentre os papeis sociais dos redatores alfabetizados, temos um agricultor (LJL), um comerciante (JRC), um operário de fábrica (JPB) e um funcionário público municipal (VVS). Para os escreventes com o ensino médio concluído, observamos se tratar de uma funcionária pública do Distrito Federal (MNRSR) e de uma informante "dona de casa" (JSL). Nessa amostra, temos dois redatores ainda em processo de conclusão do ensino superior, ou seja, atuando como estudantes do ensino superior ainda inconcluso (VB e AVV). Quanto aos informantes com o ensino superior, temos um farmacêutico (JMF), um advogado (ABSB) e um jornalista (JATF). Além disso, temos as cartas de PNQ, um autor autodidata, que concluiu as Primeiras letras, em Oeiras-PI, e um breve curso de Escrituração Mercantil, em Teresina-PI.

Com base na análise de cartas piauienses de sincronias passadas (século XX), a proposta é acompanharmos o nível de inserção do *você* no sistema pronominal do PB. Passamos, na sequência, aos resultados gerais da produtividade variável das formas *tu* e *você* no Brasil do século XX.

4. As formas *tu* e *você* na escrita piauiense novecentista: resultados gerais

Com o intuito de descrevermos a produtividade das formas de referência ao sujeito de 2SG, a hipótese é a de que o *tu* também se mostre em alternância com o *você* na escrita piauiense novecentista. Voltamo-nos ao passado a partir do presente, contexto em que a produtividade do *tu* está abaixo de 60%, com concordância entre 10% e 39% na fala piauiense (subsistema 5), *cf.* a discussão de Scherre *et al.* (2021, p. 145) à luz dos dados do Atlas Linguístico do Brasil – Alib: inquéritos geolinguísticos (Cardoso, 2008; Deus, 2009). A fim de darmos conta de tal hipótese, passamos à distribuição, em termos percentuais, das formas *tu* e *você* (tabela 1) como uma regra variável binária (Guy; Zilles, 2007).

Tabela 1: As formas de referência ao sujeito de 2SG nas cartas piauienses (século XX)

Formas de referência ao sujeito de 2SG	
Tu	**Você**
233/583	350/583
(39,97%)	(60,03%)

Fonte: Barros (2024, p. 88)

Em termos gerais, observamos, com base na análise da tabela 1, uma instigante disputa entre as formas *tu* (233/583, 39,97%) e *você* (350/583, 60,03%) na produção escrita piauiense do século XX. Convém esclarecermos que desses quinhentos e oitenta e três (583) dados de *tu*/*você* foram excluídas vinte e três (23) ocorrências de *vós* em referência a 2SG, uma vez que voltamos à regra variável em questão que está voltada à 2SG. Além disso, não nos preocupamos com os dados de *tu nulo* e *você nulo*, ambos resguardados pela morfologia da forma verbal em sua expressão imperativa, mas levantamos e codificamos os dados de *você* nulo correlacionado a uma primeira ocorrência de *você pleno*, bem como os dados de *tu nulo* (desinencial ou não expresso) e de *tu pleno*. Em (7) e (8), apresentamos evidências das formas de referência ao sujeito de 2SG (*tu* versus *você*) nas cartas piauienses analisadas.

(7) "[...] Eu não esperava nem encontrava outro rapaz milhor que *tu*, portanto filhinho te pesso que não diga mais que *tu* es indigno de minha pessoa [...]" (CRJM. Curral de Pedras, 11.03.1921.)

(8) "[...] Aí, *você* pode deixar de andar em festas. Muitas vezes *você* saía com Vitória para festa [...] Mas pela manhã [...] já *você* havia chegado [...]" (PNQ. Oeiras-PI, 12.01.1988.)

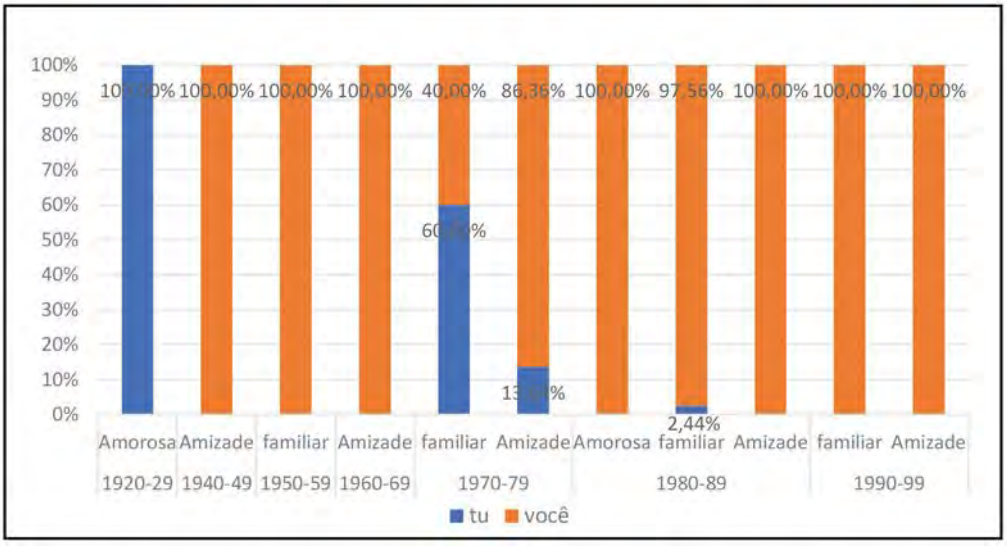
Considerando as quinhentas e oitenta e três ocorrências das formas *tu* e *você*, observamos, em termos gerais, uma estimulante concorrência entre as formas *tu* (233 oco) e *você* (350 oco), o que nos leva a pensar qual seria a distribuição de tais dados em relação ao eixo do tempo e ao subgênero da carta pessoal (amizade, amor, familiar).

Tabela 2: Distribuição das formas de referência ao sujeito de 2SG (*tu* e *você*) pelas cartas piauienses no eixo do tempo: de 1920 a 1990

Períodos/Cartas	Tu	Você	Total
1920-29 (36 cartas)	204/204	-	204/583
Amor	204/204 (100%)	-	204/583 (34,99%)
1940-49 (1 carta)	-	**1**/583	1/583
Amizade	-	1/1 (100%)	1/583 (0,17%)
1950-59 (5 cartas)	-	**17**/17	17/583
Familiar	-	17/17 (100%)	17/583 (2,92%)
1960-69 (1 carta)	-	**1**/1	1/583
Amizade	-	1/1 (100%)	1/583 (0,17%)
1970-79 (19 cartas)	27/96 (28,13%)	69/96 (71,88%)	96/583 (16,47%)
Familiar	18/30 (60%)	12/30 (40%)	30/583 (5,15%)
Amizade	9/66 (13,64%)	57/66 (86,36%)	66/583 (11,32%)
1980-89 (71 cartas)	2/259 (0,77%)	257/259 (99,23%)	259/583 (44,43%)
Amorosa	-	76/76 (100%)	76/583 (13,03%)
Familiar	2/82 (2,44%)	80/82 (97,56%)	82/583 (14,06%)
Amizade	-	101/101 (100%)	101/583 (17,32%)
1990-99 (3 cartas)	-	5/5 (100%)	5/583 (0,86%)
familiar	-	2/2 (100%)	2/583 (0,34%)
Amizade	-	3/3 (100%)	3/583 (0,51%)
Total Geral	233/583 (39,97%)	350/583 (60,03%)	583/583 (100%)

Fonte: Inspirado em Barros (2024, p. 111)

Gráfico 2: Distribuição das formas *tu* e *você* em relação aos subtipos de cartas pessoais (amorosa, amizade, familiar) no eixo do tempo (1920-1990)



Fonte: Elaborado pelos autores

Apresentamos, através de uma análise conjugada entre o gráfico 2 e a tabela 2, um panorama do português novecentista a partir de amostras quantitativamente desproporcionais em relação ao número de cartas pessoais, uma vez que ficamos à mercê dos textos que sobreviveram à ação do tempo no interior dos arquivos, o que é típico dos trabalhos no âmbito da sociolinguística histórica. Isso posto, passamos às distribuições dos dados de *tu* e *você* em relação aos subgêneros das cartas pessoais e às décadas de 20, 40, 50, 60, 70, 80 e 90 do século XX. Nesse cenário, confirmamos a prevalência do *você*, *cf.* tabela 1, tendo sempre em vista as amostras de cartas novecentistas em análise. Considerando que o *você* prevalece independentemente do subgênero da carta pessoal, passamos à descrição pormenorizada dos pouquíssimos dados de *tu*, analisando-os qualitativamente, a partir das evidências levantadas, nas cartas piauienses das décadas de 20, 70 e 80.

Com base na análise das trinta e seis cartas (36) pessoais da década de 20 do século XX, temos resultados destoantes em relação aos demais períodos, uma vez que se trata do único período em que o *tu* se mostra categórico (204 ocorrências). Essas cartas amorosas dos anos 20 foram produzidas por um redator (o noivo) com formação superior completa, ao passo que a redatora (a noiva) apresenta-se tão somente escolarizada. A prevalência do *tu* nesse subgênero textual parece estar em consonância com os resultados de Silva (2012, p. 97) que detectou, na produção escrita do noivo Jayme, a preferência pelo *tu* para tratar a sua noiva Maria, em 73% dos dados (147/168 ocorrências), na década de 30 do século XX.

Do acervo *Moura Fé*, temos as *cartas de amor*, sendo vinte e seis (26) redigidas por JMF e dez (10) por CRJM. Essas cartas foram trocadas entre o casal, durante o período de 19 de agosto de 1920 a 26 de fevereiro de 1922, registrando o relacionamento, desde o pedido de namoro/noivado, feito por JMF a CRJM, até a data do casamento, 26 de fevereiro de1922. Temos em cena, no Acervo

Moura Fé, a produção escrita de 2 (dois) escreventes, sendo um redator do sexo masculino (JMF), com nível universitário, adulto, entre 31 e 32 anos e uma redatora (CRJM) do sexo feminino, apenas alfabetizada, mas uma hábil redatora jovem, entre 19 e 21 anos.

Ainda que tenhamos, na década de 20 do século XX, a produção de redatores bem "treinados" na prática da escrita, o subgênero *carta amorosa* também parece se confirmar como um contexto promissor ao *tu*-íntimo nesse contexto. Em (9) e (10), ilustramos o *tu* nas cartas amorosas produzidas, entre os anos de 1920 e 1929, por punhos piauienses.

(9) "[...] Orgulhoso por te amar e merecer a tua affeição, por *seres* uma pessoa virtuosa afavel, carinhosa, meiga e terna, e, vaidoso por *seres* um thesouro de belleza. [...]" (JMF. Simplício Mendes-PI, 16.09.1920.)

(10) "[...] Mas, finalmente se aproxima a hora da felicidade suprema para mim e *tu* meu amorsinho, pois sei que me *amas* sinseramente. [...]" (JMF. Simplício Mendes-PI, 12.02.1922.)

Nas décadas de 40 (em 1 carta de amizade), 50 (em 5 cartas familiares) e 60 (em 1 carta de amizade), a forma *tu* prevalece categoricamente a partir de pouquíssimos dados. São eles: uma (1) ocorrência, dezessete (17) ocorrências e mais uma única (1) ocorrência, respectivamente. Por outro lado, nas dezonove cartas familiares trocadas, entre os anos de 1970 e 1979, o *tu* mostra-se produtivo em 60% dos dados (18/30). Passemos à análise desses dados de *tu*. É interessante observar que, nas cartas familiares, as 18 ocorrências de *tu* foram produzidas por uma única redatora (MNRSR) que era estudante universitária e trabalhava como funcionária pública no Distrito Federal. Nessas cartas, a redatora, entre 23 e 26 anos, escreveu à prima sobre o seu cotidiano como uma jovem universitária, a prática de esportes e namoro. Parecem-nos um tanto quanto inesperadas as evidências do *tu* com concordância (*sabes, entendes*) na produção escrita de uma redatora com o ensino médio, atuante como funcionária pública do Distrito Federal e nascida no Piauí, espaço onde atualmente, em sua capital, prevalece o *você* em convivência com o *tu* com concordância média (entre 10% e 39%), *cf.* Scherre *et al.* (2015, p. 141). O fato de essas cartas familiares serem íntimas parece ratificar a opção pelo *tu*, que nessa dinâmica de simetria social, o contextualiza também nas cartas piauienses novecentistas. Segundo Scherre *et al.* (2015, p. 143), ainda que a região centro-oeste seja, de um modo geral, marcada pelo subsistema *só você*, temos, especificamente, no espaço do Distrito Federal, a inserção do subsistema *você/tu sem concordância* a partir, possivelmente, dos anos 2000. É possível, pois, que esses dados de *tu* (*tu* com concordância) evidenciem, já na década de 70 do século XX, um leve anúncio do que vige atualmente no Distrito Federal (*você/tu sem concordância*, *cf.* Scherre *et al.* (2015, p. 143)), antecipando-o, em sincronias passadas. De (11) a (13), apresentamos algumas evidências do *tu* (*sabes*) que convive com o *você* (*você* sabe).

(11) "Cara prima [...] *sabes*, aconteceu comigo algo mais ou menos assim: - sempre fui meio focada, namoro sem botar chifre, o tipo de garota, que poderia se dizer "certinha", não é mesmo?!" (MNRSR. Brasília, 09.05.1978.)

(12) "Olá prima! [...] resolvi dividir este maravilhoso mundo com você. [...] parece não ter nenhum objetivo a atingir, nenhuma meta, *sabes*, é uma pessoa passando pela vida [...]" (MNRSR. Brasília, 20.08.1978.)

(13) "Cara prima [...] Acho que a Faculdade está me deixando meio pirada [...] *Você* sabe, não é? [...] *Sabes* Prima, as coisas aqui em casa andaram tão ruins [...]" (MNRSR. Brasília, 08.10.1978.)

(14) "[...] Cara Olivette. Olá! Beijos a todos. Acho que demorei um pouco a te escrever, mas tenho certeza que *entendes*. Logo que cheguei, reiniciei as atividades normais, trabalho, faculdade, inglês, e no serviço público você sabe que nessa época é fogo. [...]" (MNRSR. Sobradinho, Brasília, DF, 17.03.1979.)

Nas cartas de amizade da década de 70, levantamos, em 13,64% dos dados (9 ocorrências), algumas poucas evidências do *tu*. Para essas cartas de amizade, temos o redator JPB, responsável por oito (8) ocorrências de *tu*, estando três desses dados expostos de (15) a (17). Trata-se de um escrevente que, com o ensino médio, atuava profissionalmente como Digitador/Funcionário da Ford, em São Bernardo do Campo - SP, e, em suas duas (2) cartas ao amigo Rodney (Rodim), alterna entre o *tu* (com e sem concordância) e o *você*. A semântica dessas cartas é de intimidade, até mesmo em tom de cobrança, em virtude da ausência de notícias do seu amigo e da realidade piauiense que o redator parece não ter abandonado, ao migrar para São Paulo, *cf.* ilustramos de (15) a (17). Nessas cartas de JPB, temos em cena um redator que parece evidenciar traços da alternância *tu/você*, sendo que o *tu* mostra-se com (*imaginavas, ameaças*) e sem concordância (*tu não está, tu pensa*), o que, de alguma forma, transparece o atual sistema de alternância *tu/você* com concordância média (entre 10% a 39%), na capital do Piauí, *cf.* Scherre *et al.* (2015, p. 145). Ainda na década de 70, temos, em tão somente uma única (1) ocorrência, o *tu com concordância* realizado por VB, em carta voltada à sua amiga Olivette. Nessa carta, o escrevente trata do seu cotidiano em Fortaleza, incluindo a preocupação com as provas do seu curso superior (Geologia), *cf.* constatamos em (18). Nesse caso, VB evidencia o uso do *você* e do *tu* (com concordância), o que, de alguma forma, apesar do baixo número de cartas e, consequentemente, de dados, parece anunciar o atual sistema vigente em Fortaleza que é o subsistema de alternância *tu/você* com concordância média também em conformidade com o mapeamento de Scherre *et al.* (2015, p. 145).

(15) "Oi Rodim! Tudo legal [...] mais *você* não venha com essa, pois acredito que novidade aí não está faltando, estou quase na certeza que *você* esta se importunando com essas horríveis cartas que lhe remeto. [...] Olha! *tu* não está sendo como nós eramos, porque está escondendo as novidades, mais *tu* pensa que não sei? como se foi de festa na casa de toinho de chico e nas demais que estão se ocorrendo por aí? [...]" (JPB. São Bernardo do Campo - SP, 18.04.1974)

(16) "[...] Alô Rodim! Tudo bem? Existe um velho ditado que diz "tudo tem seu dia" e hoje estou a responder a sua carta, a qual me deixou muito feliz. Primeiro porque foi uma surpresa muito boa que *você* me fez. Pois não *imaginavas* que fosses me escrever mais. por que sempre é o que me *ameaças* [...] " (JPB. São Bernardo do Campo - SP, 05.06.1978.)

(17) "[...] pois não *imaginavas* que fosses me escrever mais. por que sempre é o que me *ameaças*. apesar de ter prometido, mas é sempre o que vejo acontecer, é quando agente fica ausente daí, vocês esquecem quem sai, e as noticias tornam-se difíceis, daí a razão da minha surpresa. [...]" (JPB. São Bernardo do Campo - SP, 05.06.1978.)

(18) "Caríssima Olivette! [...] Não sei se esta vai lhe encontrar em período de aula, mas acho que sim confiando na eficiência dos Correios. [...] Minha última prova será dia 01/07 quando partirei para Picos, embora só fique, nessa viagem, durante 2 dias, voltando em seguida para Fortaleza. Se *chegares* em Picos antes do dia 2 ou 3, ligue pra mim nos dias mencionado [...] Eu tenho quase que certeza que *você* chegará em Picos pelo dia primeiro de julho [...]" (VB. Fortaleza - CE, 25.06.1978.)

Nas setenta e uma cartas piauienses da década de 80, observamos poucas evidências do *tu*, em tão somente 2,44% dos dados (2/82), *cf.* dados expostos em (19) e (20). Em (19), temos a mesma redatora (MNRSR) que, na década de 70, também nos expôs algumas evidências de *tu* com concordância, mantendo o mesmo comportamento (ainda que em menor quantidade de dados), na relação entre primas, que é o uso do *tu* por influência, talvez, do espaço em que vivia, na década de 80, que ainda era o Distrito Federal. Além desse dado de *tu*, levantamos, ainda na década de 80, um único (1) dado de *tu*, em carta familiar trocada entre avô e neto, conforme apresentamos em (20). Nesse caso, trata-se de um redator (PNQ) que atuou como funcionário público municipal e que, como um autodidata, parece ter feito questão de se expressar, pois, em consonância com a norma-padrão. Acionou o *tu* com concordância (*queres*) em uma sentença de tom aconselhativo através da qual sugeriu ao neto os concursos públicos como a expressão de uma carreira *honesta e lucrativa*, segundo o redator, na realidade do Brasil novecentista.

(19) "Olivette. Cara prima, me conta como andam as coisas por esta terra? [...] Como *vês*, ando tão atarefada que não tenho nem mesmo tempo pra namorar [...] Veja que coisas lindas nos escreve o artur da Tavola [...] divido com você o prazer desta leitura." (MNRSR. Sobradinho-DF, 28.07.1981.)

(20) "[...] "Querido e sempre lembrado Neto CARLOS [...] Faça o concurso para o BANCO DO BRASIL, CAIXA ECONÔMICA ou qualquer atividade honesta e lucrativa. [...] Se *queres* vencer, desperta e luta. [...]" (PNQ. Oeiras-PI, 18.07.1987.)

Considerando o panorama apresentado, em termos estatísticos, dos níveis da alternância *tu/você*, passamos aos índices probabilísticos da regra variável em análise para a geração dos índices probabilísticos, passando, antes, pela apresentação desses parâmetros de análise (*peso relativo*, *p-value* e *logodd*).

5. A regra variável *tu/você* e os seus condicionamentos: uma análise da escrita piauiense novecentista.

Convém esclarecermos que a partir da análise da regra variável *tu/você*, em termos percentuais, produzida por Barros (2024) em relação às variáveis linguísticas (a *expressão do sujeito*, a *pessoa*

verbal e o *paralelismo formal e semântico*) e às extralinguísticas (o *sexo,* a *faixa etária,* as *relações interpessoais entre os missivistas,* o *período, a escolaridade*, *os níveis de prática de escrita*, o *subgênero da carta pessoal* e os *acervos*), passamos, nesta seção, especificamente à descrição dos parâmetros selecionados pelo Rbrul/Rstudio para a produtividade do *você* nas cartas piauienses novecentistas. Assim sendo, concentramo-nos nos resultados probabilísticos em função da correlação entre os índices do *peso relativo,* do *p-value* e do *logodd* para cada uma das variáveis selecionadas, nos níveis *up* and *down* (rodada multivariada), pelo RBrul em sua interface com o RStudio (Oushiro, 2014, p. 136), considerando a alternância *tu/você* como regra variável e a variante *você* como valor de aplicação (*application value*). Isso posto, passamos à pormenorização dos parâmetros *peso relativo, p-value* e *logodd* em relação à interpretação dos resultados da regra variável *tu/você*.

O *peso relativo* é a indicação da força dos contextos linguísticos e extralinguísticos, no interior dos grupos de fatores, em relação ao uso da variável dependente. Trata-se de um peso 'relativo' ao nível de aplicabilidade geral (*input*) de uma dada forma variante (variável dependente), conforme Guy e Zilles (2007). Os pesos relativos acima de 0.50 evidenciam uma tendência à aplicação da forma variante assumida como valor de aplicação (*você*), ao passo que pesos abaixo de 0.50 expõem uma tendência à inaplicabilidade da regra variável em cena, tendo sempre em vista o valor de aplicação em relação à interação entre todas as variáveis (linguísticas e extralinguísticas). Os *pesos relativos* revelam tendências probabilísticas que entram no âmbito da equação logística, correlacionados, portanto, aos demais pesos relativos, ao *input* da variável analisada, ao *p-value* do grupo de fatores, aos *logodds* dos fatores internos dos grupos de fatores.

O *p-value* está voltado à correspondência entre a variável dependente (variável preditora) e o conjunto das variáveis independentes (contextos linguísticos e extralinguísticos ou grupos de fatores). Nesta análise, utilizamo-nos do *p-value* para mensurar a força de um grupo de fatores em relação à aplicação da regra variável a partir do índice 0,05 (5%). Nesse sentido, o entendimento é o de que abaixo desse valor (< 0,05) temos uma variável (grupo de fatores/variável independente) impulsionadora da variável dependente. Por outro lado, o *p-value* igual ou acima de 0,05 (= ou > 0,05) coloca-nos diante um contexto refreador da regra variável em análise, ou seja, trata-se da não-correlação entre a variável preditora (variável dependente) e as independentes em análise.

Para os *logodds*, a interpretação é a de que se trata de um índice de aplicação geral de cada um dos fatores no interior de uma dada variável independente selecionada como relevante na rodada multivariada. Assim sendo, os resultados probabilísticos a partir dos *logodds* levam-nos aos índices positivos e negativos como indicação dos fatores que tendem a impulsionar e a refrear, respectivamente, a dinâmica variável em análise. Uma vez expostos os critérios desta análise probabilística no nível da rodada multivariada, passamos à descrição analítica de tais resultados.

Tabela 3: Síntese das variáveis independentes estatisticamente relevantes à aplicação do *você-sujeito* na ordem de seleção do RBrul/RStudio

Variáveis Linguísticas	Variáveis Extralinguísticas
1º) Pessoa *p-value*: 3.14e-151 3ª pessoa do singular (3SG): **0.999** (61,92%, 361/583oco) *logodd*: 103.506	4º) Faixa etária *p-value*: 0.00154 Idoso: **0.999** (40,48%, 236/583) *logodd*: 63.161 Adulto: **0.996** (39,45%, 230/583) *logodd*: 5.504
2º) Paralelismo formal e semântico *p-value*: 2.76e-08 Precedido por *você-sujeito*: **0.999** (78%, 191/245oco). *logodd*: 80.714 Precedido por *vocativos*: **0.993** (10%, 25/245). *logodd*: 4.948 Precedido por *construções imperativ*as de 2SG com *formas de subjuntivo*: **0.999** (5%, 12/245). *logodd*: 107.036 Precedido por formas de *tu não-sujeito*: **0.985** (4%, 10//245). *logodd*: 4.199245).	5º) Escolarização *p-value*: 0.00839 Ensino superior: **0.999** (47,68%, 278/583). *logodd*: 16.556 Ensino médio: **0.999** (20,41%, 119/583). *logodd*: 19.071
3º) Expressão do sujeito *p-value*: 0.00154 *Sujeito pleno*: **0.999** (46,83%, 273/583) *logodd*: 55.105	-
Regra variável: *tu/você* Valor de aplicação: *você*.	

Fonte: Barros (2024, p. 129)

A opção por controlar a variável *pessoa* está motivada pelo fato de o *tu* ser uma forma que etimologicamente é conduzida para as referências formal e semântica de 2SG, mas atualmente há a possibilidade de tal pronome estar correlacionado às formas verbais de 2SG ((*tu*) *sabes*) ou de 3SG (*você sabe*). Os dados de *tu* correlacionados às formas verbais de 3SG oferecem-nos resultados qualitativamente mais interessantes. Trata-se de ocorrências que estão restritas ao âmbito estatístico

desta análise, já que foram levantadas cerca de tão somente oito (8) ocorrências do *tu* combinadas às formas verbais de 3SG ("que *tu tenha sido feliz...*", "*tu não está sendo...*"), como ilustramos em (21) e (22). Assim sendo, ao submeter os dados da regra variável *tu/você* a partir do *você* como valor de aplicação, a variável *pessoa* foi o primeiro fator selecionado, mostrando-se, pois, como um contexto promissor ao uso do *você* nas cartas novecentistas analisadas.

(21) "[...] Desejo-te saude e a todos os teus e que *tu tenha* sido feliz na viagem. [...]" (CRJM. Fazenda Tatu-PI, 14.11.1921.)

(22) "[...] Olha! *tu* não *está* sendo como nós eramos porque *está* escondendo as novidades [...]" (JPB. São Bernardo do Campo-SP, 18.04.1978.)

A *pessoa verbal*, ao evidenciar o *p-value* de 3.14e-151 (< 0,05), foi a primeira variável linguística selecionada na interação entre as variáveis. Ao correlacionarmos o índice percentual (61,92%), o peso relativo (0.999) e o *logodd* (103.506), é possível interpretarmos a 3ª pessoa verbal (3SG) como uma variável impulsionadora do *você* nas cartas piauienses em análise. O altíssimo peso relativo no valor de 0.999 (> 0.5) aliado ao *logodd* positivo (103.506) representam as evidências concretas do contexto de 3ª pessoa verbal como propulsor do *você* na amostra de cartas pessoais em análise. Expomos, em (23) e (24), evidências do *você* em relação de concordância com a 3ª pessoa verbal.

(23) "[...] Vão também, mandar os retratos que *você mandou* pedir [...]" (PQ. Oeiras-PI, 12.03.1975.)

(24) "[...] Ante os empecilhos aparecidos, como *você relatou*, eu vinha propor uma sugestão [...]" (BB. RJ, 10.05.1987.)

Nesta análise, a variável *paralelismo formal e semântico* é testada em relação à alternância *tu/você* na escrita novecentista dos piauienses, tendo em vista a interpretação do *paralelismo* como uma "tendência geral de formas gramaticais particulares ocorrerem juntas" e a "sua forma de atuar cria uma harmonia discursiva formal, tornando o discurso mais coeso" (Scherre, 1998, p. 42 *apud* Scherre, 1988, p. 301, 385-6). Acreditamos, à luz de Scherre (1998, p. 42), que "a variável Paralelismo formal mostra que os falantes são compelidos a usar formas semelhantes por algum princípio mental associativo, que pode estar ligado a uma das formas da mente humana operar, refletido no comportamento humano em geral.". Assim sendo, entendemos se tratar de uma variável que pode também influenciar o uso das formas pronominais *tu* e *você*, já que a "mistura pronominal", tão fortemente rechaçada pela tradição gramatical, se dá através do *você* que semanticamente remete ao interlocutor (2ª pessoa), ainda que se harmonize formalmente com a 3ª pessoa. Trazemos, por exemplo, no seguinte trecho de uma carta carioca oitocentista "[8]Agradeço muito a *tua* cartinha [...] ella diz que cada vez gosta mais *de ti* [...] Desejo que esta *te* ache já melhor [...] Foi bom, *Você*

[8] Excerto da carta carioca de João Pedreira do Couto Ferraz, com 60 anos, a sua neta, Maria Leonor, com 06 anos. RJ, 15.04.1886, *cf*. Rumeu (2013, p. 133)

voltar para Santa Fé [...]", evidências da convivência do *você-sujeito* com o *te-complemento*, com o possessivo *tua* e com o sintagma preposicionado *de ti* como formas do paradigma de *tu*, ainda que formalmente destoem das formas do paradigma de *você* (*lhe, seu/sua, de você*).

Temos o *paralelismo formal e semântico* como a segunda variável linguística selecionada, o que está consubstanciado no *p-value* de 2.76e-08, índice menor do que 0,05. No que diz respeito ao número de ocorrências, aos índices percentuais, aos probabilísticos e aos *logodds*, observamos que o *você*-sujeito é favorecido pelos contextos de precedência pelo *você-sujeito* (191/245, 78%, 0.999, 80.714), pelo vocativo (25/245, 10%, 0.993, 4.948), por construções imperativas com formas de subjuntivo (12/245, 5%, 0.999, 107.036) e pelo *tu não-sujeito* (10/245, 4%, 0.985, 4.199), respectivamente. De (25) a (28), trazemos evidências de tais contextos que se mostram impulsionadores do *você* no conjunto das cartas piauienses em análise.

Dado de *você* precedido pelo *você-sujeito*

(25) "[...] Se [você] é amigo do atual prefeito Dr. B. Sá, peça a este uma colaboração naquilo que [você] precise. Contrate logo o padre, cujo nome *você* já fez referência [...]" BB. RJ, 10. 05.1987.

Dado de *você* precedido de vocativo

(26) "[...] [Rodim] pelo que me parece, o que *você* relata nas suas cartas, eu já notei que você está pensando que eu estou usando as minhas cartas como armas de desabafo. Imagino como está magoado e também mudado. Sai Dessa cara! [...]" (JPB. São Bernardo do Campo - SP, 18.04.1978.)

Dado de *você* precedido por construções imperativas com formas de subjuntivo

(27) "[...] Tenho um pouco de pressa na aquisição. [Fale] o tabelião respectivo. Não se trata do Registro de Imóveis. Este eu tenho e de todos os pedaços de posse comprados posteriormente a Março de 1.943. Se existir outras compras lavradas em cartório *você* pode incluir tudo [...]" (ABSB. Rio de Janeiro-RJ, 09.01.1989.)

Dado de *você* precedido pelo *tu não-sujeito*

(28) "[...] Eu gostaria muito, de estar numa praia curtindo [contigo] [...] *Você* **é maravilhosa. [...]"** (VB. Fortaleza - CE, 15.05.1978.)

Como terceira variável selecionada na interação em análise, temos a *expressão do sujeito* cujo valor de aplicação aponta para a sua relevância (*p-value* 0.00154), uma vez que é menor que 0,05, atuando como uma variável dinamizadora da forma *você*. O *você* parece ter a sua produtividade estimulada pelo contexto da sua expressão formal (sujeito pleno), o que está consolidado nos índices de peso relativo (0.999) e de *logodd* (55.105). O fato de o *você* ter a sua produtividade amparada pelo contexto da sua expressão fonética (sujeito pleno) dialoga com a sua origem nominal, através da forma nominal de tratamento *vossa mercê*, que genuinamente é marcada pela sua representação plena.

Para a análise dessa variável linguística, a hipótese é a de que apesar de o *você* assumir os mesmos espaços funcionais que o *tu* (etimologicamente pronome-sujeito), resguarda ainda um resquício da forma nominal de tratamento originária (*vossa mercê*) cuja expressão é essencialmente plena (típico dos nomes), corroborando os estudos de Lopes *et al.* (2009), Rumeu (2013) e Souza (2012) para as missivas cariocas e Souza (2021) para as cartas mineiras. Apresentamos, em (29) e (30), evidências do *você* em contexto de sujeito pleno.

(29) "[...] Primeiro porque foi uma surpresa muito boa que [você] me fez. Pois não imaginavas que *fosses* me escrever mais. por que sempre é o que me *ameaças*. [...]" (JPB. São Bernardo do Campo - SP, 05.06.1978.)

(30) "[...] Só os homens superiores compreendem o dever de ajudar o próximo - e *você*, Mestre, é um deles [...]" (AT. Teresina-PI, 31.01.1987.)

A quarta variável selecionada é a *faixa etária* dos redatores. À luz do *p-value* 0.00154 (menor do que 0,05) entendemos que esse contexto extralinguístico parece se mostrar como um contexto impulsionador do *você*. Em relação à *faixa etária* do redator, observamos que os redatores idosos (236/583, 40,48%, 0.999, 63.161) e adultos (230/583, 39,45%, 0.996, 5.504) mostram-se como propulsores do *você* nas cartas piauienses analisadas, o que se harmoniza às constatações de Souza (2021, p. 117) em relação às cartas mineiras oitocentistas e novecentistas. Os altos pesos relativos 0.999 e 0.996 expõem os idosos e adultos piauienses como redatores propensos ao uso do *você*. De (31) a (34), observamos dados de *você* como expressão da produção escrita de idosos e adultos piauienses.

Dado de *você* por um homem adulto:

(31) "(...) soube que *você* havia ligado. Fiquei aborrecido, pois quase não saio de casa e no momento que o faço acontece de *você* ligar. (...)" (AVV. Brasília, 28.02.1984.)

Dado de *você* por uma mulher adulta:

(32) "[...] Comadre, por aqui nada de importante tudo como sempre [...] pois *você* sabe como estão as coisas. [...]" (JSL. Santo André - SP, 15.02.1964.)

Dado de *você* por homem idoso:

(33) "[...] "[...] Se *você* puder vir e, bem assim algum colega, será uma satisfação para mim. [...]" ABSB (redator com 69 anos). (Fazenda Riacho Fundo, São Pedro, PI, 06.05.1977.)

(34) "[...] Esse almoço do dia 21 de maio corrente, comemorativo do meu natalício, devia realizar-se em Oeiras conforme eu andei tentando preparar, como *você* sabe, eis que em Dezembro do ano passado eu dei alguns passos nesse sentido. [...]" (Fazenda Riacho Fundo, São Pedro, PI, 06.05.1977.)

Ainda no âmbito das variáveis extralinguísticas, observamos a *escolarização* como a quinta (e última) variável selecionada como relevante à produtividade do *você*, o que justifica o *p-value* de 0.00839 desse grupo de fatores, menor do que 0,05, na interação entre as variáveis. Dentre os níveis de

escolarização dos redatores das missivas novecentistas piauienses, temos os que concluíram o ensino médio (119/583, 20,41%, 0.999, 19.071) e o ensino superior (278/583, 47,68%, 0.999, 16.556) como os mais propensos à produtividade do *você* nas cartas em análise. Dentre os redatores com o ensino universitário e o ensino médio, verificamos um altíssimo índice probabilístico (0.999) para ambos os níveis de escolarização. Em termos de *logodds*, observamos que os redatores com o ensino médio (19.071) tendem a se destacar, ainda que tenhamos um menor número de ocorrências (119 oco), se as comparamos à quantidade de dados dos redatores com o nível universitário (278 oco, *logodd* 16.556). Evidenciamos, em (35) e (36), dados da forma *você* produzidos por redatores que concluíram o ensino médio e o ensino superior, respectivamente.

Dado de *você* na produção escrita de redatora com o nível médio:

(35) "[...] Cara prima! [...] Realmente estou ficando velha. *Você* pensou que fosse dia 09/05, né, errou só por 4 dias, faço aniversário é hoje." (MNRSR. Brasília, DF, 05.05.1978.)

Dado de *você* na produção escrita de redator com o nível superior completo:

(36) "[...] Dizem que os homens de inteligência são bons. *Você* confirma a sentença. [...]" (JATF. Teresina - PI, 27.07.1986.)

Com o intuito de entendermos melhor o fato de os redatores idosos e adultos com o ensino superior (graduados) e com o ensino médio terem impulsionado o *você*, passamos ao cruzamento das variáveis *gênero*, *faixa etária* e níveis de *escolarização* dos treze redatores (10 homens e 3 mulheres), cf. a tabela 4.

Tabela 4: As formas *tu* e *você* em relação ao sexo, à faixa etária e ao nível de escolaridade dos redatores

Níveis de escolarização em relação ao gênero e à faixa etária	**Formas de referência ao sujeito de 2SG**	
Homem (10)		
Jovem	**Tu**	**Você**
Alfabetizado	-	**10**/10 (100%)
Ensino médio	9/37 (24,32%)	**28**/37 (75,68%)
Total	9/47 (19,15%)	38/47 (80,85%)
Adulto	**Tu**	**Você**
Alfabetizado	-	7/7 (100%)
Ensino médio	-	**55**/55 (100%)
Ensino universitário	161/167 (96,41%)	6/167 (3,59%)
Total	161/229 (70,31%)	68/229 (29,69%)

	Idoso	**Tu**	**Você**
Alfabetizado		1/125 (0,80%)	124/125 (99,20%)
Ensino universitário		-	111/111 (100%)
Total		1/236 (0,42%)	235/236 (99,58%)
Mulheres (3)			
	Jovem	**Tu**	**Você**
Alfabetizada		**43**/43 (100%)	-
Ensino médio		**19**/27 (70,37%)	8/27 (29,63%)
Total		62/70 (88,57%)	8/70 (11,43%)
	Adulta	**Tu**	**Você**
Alfabetizada		-	**1**/1 (100%)
Total			1/1 (100%)
TOTAL		233/583 (39,97%)	350/583 (60,03%)

Fonte: Barros (2024, p. 101-102)

Ao cruzarmos as variáveis sociais *sexo, faixa etária* e *escolaridade* dos redatores, observamos o uso categórico e prevalente do *você* pelos homens jovens alfabetizados e com o ensino médio a partir dos índices percentuais de 100% (10/10 oco) e 75,68% (28 oco), respectivamente. Em relação à produção escrita dos homens jovens, observamos que a alternância *tu/você* mostrou-se tão somente entre os redatores que tinham o ensino médio, já que os alfabetizados mostraram-se categoricamente voltados ao uso do *você*, ainda que tenhamos levantado um baixo número de ocorrências (10 ocorrências). No contexto dos jovens com o ensino médio, verificamos que o *você* prevalece, em 75,68% dos dados (28/37), ao passo que o *tu* apresenta-se, em tão somente 24,32% (9/37), na produção escrita dos redatores piauienses. O fato de os redatores jovens (homens) mostrarem-se mais propensos ao *você* (80,85%, 38/47) parece ser um indício sintomático de um processo de mudança linguística, já que é esperado que os jovens deem continuidade aos fluxos e contrafluxos da variação e da mudança linguísticas.

Para os homens adultos, a alternância *tu/você* mostrou-se produtiva tão somente entre os redatores com o ensino superior, cabendo ao *tu* uma substancial prevalência, em 96,41% dos dados (161/167). Nesse contexto, a opção pelo *tu* assumida pelos homens com o ensino superior pode ser justificada pelo fato de a adultez mostrar-se como o momento da vida em que se tem um maior nível de comprometimento com a norma-padrão. Nessa fase da vida humana, à luz de Labov (1994), entendemos que há, um período de intenso envolvimento com as atividades profissionais, o que tende a conduzir o informante a uma busca por alcançar exitosos desempenhos intelectual e financeiro. Observamos ainda que, dentre os redatores adultos, os alfabetizados (7/7) e com o ensino médio (55/55) mostraram-se categoricamente voltados ao uso categórico do *você* tão somente, ou seja, não houve dados de *tu* em tais cartas piauienses.

Dentre os homens idosos, temos a prevalência do *você* entre os alfabetizados (124/125, 99,20%). O uso categórico do *você* mostrou-se entre os idosos com o ensino superior (111/111). Independentemente do nível de escolarização dos homens idosos (alfabetizados ou graduados) verificamos o domínio do *você*, o que pode nos evidenciar que, nessa fase da vida menos atrelada aos compromissos das atividades profissionais, é possível evidenciar-se uma expressão linguística mais distanciada dos parâmetros da norma-padrão que prevê o *tu* como pronome pessoal do caso reto para a referência ao sujeito de 2ª pessoa do discurso.

Na produção escrita das mulheres jovens, temos a prevalência do *tu*, em 88,57% dos dados (62 ocorrências) distribuídas entre a redatora alfabetizada que preferiu categoricamente o *tu* (43/43) e a redatora com o ensino médio, em 70,37% dos dados (19/27). Observamos pouquíssimas evidências do *você*, em tão somente 8 ocorrências produzidas por uma única missivista com o ensino médio e atuante, profissionalmente, como funcionária pública no Distrito Federal (MNRSR). Verificamos, na produção escrita da redatora adulta, tão somente o *você*, em uma única ocorrência em carta de JSL, redatora alfabetizada e voltada às atividades "do lar". Esse panorama está diretamente relacionado também ao fato de somente termos tido a oportunidade de trabalhar com a produção escrita de tão somente três missivistas mulheres.

Em síntese, podemos considerar que temos uma amostra de cartas não equanimemente distribuída por todas as *faixas etárias, níveis de escolarização* e *sexo* dos redatores. Isso posto, os resultados quantitativos nos mostram, de um modo geral, a prevalência do *você*, o que está sustentado pela produção escrita dos homens piauienses (10 redatores).

Considerações finais

O foco principal deste estudo é a análise da dinâmica *tu/você* na escrita piauiense novecentista, tendo em vista não só os encaminhamentos metodológicos da SH para a formação de amostras históricas confiáveis, mas também os índices estatísticos e probabilísticos relacionados à regra variável em questão. Voltamo-nos ao "presente para explicar o passado" (Trudgill; Hernández-Campoy, 2007, p. 299 *apud* Bagno, 2017, p. 430), visto que, na atual realidade da fala piauiense, é o *você*, em alternância com o *tu* (subsistema 5, *cf.* Scherre *et al.*, 2021), a forma preferida para a referência ao sujeito de 2SG. Nesse sentido, temos clareza de que os índices estatísticos e probabilísticos evidenciam uma "imagem" do PB novecentista escrito por punhos legitimamente piauienses. Isso posto, passamos às seguintes generalizações sobre a variação *tu/você* conduzidas pelas duas questões que também impulsionaram este estudo. São elas:

(1ª) As formas *tu* e *você* também se mostrariam profícuas na escrita novecentista dos piauienses, considerando o fato de que o *você* prevalece na fala atual do Piauí, ainda que em alternância com o *tu* (com concordância média entre 10% e 39% = subsistema 5), *cf.* Scherre *et al.* (2021, p. 165)?

Confirmamos a hipótese de que o *você* prevaleceria em relação ao *tu* como já observado para a fala atual do Piauí por Scherre *et al.* (2021). Apesar de a alternância *tu* (39,97%) e *você* (60,03%) ter se mostrado intensa, evidenciamos a preferência dos piauienses pelo *você*, anunciando, na produção escrita de sincronias passadas, aquilo que atualmente se revela na fala piauiense, *cf.* Scherre *et al.* (2015, p. 165) à luz das análises embasadas no ALiB (Deus, 2009; Cardoso, 2008).

(2ª) Quais contextos linguísticos e extralinguísticos promoveriam a produtividade da forma *você* na produção novecentista piauiense?

Nas cartas novecentistas analisadas, os contextos de *pessoa verbal*, *paralelismo formal e semântico*, *expressão do sujeito*, *faixa etária* e *nível de escolarização* mostram-se com contextos propulsores do *você*. Passamos às seguintes sistematizações para cada um desses parâmetros selecionados pelo RBrul/Rstudio.

(2a) A combinação do *você* com as formas verbais de 3SG apresenta-se como um contexto promissor a sua produtividade (0.999), ratificando os resultados de Souza (2012) e de Souza (2021) em relação às amostras carioca e mineira, respectivamente, do PB escrito em sincronias passadas (séculos XIX e XX);

(2b) Em relação ao *paralelismo formal e semântico* constatamos que o *você* precedido pelo *você-sujeito* (0.999) mostra-se como um contexto favorável à sua produtividade, acompanhando os resultados de Souza (2021) para as cartas mineiras oitocentistas e novecentistas. Além desse contexto, identificamos que o *você* tende a se mostrar mais produtivo se precedido por formas vocativas (0.993), por construções imperativas com formas de subjuntivo (0.999) e por formas de *tu não-sujeito* (0.985). Temos em evidência contextos formalmente simétricos como os que tendem a conduzir a produtividade da variável dependente (*você*), o que também foi observado por Rumeu (2013) para o *tu-sujeito* das cartas cariocas, por Souza (2012) e por Souza (2021) para o *você* das cartas cariocas e mineiras, respectivamente;

(2c) Dentre as possibilidades de expressão do sujeito (nulo ou pleno), verificamos que o *você* tende a ser favorecido pelo *você* como sujeito pleno. Isso confirma tal traço formal como uma evidência que reverbera a FNT que o originou (*vossa mercê*), cuja tendência típica dos *nomes* é a de expressão plena, o que já está consubstanciado nos resultados de análises embasadas em distintos e diversificados *corpora* históricos (Souza, 2021; Rumeu, 2013; Lopes *et al.*, 2009). Ao testarmos essa variável em relação à alternância *tu/você*, mostram-se mais interessantes as evidências do *tu* combinadas às formas verbais de 3SG que, apesar de terem se mostrado mais rarefeitas nas amostras em análise (*tu não está, tu pensa*), já parecem anunciar o subsistema *tu/você* com concordância média (entre 10% e 39%) atualmente vigente na capital do Piauí (Teresina), *cf.* Scherre *et al.* (2015, p. 141);

(2d) Passemos às sistematizações voltadas aos parâmetros extralinguísticos selecionados como relevantes à proficiência da regra variável em análise. São eles: a *faixa etária* e o *nível de escolarização* dos redatores. Ainda que estejamos conscientes em relação às amostras desarmonizadas, em algum nível, em relação aos parâmetros [9]*sexo, faixa etária* e *nível de escolarização* dos redatores, observamos, no que diz respeito à *faixa etária*, a congruência dos resultados das análises embasadas em cartas mineiras (Souza, 2021) e piauienses (Barros, 2024). Essa constatação está amparada no fato de que, em ambos os trabalhos, temos os redatores idosos (0.999) e adultos (0.999) como os propulsores do *você*, tendo sempre em vista os respectivos conjuntos de cartas pessoais analisadas (cartas mineiras e cartas piauienses). No que se refere ao *grau de escolarização* dos redatores piauienses, identificamos os níveis *universitário* (0.999) e *médio* (0.999) como mais propensos ao uso do *você*. Ao correlacionarmos *sexo, faixa etária* e *escolarização*, evidenciamos que são os homens com o *nível superior* e com o *nível médio* os responsáveis por alavancar o *você* nas cartas piauienses novecentistas em análise.

Em suma, os resultados desta análise acerca da dinâmica *tu/você* apontam para a prevalência do *você*. A produtividade da forma *você* é favorecida não só pelas variáveis linguísticas *pessoa* (*terceira pessoa verbal*), *paralelismo formal e semântico* (precedido por formas de *você-sujeito*) e *expressão do sujeito* (*pleno*), mas também pelas variáveis extralinguísticas *faixa etária* (*redatores idosos e adultos*) e *escolaridade* (*redatores com o ensino superior e com o ensino médio*) a partir de acervos de cartas pessoais, produzidas por punhos genuinamente piauienses na realidade novecentista do português brasileiro.

Essa breve incursão pelos pronomes de 2ª pessoa do singular na era novecentista do PB parece já anunciar o atual sistema *tu/você* com concordância média (**índices de** *tu* abaixo de 60% e concordância de 10% a 39%) no atual espaço piauiense, *cf.* amplamente descrito e estudado por Scherre *et al.* (2015) e Scherre *et al.* (2021).

Referências

BARROS, L. E. de S. *A dinâmica variável tu/você em cartas piauienses novecentistas:* um percurso histórico. 2024, 174 f. Tese (Doutorado em Estudos Linguísticos), Faculdade de Letras – FALE, Universidade Federal de Minas Gerais, Belo Horizonte, 2024. Disponível em: https://hdl.handle.net/1843/73403. Acesso em: 20 jan. 2026.

BATISTA, P. E. S. T. *O uso de Tu e Você em cartas baianas pessoais no século XX em relações de simetria*. Dissertação (Mestrado em Estudos Linguísticos), Universidade Estadual de Feira de Santana, Bahia, 2017. Disponível em: http://tede2.uefs.br:8080/bitstream/tede/598/2/Disserta%C3%A7%C3%A3o%20-%20Depois%20da%20defesa.pdf. Acesso em: 20 jan. 2026.

BERGS, A. Social networks and historical sociolinguistics*:* studies. *In*: *Morphosyntactic variation in the Paston letters (1421-1503)*. Berlim: Walter de Gruyter, 2005.

[9] Nesta análise, temos treze (13) redatores, sendo dez (10) homens e três (3) mulheres. Todos nascidos no espaço geográfico do Piauí.

BROWN, R.; GILMAN, A. The pronouns of power and solidarity. *In*: SEBEOK, T. A. (ed.). *Style in language*. Cambridge: MIT Press, 2003. pp. 253-276. Originalmente publicado em 1960.

BROWN, L.; LEVINSON, S. *Politeness:* some universal in language usage. Cambridge: Cambridge University, 1987.

CARDOSO, S. A. M. Caminho dos pronomes pessoais no português brasileiro: considerações a partir de dados do projeto ALiB. *In*: SIMPÓSIO MUNDIAL DE ESTUDOS DE LÍNGUA PORTUGUESA, 1., 2008, São Paulo. *Anais...* São Paulo: FFLCH USP, 2008. pp. 321-345.

CONDE SILVESTRE, J. C. *Sociolinguística Histórica*. Madrid: Gredos, 2007.

DEUS, V. G. de. *Você ou tu? Nordeste versus Sul:* o tratamento do interlocutor no português do Brasil a partir dos dados do Projeto ALiB. 2009. Dissertação (Mestrado em Linguística) – Universidade Federal da Bahia, Salvador, 2009. Disponível em: http://www.repositorio.ufba.br/ri/handle/ri/8422. Acesso em: 20 jan. 2026.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. *In*: ROBERTS, I.; KATO, M. A. (org.). *Português brasileiro:* uma viagem diacrônica. Campinas: Editora da Unicamp, 1993. pp. 107-128.

GUY, G; ZILLES, A. *Sociolinguística Quantitativa* – instrumental de análise. São Paulo: Parábola Editorial, 2007.

GOMES, V. S.; LOPES, C.R.S. Formas tratamentais em cartas escritas em Pernambuco (1869-1969): tradição discursiva e sociopragmática. *Revista de Estudos da Linguagem*, v. 24, pp. 137-165, 2016. Disponível em: https://periodicos.ufmg.br/index.php/relin/article/view/28405 Acesso em: 20 jan. 2026.

HERNÁNDEZ-CAMPOY, J. M.; CONDE SILVESTRE, J. C. *The Handbook of Historical Sociolinguistics*. [*S. l.*]: 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, pp. 63-79.

LABOV, W. *Principles of Linguistic Change:* Internal Factors. Cambridge: Blackwell Publishers, 1994, v. I.

LOPES, C. R. S. *et al*. Sobre Norma e Tratamento em cartas a Rui Barbosa. *In*: AGUILERA, V. A. (org.). *Para a História do Português Brasileiro*. Londrina: Eduel, 2009. v. VII, pp. 45-92.

LOPES, C. R. S.; CAVALCANTE, S. R. A cronologia do voceamento no português brasileiro: expansão de você-sujeito e retenção do clítico-te. *Linguística*, [*s. l.*], v. 25, pp. 30-65, 2011. Disponível em: https://www.mundoalfal.org/sites/default/files/revista/25_linguistica_030_065.pdf. Acesso em: 20 jan. 2026

LOPES, C. R. S.; MARCOTULIO, L. L.; RUMEU, M. C. B.; COELHO, I. L.; ANDRADE, A.; MARTINS, M. A.; LACERDA, M. O.; GOMES, V. S.; MONTE, V. M.; CARNEIRO, Z. O. N.; SOUZA, C. M. N.; BALSALOBRE, S.; SOUZA, J. P.; OLIVEIRA, T. L.; MOURA, K. K.; CRUZ, I. A.; CARDOSO, N. D. A reorganização do sistema pronominal de 2ª pessoa na história do Português Brasileiro: a posição de sujeito. *In*: LOPES, C. R. S.; CASTILHO, A. T. (org.). *História do Português Brasileiro*: mudança sintática das classes de palavra: perspectiva funcionalista. São Paulo: Contexto, 2018. v. 4, pp. 24-141.

MOURA, K. K.; MARTINS, M. Formas de tratamento em cartas do Rio Grande do Norte. *In*: LOPES, C. R. S.; CASTILHO, A. T. (org.). *História do Português Brasileiro*: mudança sintática das classes de palavra: perspectiva funcionalista. São Paulo: Contexto, 2018. v. 4, pp. 90-97.

OUSHIRO, L. Tratamento de dados com o R para análises sociolinguísticas. *In*: FREITAG, R. M. K. (org.). *Metodologia de coleta e manipulação de dados em sociolinguística*. São Paulo: Blücher, 2014. pp. 133-176.

ROMAINE, Suzanne. *Sócio-Historical Linguistics*: It's Status and Methodology. Cambridge: Cambridge University Press. 1982.

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.; CRUZ, I. A.; CARDOSO, N. D. Formas de tratamento em cartas de Minas Gerais. *In*: CASTILHO, A. T. (org.). *História do português brasileiro*: mudança sintática das classes de palavra – perspectiva funcionalista. São Paulo: Contexto, 2018. v. 4, pp. 42-53.

SCHERRE, M. M. P.; ANDRADE, C. Q.; CATÃO, R. C. Por onde transitam o tu e o você no Nordeste? *Revista de Letras*, v. 40, pp. 164-197, 2021. Disponível em: https://periodicos.ufc.br/revletras/article/view/71460 Acesso em: 20 jan. 2026.

SCHERRE, M.; DIAS, E. P.; ANDRADE, C.; MARTINS, G. F. Variação dos pronomes "tu" e "você". *In*: MARTINS, M. A.; ABRAÇADO, J. *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015, pp.133-172.

SCHERRE, M. M. P. Paralelismo linguístico. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 7, n. 2, pp. 29-59, jul./dez. 1998. Disponível em: https://periodicos.ufmg.br/index.php/relin/article/view/28923. Acesso em: 27 jul. 2026.

SILVA, E. N. *Cartas amorosas de 1930: o tratamento e o perfil sociolinguístico de um casal não-ilustre*. Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

SOUZA, E. Q. *As formas de referência ao sujeito de 2ª pessoa do singular em missivas mineiras dos séculos XIX e XX*: uma análise linguístico-social. 2021. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2021. Disponível em: https://hdl.handle.net/1843/35833. Acesso em: 20 jan. 2026.

SOUZA, J. P. F. *Mapeando a entrada do você no quadro pronominal:* análise de cartas familiares dos séculos XIX-XX. 2012. Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012. Disponível em: https://laborhistorico.letras.ufrj.br/Mestrado/SouzaJPF.pdf. Acesso em 20 jan. 2026.

TRUDGILL, P.; HERNÁNDEZ-CAMPOY, J. M. *Diccionario de sociolingüística*. Madri: Gredos. 2007.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. Empirical foundations for a theory of language change. *In*: LEHMANN, W.; MALKIEL, Y. *Directions for historical linguistics*. University of Texas Press, 2008 [1968].

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72672

Tu, você, nós e a gente: uma análise sociolinguística das variantes pronominais no português falado em São Luís – MA

Tu, você, nós e a gente: a sociolinguistic analysis of pronominal variants in the Portuguese spoken in São Luís – MA

Elimária Oliveira Lima[1]

Letícia Gantzias Abreu[2]

RESUMO

O presente estudo tem como objetivo analisar os processos de variação linguística das variantes pronominais *tu*, *você*, *nós* e *a gente* no português falado em São Luís, Maranhão. Busca-se compreender de que modo essas formas pronominais se distribuem no discurso e quais fatores influenciam suas escolhas pelos falantes. Nesse sentido, a pesquisa é qualitativa, pautada na interpretação das pesquisadoras, e tem como metodologia leitura da bibliografia e análise sociolinguística e comparativa dos dados. Para tanto, o estudo fundamenta-se em aportes teóricos da Sociolinguística Variacionista, que permitem compreender a relação entre uso linguístico, contexto social e significado. A investigação evidencia que, no português maranhense, o emprego dessas variantes é marcado por uma variação sistemática, condicionada por fatores regionais, linguísticos e pragmático-discursivos. Dessa forma, os resultados indicam que as alternâncias *tu/você* e *nós/a gente* na fala ludovicense, embora acompanhem tendências gerais observadas no português brasileiro, apresentam particularidades locais relevantes, contribuindo para a compreensão mais ampla do quadro pronominal do português brasileiro.

PALAVRAS-CHAVE: Variação. Tu e Você. Nós e A gente. Fala ludovicense.

ABSTRACT

This study examines patterns of linguistic variation involving the pronominal variants *tu*, *você*, *nós*, and *a gente* in the Portuguese spoken in São Luís, Maranhão. It investigates how these forms are distributed in discourse and which factors condition speakers' choices. The research adopts a qualitative approach, based on the researchers' interpretations, and employs bibliographic review alongside sociolinguistic and comparative analyses of the data. The study is grounded in the theoretical framework of Variationist Sociolinguistics, which enables an analysis of the relationship between linguistic usage, social context, and meaning. The findings demonstrate that, in Maranhão Portuguese, the use of these variants exhibits systematic variation conditioned by regional, linguistic, and pragmatic-discursive factors. Furthermore, the alternations *tu/você* and *nós/a gente* in the speech of São Luís, while aligning with broader trends observed in Brazilian Portuguese, display locally specific patterns that contribute to a more comprehensive understanding of the pronominal system of Brazilian Portuguese.

KEYWORDS: Variation. Tu and Você. Nós and A gente. Ludovicense speech.

[1] Instituto Federal do Maranhão (IFMA), elimaria.lima@ifma.edu.br, https://orcid.org/0000-0002-9067-6009.

[2] Universidade Estadual do Maranhão (UEMA), leticiaabreu@professor.uema.br, https://orcid.org/0000-0002-7159-0226.

Recebido em 10/02/2026 - Aceito em 10/06/2026

1. Introdução

A reorganização do sistema pronominal do português brasileiro (PB) tem sido amplamente discutida à luz de processos de variação e gramaticalização que afetam, de modo particular, as categorias de pessoa e número. Entre os fenômenos mais recorrentes nesse debate estão os processos de variação dos pronomes *tu*/*você*, para se referir a segunda pessoa do singular, e os pronomes *nós/a gente* na expressão da primeira pessoa do plural. Essas duas variações, embora associadas a pessoas do discurso distintas, compartilham propriedades estruturais e funcionais que desafiam a correspondência tradicional.

A literatura tem mostrado que tanto *você* quanto *a gente* resultam de processos de gramaticalização que implicam a neutralização parcial de marcas morfológicas de pessoa e número, ao mesmo tempo em que preservam traços semânticos e pragmáticos relevantes para a interpretação referencial. Assim, utilizamos como respaldo teórico autores como Alves (2015), Benveniste (1995), Cintra (1972), Ilari (1996), Lopes (1999), Nascentes (1956), Scherre (2015), Labov (2008), Mendonça e Freitag (2018), entre outros estudiosos da temática. Partindo dessa perspectiva, este artigo tem como objetivo analisar e comparar o comportamento das variantes *"tu", "você", "nós" e "a gente"* no falar maranhense, em especial na capital São Luís, buscando evidenciar semelhanças e diferenças, a partir de uma revisão teórica crítica sobre a reorganização do sistema pronominal do português brasileiro, articulando a discussão a dados empíricos já descritos em pesquisas de base sociolinguística. Desse modo, defende-se a hipótese de que essas formas, embora associadas a pessoas do discurso distintas, convergem funcionalmente ao apresentar uma dissociação sistemática entre traços morfológicos e traços semântico-discursivos.

O estudo é de natureza qualitativa, fundamentado em uma abordagem interpretativista, e tem como procedimento metodológico a leitura e análise da bibliografia especializada, aliada à análise sociolinguística de dados descritos em pesquisas anteriores sobre o português maranhense, em especial aqueles provenientes do banco de dados do projeto Atlas Linguístico do Maranhão (ALiMA). A investigação pretende contribuir para a compreensão da reorganização do sistema pronominal do PB, evidenciando como os processos de gramaticalização se refletem em usos atuais da língua falada. Além disso, ao focalizar dados do falar ludovicense (indivíduo nascido em São Luís - MA), o trabalho amplia o escopo empírico das discussões sobre pronomes pessoais no português brasileiro, ressaltando a importância das variedades regionais na descrição e interpretação dos fenômenos linguísticos.

O artigo está estruturado da seguinte maneira: inicialmente, apresenta-se uma revisão teórica e analítica sobre a variação pronominal no Brasil, focando especialmente no caso do *tu* e do *você* em São Luís - MA. Em seguida, desenvolve-se uma revisão teórica e analítica acerca dos processos de gramaticalização e variação das formas *nós* e *a gente* também com foco no contexto ludovicense. Na sequência, realiza-se uma análise sociolinguística do funcionamento das estruturas pronominais de segunda pessoa do singular e de primeira pessoa do plural na língua falada em São Luís, buscando evidenciar padrões de uso, condicionamentos discursivos e possíveis convergências no comportamento

dessas formas. Por fim, são apresentadas as considerações finais, nas quais se discutem as principais contribuições da pesquisa e sua relevância para os estudos sobre o sistema pronominal no português falado no Maranhão e, de modo mais amplo, no português brasileiro.

2. O português falado em São Luís - MA: a variação do "*tu*" e do "*você*"

Segundo Benveniste (1995), os pronomes pessoais na língua revelam a expressão da subjetividade na linguagem e tal subjetividade é devido à natureza desses elementos. Sobre isso, o teórico nos mostra dois princípios característicos: uma correlação de subjetividade que diz respeito à 1ª pessoa em oposição às outras e uma relação de pessoalidade, que corresponde a 1ª e 2ª pessoas em oposição a uma 3ª pessoa. Priorizamos aqui a "correlação de subjetividade", onde o *tu* é tido como uma pessoa não subjetiva ao lado da pessoa subjetiva *eu*. Assim, subjetividade é a capacidade do falante de se admitir como sujeito (*status* linguístico do indivíduo) e a única forma de realização do discurso e da linguagem de maneira geral. Desse modo, o falante se coloca como "sujeito", chamando a si mesmo de *eu* no seu discurso. Além disso, Benveniste (1995) também conceitua instância do discurso como "o ato discreto e único mediante o qual a língua é atualizada em fala" e divide os pronomes em dois tipos diferentes de signos: os que pertencem à sintaxe da língua, uma dimensão sintática (os pronomes relativos, por exemplo), e os que caracterizam instâncias do discurso, uma dimensão pragmática. Neste último, estão presentes os pronomes pessoais que atuam como indicadores de instâncias do discurso, dando ao falante o instrumento para a conversão da língua em discurso, sendo essenciais em uma conversação. Assim, o pronome pessoal não constitui uma classe unitária, pois abarca signos que caracterizam instâncias do discurso e signos que fazem parte da sintaxe da língua.

Nesse preâmbulo, trazendo à tona aspectos históricos, Cintra (1972) revela que a forma arcaica do *você, Vossa Mercê*, era usada para se referir ao rei e à rainha, e entrou em decadência, sendo substituída por, sucessivamente, *Vossa Alteza* e *Vossa Majestade*. A variação *você*, nesse caso, começou a se popularizar no século XVII, chegando, inicialmente, em Portugal, onde, para muitas camadas da população, a expressão *você* era considerada ofensiva (Cintra, 1972, p. 25-36). Os estudos de Nascentes (1956, p.114) nos mostram que a forma *Vossa Mercê* "degradou-se, fonética e semanticamente, a tal ponto que eliminou extraordinariamente a sua forma e, de tratamento real, "pronominalizando-se", chegou a tratamento empregado com inferiores". Nesse âmbito, *Vossa Mercê* sofreu uma simplificação fonética devido a reduções de segmentos e sílabas átonas e propiciou o surgimento de muitas variantes. Ainda segundo Nascentes (1956, p. 119-121), há dezoito registros de formas simplificadas para *Vossa Mercê*. Além do *você*, são variantes também: *Cê, mecê, mincê, ocê, oncê, sucê, suncê, vacê, vainicê, vancê, vansmincê, vassuncê, voncê, vosmecê, vossemecê, vosmincê, vossuncê, ucê*. Entretanto, foi a expressão ***você*** que conseguiu se fixar no PB, tornando-se um pronome de segunda pessoa do singular, por meio do processo "pronominalização" e aceitação de novas formas de tratamento. Sobre isso, Lopes e Duarte (2003) expressam que:

> A gramaticalização de Vossa Mercê > *você* não foi um processo isolado, mas uma consequência de uma mudança encaixada linguística e socialmente. Há uma emergência gradativa de formas nominais de tratamento que passam a substituir o tratamento cortês vós, a partir do século XV em português, num primeiro momento pela ascensão da nobreza e mais tarde da burguesia, que exigia tratamento diferenciado. Essa propagação, que começa de cima para baixo, se dissemina pela comunidade como um todo e as formas perdem sua concepção semântica inicial, gramaticalizando-se – algumas de forma mais acelerada que outras, como é o caso de Vossa Mercê>vosmecê>*você* (Lopes; Duarte, 2003, p. 20).

Por consequência disso, a forma *você* apresentou um comportamento sintático diferente da sua forma original *vossa mercê*. Enquanto esta mostrava maior flexibilidade na frase, a outra começou a ocupar posições mais fixas, de sujeito, mas especificamente a forma começou a ter a função de sujeito pré-verbal (Lopes; Duarte, 2003). Sobre o *você* no PB, Scardua e Yacovenco (2018, p. 175) enfatizam:

> É preciso também considerar que o pronome *você*, advindo de um sintagma nominal, portanto, de terceira pessoa, ocorre com a desinência zero de terceira pessoa. Contudo sua referência é de segunda pessoa do discurso, isto é, relaciona-se à pessoa com quem se fala. Dessa forma, é comum, no PB, o uso de *você* com as formas te e teu, fato que é visto com restrições nos livros didáticos. Entretanto, nas palavras de Lopes (2008), o PB é formado por um quadro pronominal misto que reflete o sincretismo entre os paradigmas dessas duas pessoas do discurso. Em outras palavras, no plano discursivo, *você* é um pronome de 2ª pessoa do discurso, porém, gramaticalmente, comporta-se como um pronome de 3ª pessoa, com a forma verbal, portanto, de 3ª pessoa. Dessa forma, é comum haver uma confluência entre formas de 2ª e 3ª pessoas.

Nesse âmbito, vemos que o *você* (no singular e no plural) está totalmente inserido no quadro pronominal do PB. Entretanto, Lopes (2008, p. 3) enfatiza que:

> *Você* e tu coexistem no singular e vocês é praticamente categórico no plural na posição de sujeito, nas demais posições, contudo, nem o pronome complemento o/a/os/as nem o possessivo vosso se mantiveram produtivos, em seu lugar, se empregam com maior frequência te variando com *você*, lhe e objeto nulo; teu/tua variando com seu/sua, de você(s) e flexões e o uso do imperativo formado a partir do presente do indicativo (imperativo de 2ª pessoa) variando com o de subjuntivo (imperativo de 3ª).

Nessa perspectiva, Scherre *et al.* (2015), ao observarem a variação dos pronomes *tu/você/cê/ocê*, declaram que o *você* é a forma mais utilizada em todo o território brasileiro, sendo presente em várias regiões onde não há uso do *tu*, como na região Centro-Oeste. Após análises de variadas pesquisas, os autores classificaram essa variação em seis subsistemas:

> O *tu/você* com concordância baixa – uso de *tu* abaixo de 60% e concordância abaixo de 10%. O estado do Tocantins e parte dos do Maranhão e de Santa Catarina são os que possuem este subsistema; O *tu/você* com concordância média – uso de *tu* abaixo de 60% e concordância entre 10 e 39%. Este subsistema é característico de muitos estados do Nordeste brasileiro

> (Maranhão, Piauí, Ceará, Paraíba e Pernambuco), parte do Amazonas e de Santa Catarina; O *você/tu* - uso de *tu* entre 1 e 90% sem concordância. Esse subsistema é encontrado no Distrito Federal, em localidades dos estados do Rio de Janeiro, São Paulo, Minas Gerais, Bahia, Maranhão, Roraima e Acre.

Partindo dessas ponderações, voltamos nosso olhar sobre a capital do Maranhão, São Luís, foco deste estudo. Os dados mostram que a variedade linguística ludovicense, quanto ao uso do pronome de segunda pessoa, é um reflexo de raízes histórico-culturais, assim como de recursos linguísticos selecionados pelo falante durante uma interação (Alves, 2015), uma vez que as relações sociais entre os indivíduos e o contexto de fala se refletem na língua. Nesse viés, o estudo realizado por Ramos (1996, p. 11) sobre o emprego dos pronomes no MA mostra indícios de que o "*tu* é mais usado como forma de intimidade entre pessoas próximas, enquanto que o *você* é usado como forma de cortesia e/ou respeito para pessoas mais distantes ou dada a situação formal". Contudo, a autora reforça que:

> No plural essa oposição é anulada, visto que o Maranhão acompanhou a evolução do paradigma flexional e substituiu também o *vós* por *você*s. Em face dessa mudança, o maranhense passou a usar o *você*s tanto como forma de intimidade quanto como forma de respeito e/ou cortesia. Nos casos em que a hierarquia requer um distanciamento mais evidente, a opção passou a ser o uso da forma os senhores/as senhoras.

Tentando também delinear o quadro pronominal falado no MA, Alves (2015) nos apresenta uma distribuição geral das formas pronominais:

Tabela 1: Totais de referências concernentes à segunda pessoa na amostra conjunta[3]

	Tu (sem concordância)	Tu (com concordância)	Você	Cê	Senhor/a
Total	741/1110 66,8%	130/1110 11,7%	157/1110 14,1%	22/1110 2,0%	60/1110 5,4%

Fonte: Alves, 2015, p. 77

Dada a distribuição na tabela, vemos que a utilização do *tu* é expressivamente maior do que o uso do *você*. Logo, o *tu* (sem concordância) é a forma mais utilizada pelos maranhenses para se referir à segunda pessoa do singular. A partir da consulta ao banco de dados[4] do ALiMA (Atlas Linguístico

[3] A amostra de Alves (2015) é constituída por interações livres, as referências à segunda pessoa somam cerca 1110 ocorrências na posição de sujeito, sendo 741 de *tu sem concordância*, 130 de *tu com concordância*, 157 de *você*, 22 de *cê* e 60 de *senhor/a*. Vale ressaltar que na composição geral do *corpus*, foram desconsideradas somente as formas isoladas, seja em trechos inacabados ou repetidos seja na função vocativa, e aquelas formas utilizadas como marcadores fáticos e reguladores (Alves, 2015).

[4] Copiamos trechos, sem alterações, de transcrições grafemáticas do ALiMA para utilizar como exemplo neste capítulo teórico. A transcrições foram realizadas por demais pesquisadores do Projeto, portanto não tivemos nenhuma interferência sobre isso.

do Maranhão)[5], selecionamos dois informantes ludovicenses (que denominamos de Informante A e Informante B) e observamos que ambos usam o *você* em várias situações, principalmente em contextos mais formais ou de menor proximidade do falante, assim como apontam os dados de Alves (2015):

> A forma pronominal *você* surge como a opção de tratamento mais comum em situações em que se deseja evitar uma proximidade inadequada para com o interlocutor. Embora a ocorrência de *você* em nossos dados, como referência definida, gire em torno de 11,7%, podemos dizer que essa frequência é relativamente alta em um sistema que tem o tu como forma predominante.

Dessa maneira, a forma pronominal é motivada pelo grau de intimidade entre o falante e o seu interlocutor. Sobre isso, citamos novamente Cintra (1972), que afirma que o *tu* é uma forma pronominal própria da intimidade e o *você* é uma forma de tratamento nem íntima nem diferencial. O *tu* não é realizado de maneira depreciativa, mas se constitui como uma marca do falar ludovicense utilizada em situações informais e com um grau de proximidade maior do falante. No caso a seguir, o Informante A disse *tu* para se referir ao interlocutor, porém, quando percebeu que usou o pronome *tu*, mudou rapidamente e disse *você*, assinalando um distanciamento com o interlocutor, uma falta de intimidade, atuando assim como uma forma que demonstra respeito:

(1) INFORMANTE A: "Eh... eu achu ingraçadu que... por exemplu, quandu **tu**... **você** falô *lá da questão lá* du.. du *açaí, né?* Purquê... eh... a genti fala... chama muitu, jussara, né?" (Informante A, ludovicense, 50-65 anos, feminino, ensino fundamental)

Já o informante B, quando questionado sobre o emprego dos pronomes, expressou seu posicionamento com o uso das duas formas e considera o *você* uma forma que permite "um tratamento melhor", "mais educada", "mais delicada", que se usa quando se está conhecendo alguém, ao contrário da forma *tu*:

(2) INFORMANTE B – "Por exemplu, quandu eu tô conversandu cum alguém, eu tenhu uma dificuldade di falá **você**, intedeu? Eu não consigu: ah, **você** qué... qué... qué issu? É algu qui im *ôtrus* istadus elis falam muitu mais, quandu, ah **você** pareci uma coisa tão, poxa **tu** ta tendu... sei lá, parece qui **tu** ta tendu um tratamento melhor ainda, sabi? ((risos))"

INQ. – "Te dá uma ideia de distanciamento com relação a outra pessoa, o **você**, ou não?"

INFORMANTE B. – "Eu achou qui da... achu qui sim, di certa forma sim... por que, por exemplu, nu casu eu.. eu... eu tava falandu du **você,** geralmenti a genti usa **você** quando ta conhecendu alguém, qui é uma forma di *sê mais* delicadu, num sei... sê mais educadu, vamos dizê assim... aí... eu... eu, geral... eu particularmenti gostu mais di usá u **tu**, pra qualqué pessoa, intendeu? Mas a pessoa usa muitu **você** pra essa questão di..." (Informante B, ludovicense, 18-30 anos, masculino, ensino superior)

[5] O Atlas Linguístico do Maranhão (ALiMA) é um projeto de pesquisa vinculado ao Departamento de Letras da Universidade Federal do Maranhão (UFMA) que tem o propósito de construir um atlas do falar dos maranhenses. Coordenado pela Professora Doutora do Departamento de Letras, Conceição de Araújo Ramos, o ALiMA está vinculado ao Atlas Linguístico do Brasil (ALiB). O projeto propõe "descrever a realidade do português do Maranhão para identificar fenômenos fonéticos, morfossintáticos, lexicais, semânticos e prosódicos que caracterizam diferenciações ou definem a unidade linguística no Estado." (Ramos; Rocha; Bezerra, 2005, p. 5)

O mesmo informante, quando questionado sobre a diferença entre a utilização do *tu* e do *você* na sua infância em relação ao tempo atual, acredita que em São Luís o uso da palavra *você* aumentou entre os falantes:

(3) INQ. – "E deixa eu te perguntar uma coisa: nós falávamos ainda agora da diferença do tu e o você, tu vês diferença da época de quando tu eras criança, com relação ao tu e **você** para hoje?"

INFORMANTE B. – "é, eu veju diferença..."

INQ. - "Qual a diferença que tu percebes?"

INFORMANTE B. – "eu achu qui as pessoas di São Luís istão usandu bem mais a palavra **você** du que... du que antis, intendeu? Achu qui as pessoas falavam bem... bem menus, intendeu? Achu qui essa é uma... uma diferença... aí aquela coisa, né? Hoji im dia, né? Aqueli circulu di... a circulação di pessoas é maió, intão u contatu é maió... num só im relação à mídia, por exemplu, né? A televisão, achu qui ela tem issu..." (Infomante B, ludovicense, 18-30 anos, masculino, ensino superior)

Desse modo, observamos que o uso do *você* e do *tu (*com e sem concordância) é condicionado pelo papel social do falante e do seu interlocutor e pelo contexto (bem como o ambiente e o tópico discursivo), tal qual a investigação de Alves (2015):

> O sistema de tratamento de São Luís apresenta uma configuração eneária tendo o *você* como uma variante que alterna ao lado do *tu com concordância* em oposição ao *tu sem concordância*. Ao que nos parece, a referencialidade do você pode assumir comportamentos diferentes a depender do contexto situacional. (Alves, 2015, p. 99)

Nesse cenário, citamos Labov (2008), ao expor que uma variante pode dominar conforme a língua, a época, os grupos sociais e o contexto. O autor declara que a mudança linguística é historicamente motivada pelos variados contextos de utilização da língua. Tais contextos viabilizam o aparecimento de diferentes sentidos à "mesma" palavra, mostrando assim que as expressões não são imutáveis. Logo, é o contexto real de uso da língua(gem) que irá determinar a palavra que possui valor para o falante, sendo, portanto, um signo variável e flexível.

Em suma, as análises mostram que *tu* (sem concordância) é amplamente predominante em São Luís como forma de referência à segunda pessoa do singular, superando de modo expressivo o uso de *você*. No entanto, além do forte fator regional, a escolha pronominal do ludovicense mostra-se diretamente relacionada ao grau de intimidade entre falante e interlocutor. O *tu* não assume valor depreciativo, mas funciona como marca identitária associada a contextos informais e de maior proximidade, enquanto o *você* sinaliza distanciamento e menor intimidade com o interlocutor, o que evidencia o uso de *você* como estratégia de respeito.

A seguir, há uma exposição teórica acerca da variação do *nós/a gente*, buscando encontrar semelhanças e diferenças com o processo de variação de *tu/você*. Partimos do seguinte questionamento: à semelhança do que se observa na alternância entre as formas pronominais *tu* e *você*, o uso de *nós* e *a*

gente é condicionado por quais fatores sociolinguísticos? Desse modo, delinearemos de que maneira essas duas formas pronominais se distribuem em São Luís, considerando não apenas sua frequência, mas também os valores sociais e pragmáticos que orientam suas escolhas no discurso.

3. O português falado em São Luís - MA: a variação do "nós" e do "a gente"

De acordo com Omena e Braga (1996) e Lopes (1999), a implementação de *a gente* no quadro pronominal brasileiro teve início entre os séculos XVII e XVIII, a partir do processo de gramaticalização do nome *gente* para o pronome *a gente*. Omena e Braga (1996) apontam que a forma *gente*, originalmente empregada como substantivo coletivo ou como recurso de indeterminação para se referir a um grupo de seres humanos, passou, por extensão de uso, a ocorrer acompanhada do artigo definido *a*, assumindo valor de primeira pessoa do discurso. Nesse processo, observa-se uma mudança de natureza semântica e gramatical, visto que a forma deixa de funcionar como nome para integrar o sistema pronominal do português brasileiro:

> Semanticamente, acrescenta-se ao significado, originalmente indeterminador, a referência à pessoa que fala, deiticamente determinada; gramaticalmente, a forma deixa de ser substantivo e passa a integrar o sistema dos pronomes pessoais, conservando porém com o verbo a mesma relação sintática de terceira pessoa gramatical (Omena; Braga, 1996, p. 189).

A manutenção da relação sintática de terceira pessoa indica uma neutralização das marcas morfológicas e propriedades sintáticas da classe-origem (nome *gente*) e adoção dos atributos da classe-destino (pronome *a gente*). Em outras palavras, *a gente* manteve suas propriedades sintáticas anteriores, desencadeando a concordância de terceira pessoa do singular. Nesse sentido, a partir da análise tradicional, há uma incompatibilidade em relação aos traços morfológicos de pessoa e número, uma vez que, embora semanticamente corresponda à 1ª pessoa do plural, morfologicamente alinha-se à 3ª pessoa do singular.

Para Menon (1996), Lopes (1999) e Zilles (2002), outro aspecto relacionado à forma pronominal *a gente* é que esta passou a se correlacionar a adjetivos no masculino (*cf.* 4) ou no feminino (*cf.* 5), de acordo com o gênero do referente; da mesma forma que o pronome *nós* (*cf.* 6 e 7) e ao contrário da forma original *gente*, que, conforme as autoras, por ser um substantivo feminino, requer modificadores femininos (*cf.* 8), mesmo quando o referente é do gênero masculino (*cf.* 9).

(4) **A gente**, quando **a gente tá junto**, rapá, vambora *lá pra casa do meu cunhado. (*inf. 1, hom, 18 a 30, ens. fund, sls).

(5) Porque eu fui educada assim, respeitando até o meu irmão mais velho, **a gente foi educada** dessa forma, né. (inf. 4, mul, 50 a 65, ens, fund, sls).

(6) Eu vou fazer essa tarefa pra ela, não te preocupa que na outra tarefa **nós vamos fazer juntos**. (inf. 3, hom, 50 a 65, ens. fund, sls).

(7) Aí **nós tamo acustumada** assim, aí tá bom dimais. (inf. 10, mul, 18 a 30, ens, fund, bdc).

(8) **Ela é gente boa**, ela, graças a Deus. (inf. 1, hom, 18 a 30, ens. fund, sls).

(9) **Um pai** de famia **gente** muito **boa** e lá tinha um rapaiz filho **dele** que era fraco do juízo. (inf. 3, hom, 50 a 65, ens. fund, sls).

Para Lopes (2003, p. 28),

> Apesar de o substantivo *gente* ter o traço formal [+ fem], o gênero semântico não é especificado, porque a sua combinação com formas sintáticas de feminino não esclarece o sexo do seu referente, ou seja, a única predicação possível [+fem] não aciona interpretação de gênero, não impondo restrição no que diz respeito ao sexo dos referentes.

Ou seja, para a autora, embora a palavra *gente* tenha uma forma gramatical feminina, seu gênero semântico não é explicitamente definido. O substantivo *gente* pode se referir tanto a homens quanto a mulheres ou a um grupo misto. Assim, usa-se o feminino sem que isso exclua ou defina especificamente o sexo do grupo.

Por outro lado, a partir de consulta realizada na amostra maranhense, foram encontradas também ocorrências do nome *gente* acompanhado de modificadores tanto no feminino (*cf.* 9 e 10) quanto no masculino (*cf.* 10 e 11), evidenciando que, na fala, há contextos em que o termo que acompanha o nome *gente* pode aparecer nos dois gêneros.

(10) Tem **gente medroso** que diz: vi uma alma. Coisa que eu num.... eu num acridito que alma faça. (inf. 11, hom, 50 a 65, ens. fund, bdc).

(11) Rapaiz, **gente novo** num gosta muito de mexê cum **gente idoso** *não.* (inf. 12, mul, 50 a 65, ens. fund, bdc).

Esses dados sugerem que, mesmo após a pronominalização, a forma *gente* não apenas se mantém no português brasileiro, como também passou a ser usada na fala acompanhada de adjetivos no masculino. Por outro lado, parece haver uma restrição quanto ao gênero do modificador a depender da posição em que ele se encontra – anteposto ou posposto ao nome.

Nos excertos (8), (9), (10) e 11), os adjetivos que acompanham o nome *gente* aparecem pospostos ao substantivo e possuem função qualificativa, indicando características de personalidade, temporalidade ou qualidade moral. Mesmo que, à primeira vista, essa combinação (*gente + adj. masc.*) pareça um pouco estranha ou inadequada segundo a gramática normativa, construções no masculino são possíveis na fala espontânea, embora ocorram em número significativamente menor em comparação às construções no feminino. Por outro lado, nos excertos (12) e (13), em que há um adjetivo quantificador anteposto ao substantivo, a substituição por "*muito gente*" soa agramatical.

(12) Que a afirmação que **muita gente** faz nessa face da terra em qualquer lugar, é senso comum puro. (inf. 8, mul, 50 a 65, ens. sup, sls).

(13) Purque eles colocaram o prazo, né, pra fazer a entrega de TCC e aí teve **muita gente** que não conseguiu entregar no prazo. (inf. 14, mul, 18 a 30, ens. sup, bdc).

Dessa forma, os dados obtidos na fala maranhense sugerem que, quando o substantivo *gente* vem acompanhado de adjetivos pospostos, a variação em gênero é possível. Quando vem acompanhado de adjetivos antepostos, a variação em gênero não ocorre. Os dados consultados sugerem também que as ocorrências do substantivo *gente* acompanhado de formas no masculino acontecem com maior frequência entre falantes menos escolarizados, enquanto falantes com maior grau de escolarização o realizam quase sempre acompanhado de formas no feminino.

Outro aspecto apontado pelos estudiosos da gramaticalização de *a gente* diz respeito ao traço indeterminado do substantivo *gente*. Como a forma plural *nós* também permite leituras interpretativas diversas, que vão desde uma determinação precisa, como *eu + tu/você* ou *eu + ele/ela*, até um grau máximo de indeterminação e generalidade, como *eu + todo mundo* ou *eu + qualquer um*, isso pode ter contribuído para que *a gente* adquirisse *status* de pronome.

De acordo com Mendonça (2022), as formas pronominais de primeira pessoa do plural apresentam diferentes possibilidades de interpretação, codificando referentes mais ou menos abrangentes. No entanto, conforme Lopes (1998), há uma diferença no uso das duas formas,

> o falante utiliza preferencialmente o pronome *nós* para se referir a ele mesmo e mais o interlocutor (eu+tu/você), ou a (eu+ele/ela) "não-pessoa" [1]: referente [+perceptível] e [+determinado]. No momento em que o falante amplia a referência, indeterminando-a, há maior favorecimento para a forma *a gente* (Lopes, 1998, p. 419).

Observa-se que em torno das formas de primeira pessoa do plural há a reunião de diferentes pessoas, sempre com a presença do locutor, somado ao interlocutor ou combinado também com a não-pessoa. De um lado, têm-se as formas pronominais associadas a uma afirmação vaga, generalizante. E do outro, associadas a uma junção de pessoas específicas.

Conforme Benveniste (1995), uma distinção importante para a interpretação da primeira pessoa do plural diz respeito à inclusão ou exclusão do interlocutor no referente da forma pronominal. Mendonça (2022) acrescenta que, no português brasileiro, embora os pronomes de primeira pessoa do plural não codifiquem essa distinção semântica morfologicamente, eles apresentam as duas possibilidades de interpretação. Em (14) e (15), o contexto pragmático possibilita a interpretação do traço semântico de *a gente* como inclusivo, visto que as formas pronominais fazem referência ao falante e seu destinatário.

(14) **Nois** tamo aqui numa boa purque tamo só **nós duas**. (inf. 12, mul, 50 a 65, ens. fund, bdc).

(15) Porque num sei se os minino *já foi. Ou* **a gente** deixa pra terminar amanhã? (inf. 13, hom, 18 a 30, ens. sup, bdc).

Em (16) e (17), por outro lado, os pronomes apresentam uma referência específica ao falante mais uma terceira pessoa, com valor semântico exclusivo, isto é, o interlocutor não está incluído no referente da forma pronominal.

(16) Eu tenho… eu tenho… os meus familiares lá da onde eu morava e **tenho um primo** que **nóis** fomo criado junto, né. (inf. 11, mul, 50 a 65, ens. fund, bdc).

(17) Todo dia de noite eu vou e sento no canto um poquinho, bato papo com **um amigo meu** lá que é pagodeiro, aí **a gente** fica conversando. (inf. 7, hom, 50 a 65, ens. sup, sls).

Nos quatro exemplos apresentados, as formas pronominais possuem referência determinada, ou seja, remetem a um grupo específico de falantes, cuja composição pode ser identificada a partir do contexto discursivo e/ou pragmático. Em (14) e (15), os falantes se referem a si próprios e à pessoa com quem estão conversando; em (16), o falante se refere a si e a um primo; e, em (17), o falante se refere a si e a um amigo.

Os traços *inclusivo* e *exclusivo* também são expressos nos usos genéricos das formas de primeira pessoa do plural.

(18) Purque **nós** temos um potencial turístico muito forte **aqui em Barra do Corda**. (inf. 15, hom, 50 a 65, ens. sup, bdc).

(19) E aí, o tema que eu escolhi, a minha banca até falou que foi bem disafiador, purque **aqui em Barra do Corda a gente** *não tem... é... a política de urbanização, né.* (inf. 14, mul, 18 a 30, ens. sup, bdc).

(20) **Lá** mesmo **nós**... **nós** temos a carência do assistente administrativo, que vale pra lá... vai também outro técnico de laboratório na parte de ciências e três professor de administração. (inf. 13, hom, 18 a 30, ens. fund, sup, bdc).

(21) **Lá** na **Vila Esperança** *não tinha ensino médio. Eu fui líder comunitário cinco ano e com ajuda dos meus colega de trabalho* **lá** na **comunidade a gente** conseguiu ensino médio pra lá. (inf. 3, hom, 50 a 65, ens. fund, sls).

Nos exemplos (18) e (19), as ocorrências pronominais codificam genericamente o grupo social de cidadãos barra-cordenses, no qual o falante e o destinatário estão incluídos. Em ambos os casos, a interpretação da forma pronominal é realizada a partir da natureza dêitica da primeira pessoa do plural mais a presença do termo âncora 'Barra do Corda'.

Em (20) e (21), o contexto linguístico atribui o traço semântico exclusivo aos referentes dos pronomes *nós e a gente*, uma vez que o interlocutor não está incluído na referência indicada pelos pronomes. No excerto (20), a forma de primeira pessoa do plural foi usada para fazer referência ao local de trabalho do falante e, em (21), para fazer referência à comunidade (bairro) onde o falante mora. A informação dêitica sobre quem é o interlocutor do entrevistado nesse momento de fala é fundamental para a identificação do traço exclusivo, pois, em (20), o interlocutor não exerce suas atividades laborais no mesmo local de trabalho do entrevistado (Inf. 13) e, em (21), o interlocutor não mora na mesma comunidade que o participante da pesquisa (Inf. 3). Os exemplos mostram que, pelo fato de os pronomes *nós* e a *gente* não serem autorreferenciais, a identificação de sua referência

depende de outra expressão presente no contexto ou de conhecimento pragmático atrelado à situação discursiva.

A partir do exposto, observa-se que os pronomes de primeira pessoa do plural, entre as suas possibilidades de interpretação, apresentam distinções semânticas em relação ao tamanho do grupo referido, com diferentes graus de abrangência. Nesse sentido, as formas *nós* e *a gente* podem expressar diferentes níveis de indeterminação do sujeito, uma vez que os recursos da indeterminação não apresentam um aspecto quantitativo, pois, na indeterminação, o referente, por não ser determinado, não pode ser quantificado. Baseando-se na teoria de Benveniste (1995), Lopes (1998) afirma que, ao englobar diferentes pessoas, um pronome pode tornar-se tão genérico a ponto de não podermos precisar qual é o seu referente.

Conforme Mendonça e Freitag (2018), a indeterminação do sujeito por meio das formas pronominais de primeira pessoa do plural inclui o falante no grupo genérico referenciado, pois embora haja uma generalização, podendo abranger as três pessoas do discurso, há a presença obrigatória da primeira pessoa "eu". Sobre esse aspecto, Silva (2004) acrescenta que as formas *nós* e *a gente* podem abarcar uma variedade de referenciais, específicos e genéricos, elaborados nas negociações entre interlocutores. Para Benveniste (1995), o uso dos pronomes de primeira pessoa do plural vincula uma globalidade indistinta de outras pessoas ao 'eu'. A determinação da natureza do 'não-eu', no contexto, é o que define a especificidade da forma pronominal de 1PP.

Como se observa, na definição da primeira pessoa do plural, de um lado, estão os traços morfossintáticos, que são inerentes às formas pronominais, e, de outro, os traços semânticos do referente à primeira pessoa do plural, que são decorrentes do contexto. De acordo com (Mendonça, 2022), a interação entre as categorias gramaticais de pessoa e número é a base morfossintática para a atribuição dos traços semânticos das formas de primeira pessoa do plural. Combinadas com as pistas do contexto linguístico, discursivo e pragmático, as noções semânticas de pessoa e número possibilitam a interpretação dos pronomes de primeira pessoa do plural em função da abrangência do referente e da inclusão do interlocutor.

Uma outra possibilidade de interpretação para os pronomes *nós* e *a gente* diz respeito à autorreferência, isto é, a partir de uma 'alteração' na categoria gramatical de número, as formas de primeira pessoa do plural podem ser utilizadas para se referir unicamente ao próprio falante. Ao usar a autorreferência, o falante desvia a atenção do eu, sugerindo uma autoria conjunta. Ilari, Franchi e Neves (1996, p. 88), com base em Benveniste, classificam esse uso como "um tipo de emprego em que um indivíduo institui sua fala como a de um grupo, mas nele não inclui nem a segunda nem a terceira pessoa (plural de modéstia)".

No excerto (22), o falante usa a forma *a gente* para fazer uma autorreferência. O referente da forma de primeira pessoa do plural é identificado a partir do assunto abordado pelo falante, isto é, o seu trabalho.

(22) INQ: Como é o lugar onde você trabalha? Fale um pouco sobre esse lugar.

INF: Tem veiz que **a gente** faiz... mais é... não... não posso dizê que trabalho nisso aí, purque na verdade não tenho profissão nisso aí, né. As veiz, quando alguma pessoa, um vizin pricisa, **a gente** vai levá e as pessoa as veiz agrada **a gente**, gratifica, né. (inf. 11, hom, 50 a 65, ens. fund, fund, bdc).

O contexto em que o pronome *a gente* é empregado ilustra como elementos gramaticais e semânticos interagem com as condições pragmáticas e discursivas. Apesar de sua natureza gramatical como uma expressão de primeira pessoa do plural, o uso de *a gente* no excerto denota uma singularidade referencial, na medida em que o falante não inclui outros indivíduos no grupo enunciado. Esse uso pode ser interpretado como uma estratégia discursiva que atenua a responsabilidade pessoal do falante, conferindo um tom mais coletivo à experiência individual relatada.

A partir das discussões levantadas e dos exemplos apresentados, os pronomes de primeira pessoa parecem intercambiáveis nos diferentes contextos; no entanto, nem sempre é o caso. *A gente* contrasta com *nós*, por exemplo, em não poder ser modificado por numerais ou nominais (*cf.* 23 e 24). Nesses casos, observa-se que os traços semânticos da forma-fonte (*gente*) foram preservados na forma gramaticalizada (*a gente*), motivando, para o uso da forma gramaticalizada, restrições sintáticas. De acordo com Omena e Braga (1996), a persistência do traço indeterminador na forma *a gente* é o que provoca certas restrições em seu uso.

(23) Acho que é purisso que **nós dois** *dá certo*. (inf. 10, mul, 18 a 30, ens. fund, fund, bdc).

* Acho que é purisso que **a gente dois** [dá/damos] certo.

(24) **Nós brasileiros** adoramos tango. (Kato; Martins; Nunes, 2023, p. 35).

* **A gente brasileiros** [adora/adoramos] tango.

Os exemplos ilustram uma diferença importante entre os pronomes de primeira pessoa do plural no português brasileiro, que está relacionada às possibilidades de modificação e uso em contextos específicos. O pronome *nós* pode ser modificado por numerais, como em "*nós* dois", ou por nominais, como em "*nós* brasileiros". Já *a gente*, embora funcione como um pronome na prática, mantém uma estrutura mais próxima de um sintagma nominal (como "a pessoa" ou "o grupo") e, por isso, não aceita modificadores dessa natureza. Omena e Braga (1996) observam ainda que elementos como *"nenhum", "cada um"* e "todos", por exemplo, podem acompanhar *nós*, mas não são compatíveis com *a gente*.

Nesse contexto, evidencia-se a afirmação de Lopes (2004, p. 52) de que "nem todas as propriedades formais do nome *gente* foram perdidas, assim como não foram assumidas todas as propriedades intrínsecas aos pronomes pessoais". A coexistência entre novos valores/usos ao lado dos antigos e a permanência das propriedades lexicais na forma gramaticalizada confirmam o princípio da *persistência* de Hopper (1991, p. 124), o qual sustenta que "alguns traços do significado lexical original de um item tendem a aderir à nova forma gramatical e detalhes de sua história lexical podem refletir-se na sua distribuição gramatical".

3. As variações Tu/Você e Nós/A Gente em São Luís-MA: semelhanças e diferenças

Do ponto de vista histórico, tanto *você* quanto *a gente* resultam da gramaticalização de expressões nominais – *vossa mercê* e *gente*, respectivamente – que, ao longo do tempo, passaram a integrar o sistema pronominal do português. Esse processo implicou a perda gradual de propriedades lexicais e a incorporação de funções gramaticais, culminando na consolidação dessas formas como pronomes de segunda pessoa do singular e de primeira pessoa do plural.

Quadro 1: Paralelos funcionais entre *você* e *a gente* no português brasileiro

Critério	**Você**	**A gente**
Origem histórica	Gramaticalização de *vossa mercê*	Gramaticalização do sintagma nominal *a gente*
Pessoa gramatical	3ª pessoa do singular	3ª pessoa do singular
Pessoa do discurso	Referência à 2ª pessoa do singular	Referência à 1ª pessoa do plural
Concordância verbal	Predominantemente na 3ª pessoa	Predominantemente na 3ª pessoa
Dissociação morfologia × referência	Sim	Sim

Fonte: Elaboração dos autores

Os paralelos apresentados no quadro evidenciam que *você* e *a gente*, embora atuem em domínios distintos do sistema pronominal, compartilham trajetórias funcionais semelhantes no português brasileiro, especialmente no que se refere aos processos de gramaticalização e à consolidação como formas pronominais no uso. Essas convergências não implicam uma equivalência plena entre os fenômenos, mas permitem compreender que ambos participam de um mesmo movimento de reorganização do sistema pronominal. A partir dessa perspectiva, torna-se possível avançar na análise considerando não apenas as semelhanças, mas também os limites e as especificidades que caracterizam o funcionamento dessas formas na variedade falada em São Luís.

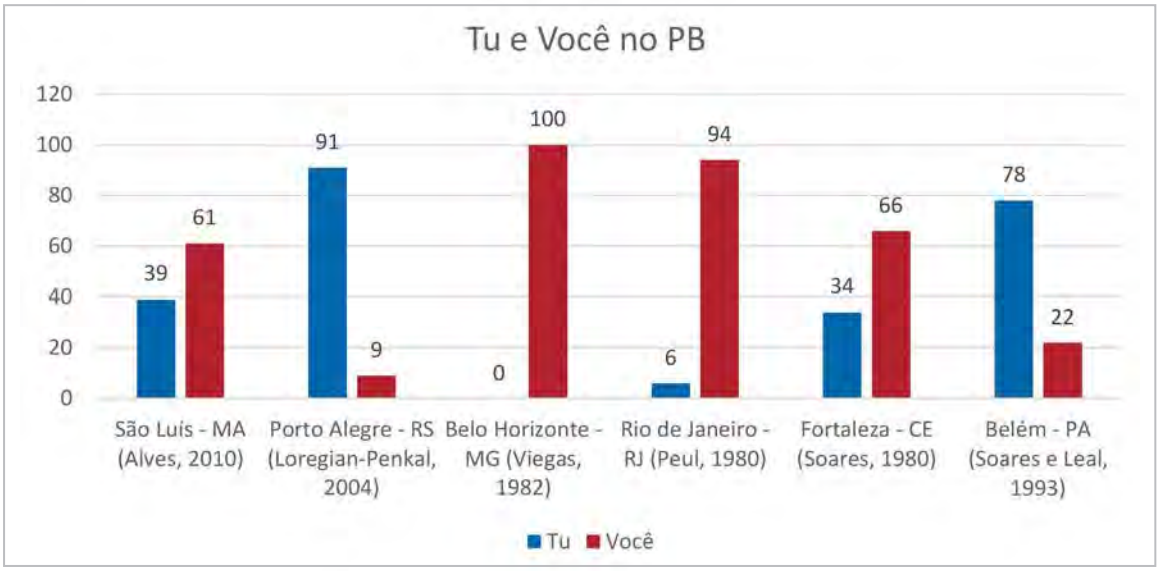
É nesse quadro mais amplo de reorganização do sistema pronominal que se inscrevem as particularidades observadas no falar ludovicense. Nesse contexto, o processo se manifesta de modo coerente com tendências gerais do português brasileiro, mas apresenta especificidades relacionadas às práticas interacionais locais, perceptíveis na maior ou menor produtividade dos pronomes inovadores e na forma como essas variantes se articulam aos usos descritos para outras regiões do país. Essas particularidades, no entanto, não podem ser analisadas de forma dissociada das representações sociais e dos discursos metalinguísticos historicamente construídos sobre a língua em São Luís.

Entre essas representações, destaca-se uma singularidade que até hoje integra o senso comum maranhense e a história do estado: a ideia de que "o lugar onde melhor se fala português no Brasil é o Maranhão". Sobre essa afirmação, Bagno (1999, p. 71) afirma que não há nenhuma fundamentação científica que a sustente, afirmando que "esse mito nasceu da velha posição de subserviência em relação ao português de Portugal". O autor atribui a disseminação desse mito à regularidade com que se usa o pronome *tu* em São Luís – MA, seguido das formas verbais características de segunda

pessoa, como *tu vais, tu queres, tu dizes, tu comias, tu cantavas etc.*, enquanto na maior parte do Brasil o pronome *tu* foi substituído por você. Bagno (1999) acrescenta que quando empregado em outras variedades, o verbo junto ao pronome *tu* assume a forma de terceira pessoa, como *tu vai, tu fica, tu quer, etc.*

Sobre a variação pronominal do *tu* e *você* no PB, montamos o gráfico abaixo, conforme dados expostos na obra *Mapeamento Sociolinguístico do Português Brasileiro* (2015), dos autores Jussara Abraçado e Marco Antonio Martins. O objetivo é comparar a situação de São Luís com outras capitais brasileiras:

Gráfico 1: Quadro comparativo da variação Tu/Você em São Luís – MA e outras capitais brasileiras



Fonte: Elaboração dos autores a partir de Abraçado e Martins (2015)

O gráfico evidencia uma variação regional significativa no uso dos pronomes pessoais *tu* e *você* nas diferentes regiões do Brasil, confirmando que essa escolha pronominal no português brasileiro é fortemente condicionada por fatores geográficos. Em Porto Alegre (RS), observa-se um uso predominante de *tu* (91%), o que reflete um traço identitário marcante da região Sul, frequentemente empregado em contextos informais e formais, ainda que muitas vezes com concordância verbal variável. Já nas cidades do Sudeste, o uso de *você* é praticamente categórico, pois Belo Horizonte (MG) apresenta 100% de *você*, enquanto o Rio de Janeiro (RJ) registra apenas 6% de *tu*, confirmando a consolidação de *você* como forma dominante nas regiões Centro-Oeste e Sudeste.

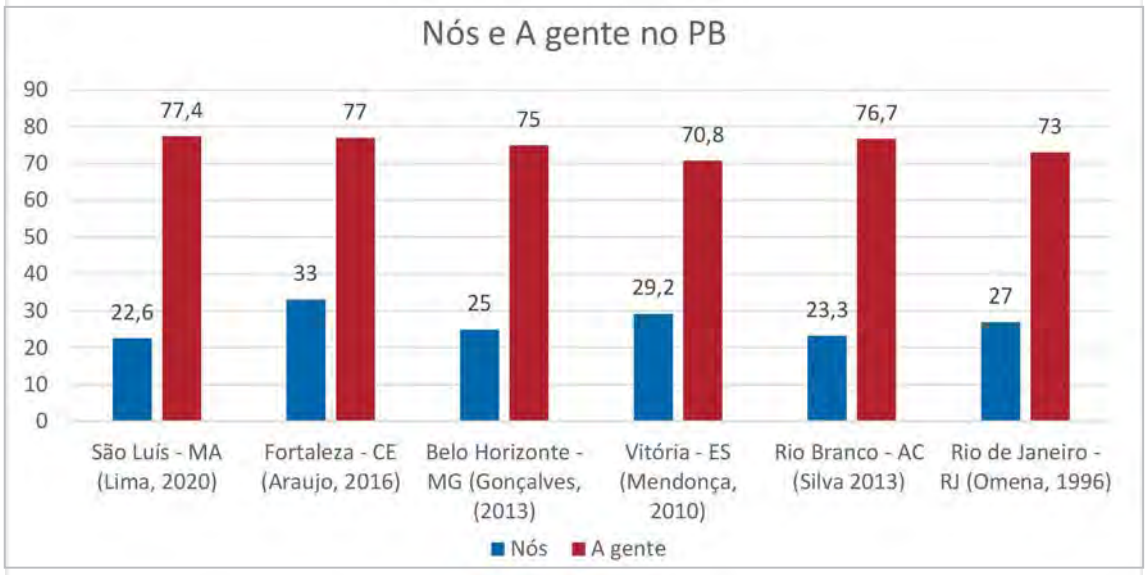
Quando analisamos Belém (PA), o *tu* também aparece de forma expressiva (78%), indicando que o pronome mantém forte presença no Norte do país, sobretudo na fala cotidiana. Diferentemente do Sul, nessa região o *tu* costuma ocorrer majoritariamente sem concordância verbal padrão (91%). No caso de São Luís, foco deste artigo, percebe-se um predomínio do pronome *você* (61%), embora o *tu* ainda tenha percentuais relevantes (39%), o que também acontece de forma semelhante em Fortaleza – CE. Tais dados sugerem um cenário de coexistência pronominal, no qual ambos os

pronomes circulam, possivelmente associados a fatores como região, idade, grau de intimidade e contexto comunicativo. Alves (2010) revela que 71,1% das ocorrências de segunda pessoa na fala maranhense são acompanhadas da forma verbal de terceira pessoa, sendo a concordância regulada principalmente pela escolaridade. Conforme a autora, a forma pronominal *tu* com concordância é favorecida entre os falantes mais escolarizados, ao passo que falantes menos escolarizados favorecem o uso do *tu* sem concordância.

A localização de São Luís, capital do Maranhão (transição entre o Norte e Nordeste do Brasil), contribui para a circulação de padrões linguísticos distintos, o que ajuda a explicar a presença significativa do *tu*, tradicionalmente mais frequente nas regiões Norte e Sul. Ao mesmo tempo, o predomínio de *você* reflete um padrão difundido no PB contemporâneo, reforçado pela escola, pela mídia e por normas de prestígio linguístico. Fatores sociais, como idade e escolaridade, também influenciam essa distribuição, assim como grau de intimidade e contexto comunicativo, já que os falantes mais jovens tendem a empregar mais o *tu* em contextos informais, enquanto o *você* é frequentemente associado a situações mais formais. Assim, a variação pronominal de segunda pessoa do singular em São Luís nos revela a interação entre fatores regionais, sociais e contextuais, mantendo o *tu* como uma forma significativa, embora o *você* se imponha como a escolha majoritária.

Dessa maneira, analisamos a variação pronominal *você* e *tu* em relação às formas *nós* e *a gente* no português falado em São Luís, à luz dos processos de gramaticalização, da variação linguística e dos traços semântico-pragmáticos que orientam seu uso. Assim como na análise dos pronomes *tu* e *você*, consideramos apenas estudos realizados em capitais, em razão da proximidade de características relacionadas ao grau de mobilidade dos falantes, à diversidade sociocultural e ao maior contato entre variedades linguísticas, fatores que favorecem a emergência e a difusão de formas inovadoras no sistema pronominal (Lima, 2025).

Gráfico 2: Quadro comparativo da variação *Nós/A gente* em São Luís - MA e outras capitais brasileiras



Fonte: Elaboração dos autores a partir de Lima (2025)

A partir do gráfico, observa-se que a variação das formas *nós* e *a gente* ocorre de maneira sistemática e não aleatória, com comportamentos semelhantes no uso dos pronomes nas diferentes capitais. Em todas as pesquisas analisadas, o pronome inovador suplanta o pronome mais antigo, o que corrobora a hipótese de que *a gente* se encontra em estágio avançado de difusão no português brasileiro. O *nós* apresenta percentuais entre 20% e 30%, enquanto *a gente* varia entre 70% e 80%.

Apesar desse panorama indicar uma tendência consistente de expansão do pronome inovador, a distribuição das variantes não é uniforme. Lima (2025), ao observar a variação de *nós* e *a gente* em diferentes variedades do português brasileiro, observou a existência de estudos em que o *nós* apresenta maior produtividade do que *a gente*. Esses resultados, contudo, mostram-se fortemente condicionados pelo período de coleta dos dados, pela localidade e pela modalidade da amostra. Isso implica dizer que: estudos com amostras mais recentes tendem a registrar maior uso de *a gente* do que pesquisas com amostras mais antigas; amostras urbanas apresentam percentuais mais elevados do pronome inovador em comparação às rurais; e dados de fala revelam maior produtividade de *a gente* do que dados de escrita.

Esses condicionamentos empíricos permitem avançar para uma interpretação sociolinguística dos resultados. Nesse sentido, entende-se que pelo fato de o pronome *a gente* não ser estigmatizado, ele tenha se tornado cada vez mais recorrente na fala, inclusive entre os falantes considerados cultos. No que diz respeito à modalidade, é oportuno apontar que a oralidade é um domínio no qual a forma inovadora encontra terreno fértil para difusão, tendo em vista sua ampla aceitabilidade entre os falantes de modo geral. Por outro lado, a maior frequência de *nós* na escrita revela a preocupação com a norma padrão e com a adequação linguística frente às convenções de escrita constantemente reforçadas no contexto escolar. Em se tratando da localidade, o isolamento geográfico tende a favorecer padrões mais conservadores, ao passo que contextos marcados por maior mobilidade e contato social propiciam um estágio mais avançado de difusão da forma inovadora.

Quanto à concordância verbal com as formas pronominais de 1PP, existe uma grande diferença entre a realização da concordância verbal com a forma pronominal *nós* e com a forma *a gente*: enquanto este último quase sempre satisfaz às regras de concordância padrão, o primeiro realiza-se predominantemente sem obedecer a esse fator. Em termos percentuais, nas diferentes variedades do português brasileiro, de acordo com Lima (2025), a concordância verbal com o pronome *a gente* fica acima de 95%. Por outro lado, a concordância com o pronome *nós* apresenta grandes oscilações, sendo condicionada principalmente pelo fator escolaridade.

Essas diferenças no comportamento morfossintático das formas pronominais contribuem para compreender padrões mais amplos de difusão e uso no português brasileiro. Embora pertençam a diferentes categorias de pessoa gramatical (segunda pessoa do singular e primeira pessoa do plural), observa-se que *você* e *nós* apresentam comportamentos semelhantes no que diz respeito à sua circulação em contextos mais formais e ao seu condicionamento por fatores sociais, como o grau de formalidade e o nível de monitoramento da fala. Ambas as formas tendem a funcionar como estratégias para

evitar maior intimidade entre os interlocutores, sendo recorrentes em situações comunicativas mais formais. No entanto, a análise dos gráficos revela que *a gente* se sobrepõe amplamente a *nós* no PB, configurando-se como a variante mais frequente da 1ª pessoa do plural na fala em todas as regiões, ao passo que *nós* se concentra em contextos mais formais e monitorados. Tal comportamento, contudo, não se reproduz de maneira equivalente na variação *tu/você*: *você* é majoritário na maioria das cidades, especialmente nas regiões Sudeste e Centro-Oeste, enquanto *tu* se mantém mais produtivo em áreas como o Sul, o Norte e o Nordeste, coexistindo, no caso de São Luís, com a forma inovadora. Assim, é possível sistematizar alguns apontamentos sobre a variação do *tu/você* e *nós/a gente* em São Luís - MA, conforme a tabela a seguir:

Quadro 2: Semelhanças e diferenças entre as variações *Tu/Você* e *Nós/A gente* em São Luís

Eixo analítico	Tu/Você	Nós/A Gente
Configuração da variação	Coexistência entre forma inovadora (*você*) e forma tradicional (*tu*)	Coexistência entre forma inovadora (*a gente*) e forma tradicional (*nós*)
Estágio de difusão	Difusão variável, com forte marcação regional; coexistência produtiva em São Luís	Difusão ampla no PB; *a gente* em estágio avançado de expansão
Sensibilidade sociopragmática	Sensível a grau de formalidade, monitoramento da fala e relações interlocutivas	Menor sensibilidade social; uso recorrente de *a gente* inclusive entre falantes cultos
Implicações morfossintáticas	Oscilações na concordância verbal no uso de *tu*, enquanto *você* se realiza quase categoricamente com concordância de terceira pessoa	Oscilações na concordância verbal no uso de *nós*, enquanto *a gente* se realiza quase categoricamente com concordância de terceira pessoa

Fonte: Elaboração dos autores

A análise das variações *Tu/Você* e *Nós/A gente* no português falado em São Luís evidencia que, embora pertençam a pessoas do discurso distintas, esses pares pronominais apresentam comportamentos sociolinguísticos que permitem uma leitura articulada de seu funcionamento no sistema pronominal. Em ambos os casos, observa-se a coexistência entre formas canônicas e formas inovadoras que resultaram de processos de gramaticalização e se consolidaram como formas pronominais.

Nesse âmbito, em São Luís, vemos que as formas menos marcadas (*você* e *a gente*) apresentam ampla expansão no português brasileiro, enquanto as formas tradicionais (*tu* e *nós*) permanecem produtivas, porém condicionadas por fatores mais específicos, como a marcação regional, o grau de formalidade e o nível de monitoramento da fala. A diferença mais notória é que o *tu* ainda mantém força regional significativa, ao passo que *nós* apresenta retração mais uniforme no território nacional, reforçando a assimetria entre os dois processos de variação.

Considerações finais

Este estudo contribui para a compreensão da reorganização do sistema pronominal do PB, especialmente no falar maranhense, ao evidenciar como os processos de gramaticalização se refletem nos usos atuais da língua falada. Para tanto, adotou-se uma pesquisa de natureza qualitativa, de abordagem interpretativista, tendo como procedimento metodológico a leitura e análise de bibliografia especializada, como Cintra (1972), Ilari (1996), Scherre *et al.* (2015), Labov (2008), Lima (2025), Lopes (2003), entre outros, aliada à análise sociolinguística de dados apresentados em estudos anteriores sobre o português maranhense, especialmente aqueles oriundos do banco de dados do projeto Atlas Linguístico do Maranhão.

A pesquisa alcançou seu objetivo de analisar e comparar o comportamento das variantes *tu*, *você, nós* e *a gente* no falar maranhense, em especial na capital São Luís, buscando evidenciar semelhanças e diferenças. Como resultado, observamos que os processos de variação na segunda pessoa do singular e na primeira pessoa do plural revelam um comportamento semelhante: as formas inovadoras (*você* e *a gente*), oriundas da gramaticalização de expressões nominais (*vossa mercê* e *gente*), consolidam-se progressivamente no uso sem eliminar completamente as formas mais antigas (*tu* e *nós*). Esses processos, por sua vez, não se restringem à variação das formas pronominais, mas envolvem também reconfigurações no plano morfossintático, uma vez que os próprios pronomes canônicos passam a apresentar variação nos padrões de concordância verbal, como se observa no uso recorrente de *tu* e *nós* com formas verbais de terceira pessoa do singular.

De modo geral, por meio das análises, os dados evidenciam que, em São Luís – MA, o uso dos pronomes *tu/você* e *nós/a gente* é marcado por variação sistemática, condicionada principalmente por fatores regionais, sociais e estilísticos. O *tu*, fortemente condicionado pelo fator geográfico, coexiste com a forma *você*, configurando uma alternância que reflete o grau de intimidade e formalidade da interação. Já com relação à variação *nós/a gente,* observa-se a predominância de *a gente* na fala oral, enquanto *nós* se realiza, majoritariamente, em contextos mais formais. Esses resultados permitem inferir que a escolha pronominal no português ludovicense responde a critérios como marcação regional, grau de formalidade e naturalidade da situação comunicativa.

No âmbito do ensino, vale salientar que ambas as variações, *tu/você* e *nós/a gente*, ainda são pouco exploradas nos livros didáticos utilizados nas escolas brasileiras. Esses materiais tendem a privilegiar os pronomes *tu* e *nós*, em conformidade com a gramática normativa, frequentemente invisibilizando formas amplamente utilizadas pelos falantes. Tal postura contrasta com o reconhecimento da variação linguística por documentos oficiais, como a Base Nacional Comum Curricular (BNCC), e acaba por reforçar um ensino de língua que pouco dialoga com os usos reais e com a diversidade linguística do país.

Nesse sentido, a comparação entre as variantes pronominais apresentada neste artigo pode ser incorporada às aulas de Língua Portuguesa, contribuindo para que o aluno compreenda o funcionamento efetivo da língua e desenvolva uma postura crítica diante da norma padrão. Logo, o professor deve

promover discussões acerca dos diferentes contextos de uso das variantes, ressaltando que sua seleção está vinculada a fatores sociais, regionais e discursivos. O reconhecimento da legitimidade dessas formas no uso cotidiano favorece uma aprendizagem mais significativa, contribui para a superação de preconceitos linguísticos e amplia a competência comunicativa dos estudantes, preparando-os para circular conscientemente entre distintos registros e situações de comunicação.

Referências

ALVES, C. C. B. *O uso do tu e do você no português falado no Maranhão*. 2010. 144 f. Dissertação (Mestrado em Linguística) – Universidade Federal do Ceará, Fortaleza, 2010.

ALVES, C. C. B. *Pronomes de segunda pessoa no espaço maranhense*. 2015. 153 f. Tese (Doutorado em Linguística) – Programa de Pós-graduação em Linguística – PPGL), Universidade de Brasília, Brasília, 2015.

BAGNO, M. *Preconceito linguístico*: o que é, como se faz. São Paulo: Loyola, 1999.

BENVENISTE, É. A Natureza dos pronomes. *In*: BENVENISTE, É. *Problemas de Linguística geral I*. 4. ed. Campinas, SP: Editora Pontes, 1995.

CINTRA, L. F. L. *Sobre formas de tratamento na língua portuguesa*: ensaios. Lisboa: Horizonte, 1972.

HOPPER, P. J. On some principles of grammaticization. *In*: TRAUGOTT, E. C.; HEINE, B. (eds.): *Approaches to grammaticalization,* Volume I, Amsterdam/Philadelphia: John Benjamins Company, 1991.

ILARI, R.; FRANCHI, C.; NEVES, M. H. M. Os pronomes pessoais do português falado: roteiro para a análise. *In*: CASTILHO, A. T. de; BASÍLIO, M. (org.). *Gramática do português falado*: estudos descritivos. Campinas: Editora da Unicamp, 1996. v. 4, pp. 79-168.

KATO, M. A.; MARTINS, A. M.; NUNES, J. *Português brasileiro e português europeu*: sintaxe comparada. São Paulo, SP: Contexto, 2023.

LABOV, W. *Padrões sociolinguísticos*. Tradução Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. -São Paulo, Parábola Editorial, 2008.

LIMA, E. O. *Uso variável da concordância verbal de primeira pessoa do plural no Maranhão*. 2025. 228 f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2025.

LOPES, C. R. S. Nós e a gente no português falado culto do Brasil. Delta, v. 14, n. 2, pp. 405- 422, 1998.

LOPES, C. R. *A inserção de a gente no quadro pronominal do português*. 1999. 174 f. Tese (Doutorado em Letras) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999.

LOPES, C. R. dos S. *A inserção de ´a gente´ no quadro pronominal do português*. Frankfurt am Main/Madrid: Vervuert/Iberoamericana, 2003, v.18. p. 174.

LOPES, C. R. dos S. A gramaticalização de a gente em português em tempo real de longa e de curta duração: retenção e mudança na especificação dos traços intrínsecos. *Fórum Linguístico*: Florianópolis, v. 4, n. 1, pp.47-80, julho de 2004.

LOPES, C. R. S. Retratos da variação entre você e tu no português do Brasil: sincronia e diacronia. *In*: RONCARATI, C.; ABRAÇADO, J. (org.). *Português brasileiro II*: contato linguístico, heterogeneidade e história. Niterói: EDUFF, v. 2, 2008, pp. 55-57.

LOPES, C. R. S.; DUARTE, M. E. L. De *vossa mercê* a *cê*: os processos de uma mudança em curso pronominalização de nominais em peças brasileiras e portuguesas setecentistas e oitocentistas. *In*: BRANDÃO, S. F.; MOTA, M. A. (org.). *Análise contrastiva de variedades do português*: primeiros estudos. Rio de Janeiro: *In-Fólio*, 2003, pp. 61-76.

MENDONÇA, J. de J. O controle dos traços semânticos de nós e a gente em estudos variacionistas. *Carderno de Estudos Linguísticos*. v. 64, pp. 1-16, 2022.

MENDONÇA, J. J.; FREITAG, R. M. K. Indeterminação do sujeito, primeira pessoa do plural e gradação referencial. *In*: CARVALHO, D.; BRITO, D. (org.). *Pronomes*: morfossintaxe e semântica. 1. ed. Salvador: EDUFBA, 2018. pp. 235-258.

MENON, O. P. da S. A gente: um processo de gramaticalização. *Estudos Linguísticos, São Paulo, v. 25*, pp. 622-628, 1996.

NASCENTES, A. O tratamento de você no Brasil. *In*: *Letras*. Curitiba/PR: Ed. UFPR, v. 6, n. 5, pp. 114-122, 1956.

OMENA, N. P.; BRAGA, M. L. A gente está se gramaticalizando?. *In*: MACEDO, A. T.; RONCARATI, C.; MOLLICA, M. C. (org.). *Variação e Discurso*. Rio de Janeiro: Tempo Brasileiro, 1996, pp. 75-84.

RAMOS, C. de M. de A. *O Português falado em São Luís*: os pronomes pessoais na posição de sujeito. São Luís, 1996. (mimeo).

RAMOS, C. de M. de A.; ROCHA, M. de F. S.; BEZERRA, J. R. M. *O português falado no Maranhão: estudos preliminares*. São Luís: EDUFMA, 2005.

SCHERRE, M.; DIAS, E. P.; ANDRADE, C.; MARTINS, F. G. Variação dos pronomes tu e você. *In*: MARTINS, M. A.; ABRAÇADO, J. *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015, pp. 133-172.

SCARDUA, J. R.; YACOVENCO, L. C. A variação pronominal de segunda pessoa: contribuições da sociolinguística para o ensino de língua portuguesa. *Working Papers em Linguística*, Florianópolis, v. 18, n. 2, pp. 171-191, jan. 2018. ISSN1984-8420. Disponível em: https://periodicos.ufsc.br/index.php/workingpapers/article/view/19848420.2017v18n2p171. Acesso em: 20 abr. 2025.

SILVA, I. da. *De quem nós/a gente está(mos) falando afinal?*: uma investigação sincrônica da variação entre nós e a gente como estratégias de designação referencial. 2004. 145 f. Dissertação (Mestrado em Linguística) – Universidade Federal de Florianópolis, Florianópolis, 2004.

ZILLES, A. M. S. Grammaticalization of 'a gente' in Brazilian Portuguese. *University of Pennsylvania Working Papers in Linguistics*, [S, l], v. 8, n. 3, pp. 297-310, 2002.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72246

Sistematização de estudos sociolinguísticos sobre a representação da primeira pessoa do plural na fala alagoana: um olhar para a variável comunidade

Systematization of sociolinguistic studies on the representation of the first-person plural in Alagoas speech: a look at the community variable

Elyne Giselle de Santana Lima Aguiar Vitório[1]

RESUMO

Focalizamos a representação da primeira pessoa do plural na fala alagoana, com o intuito de apresentar uma sistematização de resultados de pesquisas sobre a variação *nós* e *a gente* na posição de sujeito e na posição de não sujeito, bem como sobre a concordância verbal com os pronomes *nós* e *a gente*. Para tanto, adotamos métodos de análise de estatística descritiva e inferencial e controlamos a variável comunidade, de modo a responder às seguintes questões: qual o comportamento linguístico da representação da primeira pessoal do plural na variedade alagoana? Há diferenças estatisticamente significativas nesse comportamento em relação à comunidade de fala? Os resultados evidenciam que a comunidade é uma variável importante na representação da primeira pessoa na variedade alagoana, pois mostra que *a gente* tanto na posição de sujeito quanto na posição de não sujeito é mais frequente em comunidades mais urbanizadas, bem como que a variante não padrão da concordância verbal com *nós* é mais frequente em comunidades mais rurais, ao passo que a variante não padrão da concordância com *a gente* é mais frequente em comunidades mais urbanas, o que nos leva a ponderar que o processo de urbanização é um fator importante para esses contextos linguísticos variáveis.
PALAVRAS-CHAVE: Variação. Nós. A gente. Comunidade. Alagoas.

ABSTRACT

We focus on the representation of the first-person plural in alagoana speech, aiming to present a systematization of research results on the variation between *nós* and *a gente* in the subject and non-subject positions, as well as on verbal agreement with the pronouns *nós* and *a gente*. To this end, we adopted descriptive and inferential statistical analysis methods and controlled for the community variable, in order to answer the following questions: what is the linguistic behavior of the representation of the first-person plural in the alagoana variety? Are there statistically significant differences in this behavior compared to the speech community? The results show that community is an important variable in the representation of the first person in the alagoana variety, as it shows that *a gente* is more frequent in more urbanized communities, both as a subject and as a non-subject, and that the non-standard variant of verbal agreement with *nós* is more frequent in more rural communities, while the non-standard variant of agreement with *a gente* is more frequent in more urban communities. This leads us to consider that the urbanization process is an important factor for these variable linguistic contexts.
KEYWORDS: Variation. *Nós*. *A gente*. Community. Alagoas.

1. Introdução

Pesquisas sociolinguísticas mostram que, nas variedades brasileiras, a representação da primeira pessoa do plural é variável, podendo ocorrer com as variantes *nós* e *a gente* tanto na posição de

[1] Universidade Federal de Alagoas (UFAL), elyne.vitorio@gmail.com, https://orcid.org/0000-0002-6279-2379.

Recebido em 20/01/2026 - Aceito em 10/06/2026

sujeito quanto na posição de complemento e adjunto (Vianna; Lopes, 2012; Vianna; Lopes, 2015; Vitório, 2017; Silva; Vitório, 2018). Na posição de sujeito, *a gente* é a variante preferida, ao passo que, na posição de não sujeito, as formas padrão do paradigma da primeira pessoal do plural – *nós*, *nos*, *conosco* e *nosso/a* – predominam, mas *a gente* já encontra sua porta de entrada, com restrições linguísticas e sociais condicionando esses processos variáveis.

As descrições sociolinguísticas tendem a mostrar, conforme Vianna e Lopes (2015), que a inserção da variante *a gente* na posição de sujeito nas variedades brasileiras mais urbanizadas tende a ser favorecida por mulheres, falantes mais escolarizados e mais jovens, representando uma mudança em progresso (Labov, 2008). Vale ressaltar que as descrições apresentadas pelas autoras também mostram que, nas variedades com traços mais rurais, há uma inserção mais tardia dessa variante, com o pronome *nós* apresentando percentuais maiores de uso, o que pode ser um indício de que *a gente* seja uma variante externa a comunidade rural.

Diretamente relacionada à variação *nós* e *a gente* na posição de sujeito está a concordância verbal com essas variantes, que também é variável (Rubio, 2015). A realização da concordância verbal com *nós* e *a gente* pode ocorrer tanto com morferma *-mos* quanto com o morfema zero, como *nós estudamos aqui* / *nós estuda aqui* e *a gente estuda aqui / a gente estudamos aqui*. Esses processos variáveis também tendem a ser condicionados por restrições linguísticas e sociais, com as variantes não padrão *nós estuda aqui* e *a gente estudamos aqui* tendendo a ser avaliadas negativamente, principalmente em áreas mais urbanas.

A variação na concordância verbal com o pronome *nós* também tende a ser fortemente condicionada pelo nível de escolarização dos falantes, indicando que quanto mais escolarizado menor a realização da variante não padrão – verbo com morfema zero, como *nós estuda muito*. Outro fator relacionado à concordância verbal com o pronome *nós* diz respeito ao processo de urbanização, Agostinho e Coelho (2015, p. 80) argumentam que "os falantes de zona rural, ao migrarem para a cidade, estão adquirindo as desinências número-pessoais que encontraram na fala urbana e, simultaneamente, aprendendo nas escolas as desinências da norma-padrão".

Nesta pesquisa, focalizamos a representação da primeira pessoa do plural na variedade alagoana a partir da sistematização de resultados de pesquisas sociolinguísticas sobre as seguintes regras variáveis, a saber, (i) *nós* e *a gente* na posição de sujeito, (ii) *nós* e *a gente* na posição de não sujeito – complemento e adjunto, (iii) concordância verbal com o pronome *nós* e (iv) concordância verbal com o pronome *a gente*. Nosso objetivo é apresentar generalizações sobre a representação da primeira pessoa do plural na fala alagoana tomando por base descrições sociolinguísticas realizadas até o momento nessa variedade linguística.

Para tanto, recorremos à revisão sistemática que sumariza evidências científicas disponíveis sobre uma questão específica de pesquisa, oferecendo uma visão confiável sobre um tema para fundamentar decisões e gerar novas hipóteses. Também adotamos métodos de análise de estatística descritiva e inferencial (Freitag, 2021; Siqueira; Sousa, 2024) e controlamos a variável comunidade,

de modo a responder as questões: qual o comportamento linguístico da representação da primeira pessoal do plural na fala alagoana? Há diferenças estatisticamente significativas nesse comportamento em relação à comunidade de fala?

O artigo está estruturado da seguinte forma: além desta seção introdutória, que faz uma breve apresentação da proposta do trabalho, apresentamos, na próxima seção, o percurso metodológico adotado para a sistematização e análise dos dados, em seguida, analisamos e discutimos os resultados obtidos acerca do uso da primeira pessoa do plural em correlação com a variável comunidade, e, por fim, encerramos as discussões levantadas acerca do tema. Esperamos contribuir para a agenda sociolinguística oferecendo uma visão da realidade linguística da variedade em foco tomando por base a variação da primeira pessoal do plural.

2. Notas metodológicas

Para atingirmos os objetivos propostos, recorremos à revisão sistemática que é vista como "uma forma de pesquisa que utiliza como fonte de dados a literatura sobre determinado tema" (Sampaio; Mancini, 2007, p. 83). A elaboração de estudo de revisão sistemática tende a apresentar as seguintes etapas, a saber, definição da pergunta de pesquisa, busca das evidências, revisão e seleção dos estudos, análise e qualidade metodológica dos estudos e apresentação dos resultados. Para tanto, necessário também se faz a elaboração de um protocolo de pesquisa, que inclua, entre outros fatores, critérios de inclusão e exclusão das fontes primárias.

Em nosso estudo, adotamos os seguintes passos metodológicos: (i) procedemos à busca e seleção de pesquisas sobre a variação na primeira pessoa do plural na fala alagoana considerando as variáveis *nós* e *a gente* sujeito, *nós* e *a gente* não sujeito, concordância verbal com *nós* e concordância verbal com *a gente*; (ii) realizamos a análise de sistematização das informações, bem como codificação e filtragem dos dados, através do uso de técnicas estatísticas (Freitag, 2021; Siqueira; Sousa, 2024), e, por fim, (iii) interpretamos os dados considerando a distribuição geral de cada variável e a correlação com a variável comunidade.

Na etapa de busca e seleção dos estudos, delimitamos as palavras-chave – variação, *nós, a gente*, língua falada – e os critérios de inclusão – estudo sociolinguístico, primeira pessoa do plural e fala alagoana – e identificamos 13 pesquisas. No entanto, selecionamos 11 estudos, de acordo com as variáveis *nós* e *a gente* sujeito (Vitório, 2017; Oliveira; Nascimento, 2017; Feitosa; Vitório, 2018; Souza; Vitório, 2021; Silva, 2022), *nós* e *a gente* não sujeito (Vitório, 2017; Silva; Vitório, 2018; Santana; Vitório, 2022), concordância verbal com *nós* (Vitório, 2017; Silva; Vitório, 2020; Silva, 2021; Souza, 2021) e concordância verbal com *a gente* (Vitório, 2017, Rodrigues, 2018; Souza; Vitório, 2021; Silva, 2022)[2].

[2] Excluímos as pesquisas de Vitório (2015) e Vitório (2016). A pesquisa de Vitório (2015) analisa a variação *nós* e *a gente* na posição de sujeito, mas toma como amostra a fala de crianças alagoanas; e a pesquisa de Vitório (2016) analisa a variação *nós* e *a gente* na fala maceiense, mas considera uma amostra de falantes com ensino superior na fala de Maceió, nível de escolarização contemplado na pesquisa de Vitório (2017).

Essas pesquisas seguem a metodologia da Sociolinguística Variacionista (Labov, 2008), o que significa que focalizam a análise de uma regra variável e aplicam análise estatística para observar a frequência dos dados e inferir associação entre a distribuição da variável dependente e as variáveis independentes. Ao identificarmos e selecionarmos os estudos primários que compõem esta sistematização de pesquisas sociolinguísticas sobre a primeira pessoa do plural na fala alagoana, extraímos as seguintes informações de cada amostra: comunidade, estratificação da amostra e número de entrevistas sociolinguísticas, conforme quadro 1.

Quadro 1: Informações sobre as amostras das pesquisas selecionadas[3]

Comunidade	Estratificação da amostra	Nº de entrevistas
Maceió	sexo/gênero (masculino/feminino) faixa etária (15-29 anos / 30-44 anos / + 44 anos) escolaridade (EF / EM / ES)	72
Palmeira dos Índios	sexo/gênero (masculino/feminino) faixa etária (18-30 anos / 40-55 anos / + 65 anos) escolaridade (EF / ES)	24
Alto Sertão	sexo/gênero (masculino/feminino) faixa etária (18-29 anos / 30-44 anos / + 44 anos) escolaridade (E0 / EF / EM / ES)	96
Comunidade Quilombola Serra das Viúvas	sexo/gênero (masculino/feminino) faixa etária (25-50 anos / + 60 anos)	20
Zona rural de Pariconha	sexo/gênero (masculino/feminino) faixa etária (18-29 anos / 30-44 anos / + 44 anos) escolaridade (EF / EM / ES)	36

Fonte: Elaboração da autora

Essas comunidades, conforme a imagem 1, estão localizadas em Alagoas. Maceió, capital do estado, fica localizada no leste alagoano e possui uma população estimada em 1,025 milhão. Palmeira dos Índios está localizada no agreste alagoano e fica a cerca de 136 Km de Maceió, sendo a quarta maior cidade do estado, com uma população estimada em 73.213 habitantes. O Alto Sertão, representado pela amostra do Projeto Lusa (Vitório, 2020), é localizado no sertão alagoano, sendo formado pelos municípios de Piranhas, Olho D'Água do Casado, Delmiro Gouveia, Pariconha, Água Branca, Inhapi, Mata Grande e Canapi, conforme a imagem 2.

A zona rural de Pariconha, pertencente ao município de Pariconha, também está situada na região do Alto Sertão alagoano e fica cerca de 314 Km da capital Maceió. O município de Pariconha é constituído por aproximadamente 10.264 habitantes, que vivem majoritariamente na zona rural –

[3] As abreviações E0, EF, EM e ES significam, respectivamente, sem escolarização, ensino fundamental, ensino médio e ensino superior.

72,76% dos pariconheses residem em povoados e sítios (Silva, 2021). A comunidade quilombola Serra das Viúvas fica localizada no Alto Sertão de Alagoas a 7 km do munícipio de Água Branca e possui cerca de 600 moradores, que estão distribuídos em 70 famílias, com a maior parte dos moradores trabalhando na própria comunidade (Souza, 2020).

Imagem 1: Mapa de Alagoas



Fonte: https://www.baixarmapas.com.br/mapa-de-alagoas-mesorregioes/

Imagem 2: Regiões do Alto Sertão de Alagoas



Fonte: https://dados.al.gov.br/catalogo/km/dataset/regioes-de-planejamento-estado-de-alagoas/resource/361b1d5c-2a22-4f54-a125-be3efffdab8c

Após a extração dos dados relevantes para nossa pesquisa, recuperamos os dados de frequências apresentados em cada estudo. Consideramos os dados tabulares de cada variável dependente controlada – *nós* e *a gente* na posição de sujeito, *nós* e *a gente* na posição de não sujeito, concordância

verbal com *nós* e concordância verbal com *a gente* – que nos forneceram a distribuição das variantes nas comunidades analisadas. Em seguida, na plataforma R (R Core Team, 2022), executamos testes estatísticos inferenciais e de força de associação, tomando por base cada variável dependente e a sua correlação com a variável comunidade.

> As análises dos dados seguem métodos de análise estatística descritiva e inferencial. Para cada variável, realizamos análise univariada da distribuição da variável dependente, visualizada por meio de gráficos. Como teste de significância para as análises, observamos a relação entre a frequência absoluta da variável dependente (e a variável independente, existindo), se é efeito do acaso ou se é uma distribuição significativa, por meio de teste de qui-quadrado. O teste apresenta um p-valor, que é comparado com nosso p-valor pré-determinado, o α (alfa) no valor de < 0,05 (5%), que significa que, se repetirmos o teste 100 vezes, cinco dessas vezes o resultado pode ser diferente do obtido inicialmente. Nossos resultados são efeitos do acaso, não havendo significância estatística, quando p = ou > 0,05; quando o p < 0,05, a distribuição de nossa variável dependente não é efeito do acaso, havendo significância em sua distribuição (Siqueira; Sousa, 2024, p. 86).

Os resultados são apresentados considerando a distribuição de cada variável – *nós* e *a gente* sujeito, *nós* e *a gente* não sujeito, concordância verbal com *nós* e concordância verbal com *a gente* – e sua correlação com a variável comunidade, com intuito de analisar se encontramos correlações positivas entre o uso da primeira pessoa do plural e a comunidade de fala. Para tanto, analisamos qual o comportamento linguístico da representação da primeira pessoal do plural na fala alagoana e se há diferenças sociolinguísticas estatisticamente significativas desses contextos linguísticos variáveis em relação à comunidade.

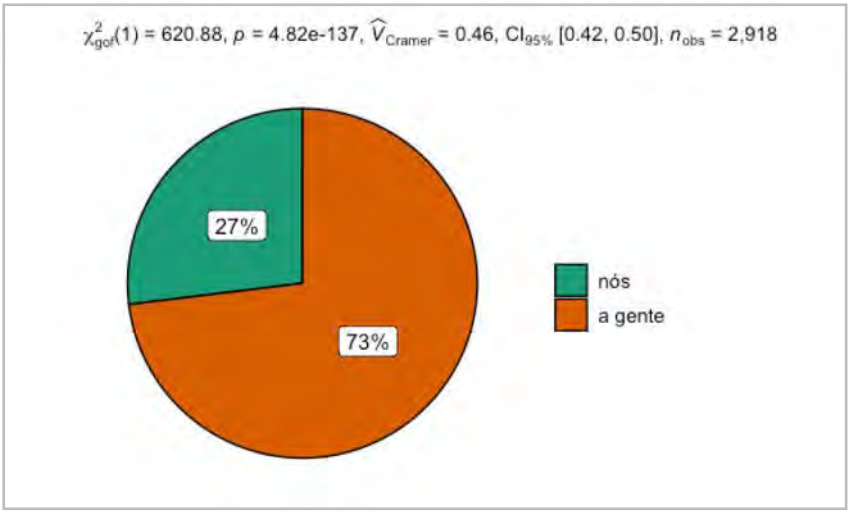
3. Descrição e análise dos dados

3.1. Sobre a variação *nós* e a *gente*

A variação *nós* e *a gente* na posição de sujeito, como *nós chegamos cedo* e *a gente chegou cedo*, tem sido um dos fenômenos linguísticos variáveis mais descritos nas variedades brasileiras (Vianna; Lopes, 2015). As descrições sociolinguísticas têm evidenciado que *a gente* é a variante preferida, mas restrições linguísticas e extralinguísticas têm agido como fatores que ainda retardam a implementação da mudança linguística (Labov, 2008). As pesquisas têm mostrado que a variante não padrão *a gente* tende a ser favorecida majoritariamente por mulheres, pessoas mais escolarizadas, mais jovens e em regiões mais urbanizadas.

Na variedade alagoana, consideramos as pesquisas nas comunidades de Maceió (Vitório, 2017), Palmeira dos Índios (Oliveira; Nascimento, 2017), Alto Sertão (Feitosa; Vitório, 2018), zona rural de Pariconha (Silva, 2022) e comunidade quilombola Serra das Viúvas (Souza; Vitório, 2021), e observamos, conforme gráfico 1, que a variante não padrão *a gente* é a forma preferida. O conjunto de dados mostra uma distribuição de 73% (n = 2132) de *a gente* versus 27% (n = 786) de *nós*, e essa distribuição é estatisticamente significativa ($\chi^2(1, N=2918) = 620.88$, $p = 0.001$), o que significa considerar que não é aleatória.

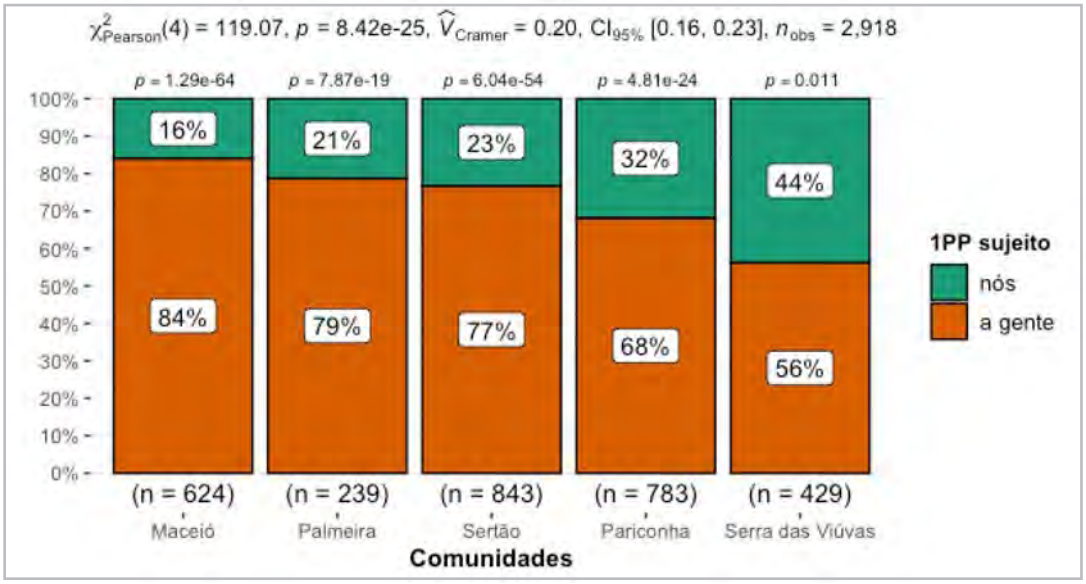
Gráfico 1: Distribuição de *nós* e *a gente* sujeito na fala alagoana



Fonte: Elaboração da autora

Ao analisarmos a distribuição da variável dependente juntamente à distribuição da variável comunidade, observamos, conforme gráfico 2, que o efeito da comunidade é estatisticamente significativo ($\chi^2(4, N=2918) = 119.07$, $p = 0.001$), com associação moderada ($V^2 = 0.20$), mostrando que o ritmo de inserção de *a gente*, na fala alagoana, também é condicionado por essa variável. Esse conjunto de dados mostra que, à medida que nos afastamos da capital Maceió, *a gente* diminui o seu percentual de uso – 84% (n = 524) em Maceió, 79% (n = 188) em Palmeira dos Índios, 77% (n = 646) no Alto Sertão, 68% (n = 533) na zona rural de Pariconha e 56% (n = 241) na comunidade quilombola Serra das Viúvas.

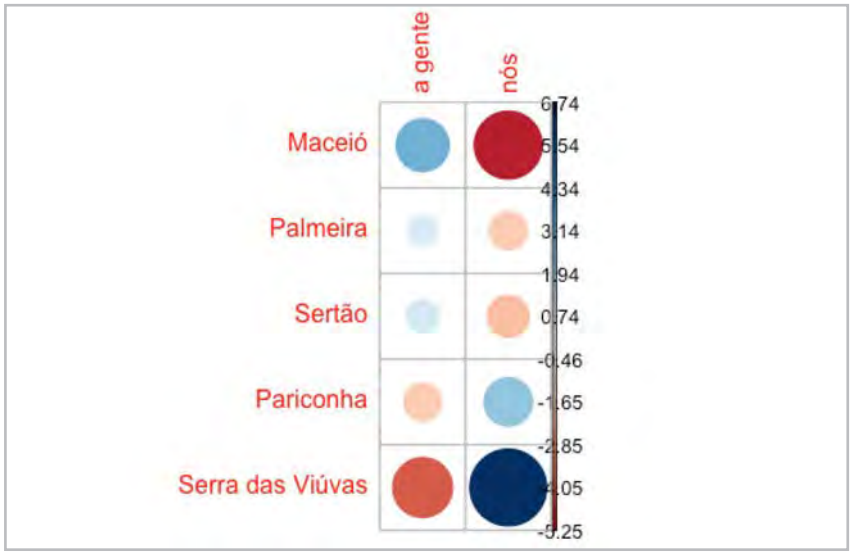
Gráfico 2: Distribuição de *nós* e *a gente* sujeito em relação à variável comunidade



Fonte: Elaboração da autora

Ao analisarmos o quanto cada amostra contribui para o escore do teste de χ2, geramos a matriz de correlação, conforme figura 1, que mostra que o tamanho do círculo é proporcional à contribuição da célula para o resultado do teste, com associações positivas entre linhas e colunas apresentadas por círculos em azul e associações negativas por círculos em vermelho. Isso significa considerar que há uma associação positiva entre o uso de *a gente* e as comunidades Maceió, Palmeira dos Índios e Alto Sertão, mas uma associação negativa entre o uso de *a gente* e a zona rural de Pariconha e a comunidade quilombola Serra das Viúvas.

Figura 1: Contribuição de cada fator para o escore do teste de χ2 quanto à comunidade



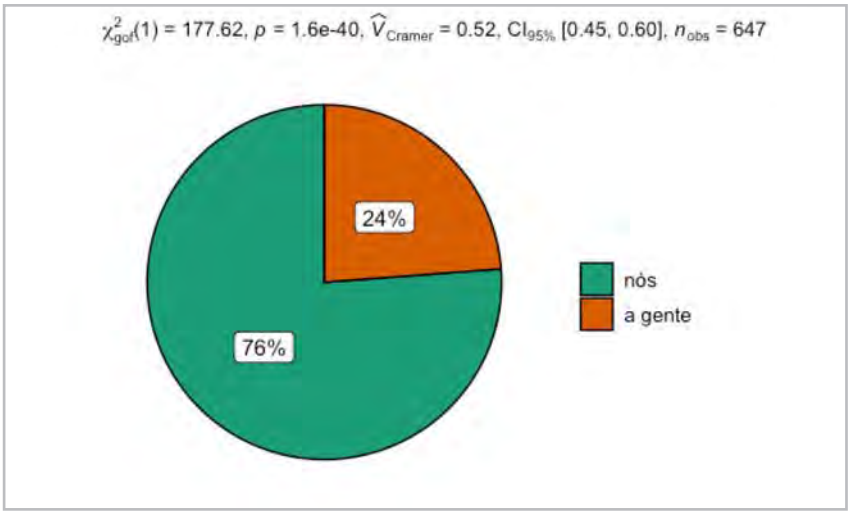
Fonte: Elaboração da autota

Pesquisas sociolinguísticas mostram que além de jovens, mulheres e pessoas mais escolarizadas favorecerem mais a inserção de *a gente*, comunidades urbanas também tendem a apresentar um comportamento mais inovador, com comunidades com traços mais rurais apresentando um ritmo mais lento de inserção dessa variante (Maia, 2003; Mendes, 2007; Foeger, 2014; Pinto; Berlinck, 2022). Ao analisar a inserção de *a gente* sujeito na fala alagoana, Vitório (2024) argumenta que a variação *nós* e *a gente* na posição de sujeito também esteja relacionada ao contínuo rural/urbano, conforme Bortoni-Ricardo (2004).

Ao considerarmos a questão da urbanização na inserção de *a gente*, ponderamos que *a gente* pode ser uma variante externa à comunidade rural, que é trazida ou pela mídia ou pelos falantes que possuem maior mobilidade espacial, ou seja, falantes que têm maior contato com grandes centros urbanos, seja para trabalhar ou estudar. Associada aos dados obtidos, essa ponderação pode ser um indício de que, em comunidades alagoanas com traços mais rurais, como a zona rural de Pariconha e a comunidade quilombola Serra das Viúvas, haja um ritmo mais lento de inserção de *a gente*, havendo um maior percentual de uso da variante padrão *nós*.

Diferentemente da variação *nós* e *a gente* na posição de sujeito, o uso das variantes de *nós* na posição de não sujeito – complemento e adjunto – são mais frequentes (Omena, 1996; Ramos; Bezerra; Rocha, 2009; Tamanine, 2010; Vianna; Lopes, 2012; Araújo; Almeida, 2014), mas já observamos a inserção de *a gente* nesses contextos sintáticos, como *o menino esperou a gente no ponto*, *a professora vai com a gente*, *a aula da gente começou tarde*. Na fala alagoana, consideramos as descrições nas comunidades de Maceió (Vitório, 2017), Alto Sertão (Silva; Vitório, 2018) e na zona rural de Pariconha (Santana; Vitório, 2022), e observamos, conforme gráfico 3, que as variantes de *nós* são as formas mais utilizadas pelos falantes.

Gráfico 3: Distribuição de *nós* e *a gente* não sujeito na fala alagoana

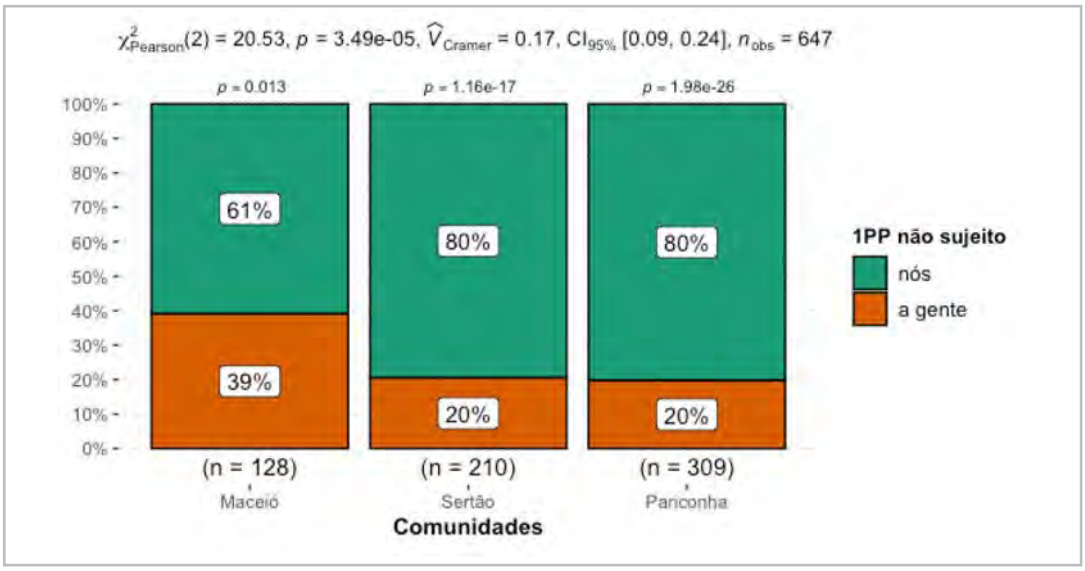


Fonte: Elaboração da autora

A distribuição geral dos dados mostra que, na função de não sujeito, as variantes padrão do paradigma pronominal de primeira pessoa do plural (*nós*, *nos*, *conosco* e *nosso/a*) são as formas preferidas – 76% (n = 493) contra 24% (n = 154) da variante *a gente* e essa distribuição é estatisticamente significativa ($\chi^2(1, N = 647) = 177.62$, $p = 0.001$). Isso significa considerar que há restrições que condicionam esse processo de variação. Na fala alagoana, as descrições sociolinguísticas mostram que *a gente* tende a ser favorecida no núcleo verbal, na relação gramatical de complementação, entre falante menos escolarizado, do sexo feminino e mais novo. No entanto, ainda não temos uma análise da correlação com a comunidade de fala.

Ao analisarmos a distribuição dos dados em relação à variável comunidade, observamos, conforme gráfico 4, que o efeito da comunidade é estatisticamente significativo ($\chi^2(2, N = 647) = 20.57$, $p = 0.001$), com associação fraca ($V^2 = 0.17$). Os dados mostram que, no Alto Sertão alagoano e na zona rural de Pariconha, há um mesmo comportamento linguístico em relação ao uso dessas variantes na posição de não sujeito, com *a gente* apresentando percentual de 20% (n = 43 e n = 61, respectivamente). Em Maceió, temos um percentual maior de uso de *a gente* – 39% (n = 50), o que indica que a capital favorece mais a sua inserção.

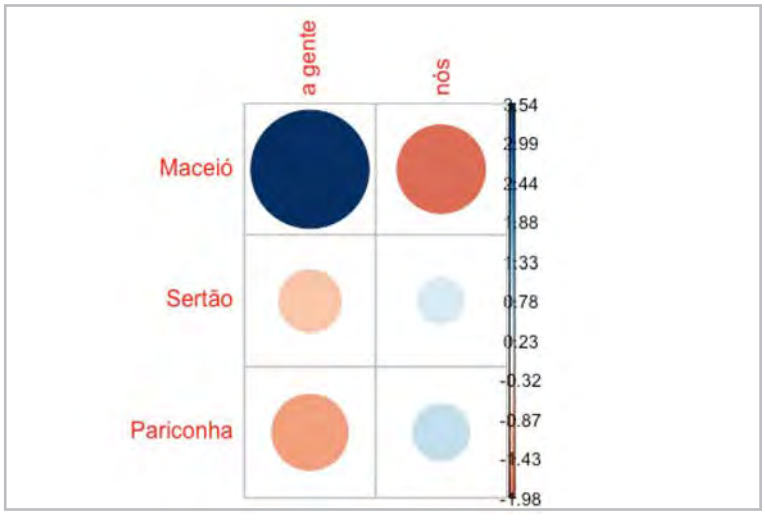
Gráfico 4: Distribuição de *nós* e *a gente* não sujeito em relação à variável comunidade



Fonte: Elaboração da autora

Para analisarmos o quanto cada amostra contribui para o escore do teste de χ2, geramos a matriz de correlação, conforme figura 2. A contribuição de cada fator é indicada pelo tamanho e cor do círculo, com círculos em azul apresentando associações positivas entre linhas e colunas, e círculos em vermelho associações negativas. Os dados mostram que *a gente* na posição de não sujeito é mais favorecido em Maceió, ao passo que, no Alto Sertão e na zona rural de Pariconha, observamos uma associação positiva para o uso das variantes padrão do paradigma pronominal de primeira pessoa do plural – *nós*, *nos*, *conosco* e *nosso/a*.

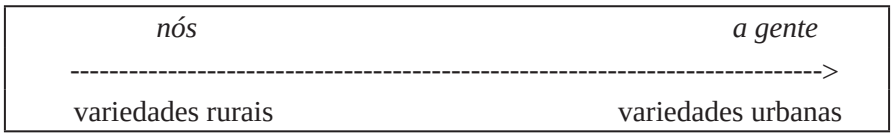
Figura 2: Matriz de correlação quanto à variação *nós* e *a gente* não sujeito



Fonte: Elaboração da autora

Esses dados nos levam a retornar a hipótese de que, em relação à inserção de *a gente* na fala alagoana, há um ritmo mais lento em comunidades com traços mais rurais. Essa hipótese parece se aplicar ao comportamento de uso de *a gente* tanto na posição de sujeito quanto na posição de não sujeito. Dessa forma, podemos assumir que, na variedade alagoana, a inserção de *a gente* está mais avançada em variedades urbanas, com comunidades rurais retardando o processo de inserção dessa variante. Esse comportamento pode ser observado no contínuo de urbanização proposto para o uso das variantes *nós* e *a gente*, conforme figura 3.

Figura 3: Contínuo de urbanização para o uso de *nós* e *a gente* na fala alagoana



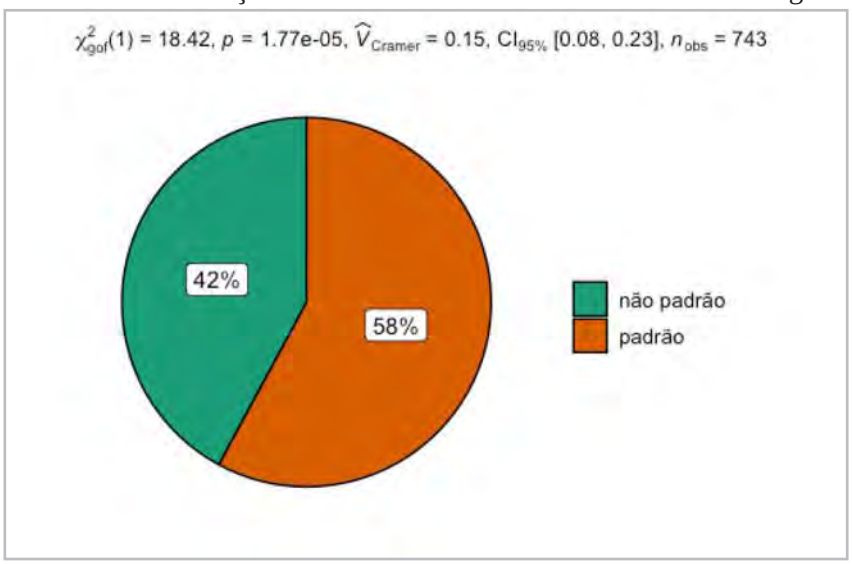
Fonte: Elaboração da autora

O contínuo de urbanização para o uso de *nós* e *a gente*, adaptado da proposta de Bortoni-Ricardo (2004), mostra que a variante *nós* está mais ligada às variedades rurais ao passo que a variante *a gente* tem mais espaço nas variedades urbanas. Essa proposta nos leva a considerar que, na descrição da inserção dessas variantes na fala alagoana, é importante considerar o papel da urbanização, que, de acordo com Bortoni-Ricardo (2005, p. 73), deve estar relacionada também com "zona de residência, acesso à escolarização, qualidade do ensino recebido, características da rede de relações informais, exposição aos meios de comunicação etc. [...]".

3.2. Sobre a variação na concordância verbal com os pronomes *nós* e *a gente*

A concordância verbal com o pronome *nós* é realizada tanto com o moferma *-mos*, variante padrão, como *nós estudamos*, quanto com o moferma zero, variante não padrão, como *nós estuda*, caracterizando-se como um fenômeno linguístico variável nas variedades brasileiras (Rubio, 2015). Para a sistematização de pesquisas sociolinguísticas sobre esse processo variável na fala alagoana, consideramos as comunidades Maceió (Vitório, 2017), Alto Sertão (Silva; Vitório, 2020), zona rural de Pariconha (Silva, 2021) e a comunidade quilombola Serra das Viúvas (Souza, 2021), e obtivemos, conforme gráfico 5, os seguintes resultados.

Gráfico 5: Distribuição da concordância verbal com *nós* na fala alagoana

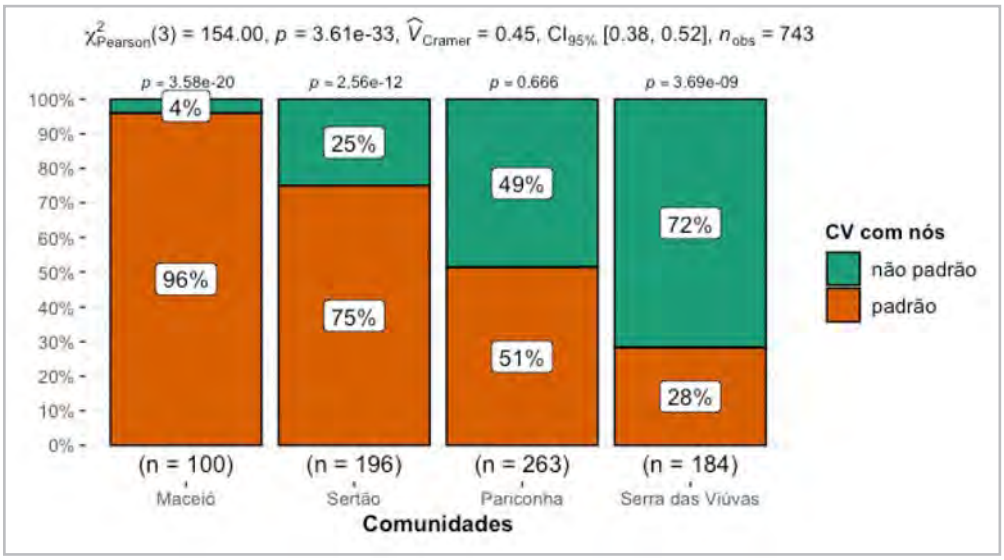


Fonte: Elaboração da autora

A distribuição geral dos dados mostra um percentual maior de realização da variante padrão – 58% (n = 430) contra 42% (n = 313) da variante não padrão, indicando que essa distribuição é estatisticamente significativa ($\chi^2(1, N = 743) = 18.42$, p = 0.001). Esse dado mostra que há restrições linguísticas e extralinguísticas que condicionam esse processo de variação, mas, nesta sistematização, focalizamos na atuação da variável comunidade. O controle dessa variável objetiva checar se há um maior favorecimento da variante não padrão em comunidades com traços mais rurais, conforme Agostinho e Coelho (2015).

Ao analisarmos a distribuição dos dados tomando por base a variável comunidade, observamos, conforme gráfico 6, que o efeito da comunidade é estatisticamente significativo ($\chi^2(3, N = 743) = 154.00$, p = 0.001), com associação moderada ($V^2 = 0.45$), bem como que há um aumento no percentual de uso da variante não padrão em comunidades com traços mais rurais. Os dados mostram os seguintes percentuais de uso da variante não padrão por comunidade: 4% (n = 4) em Maceió, 25% (n = 49) no Alto Sertão, 49% (n = 128) na zona rural de Pariconha e 72% (n = 132) na comunidade quilombola Serra das Viúvas.

Gráfico 6: Distribuição da concordância verbal com *nós* relação à variável comunidade

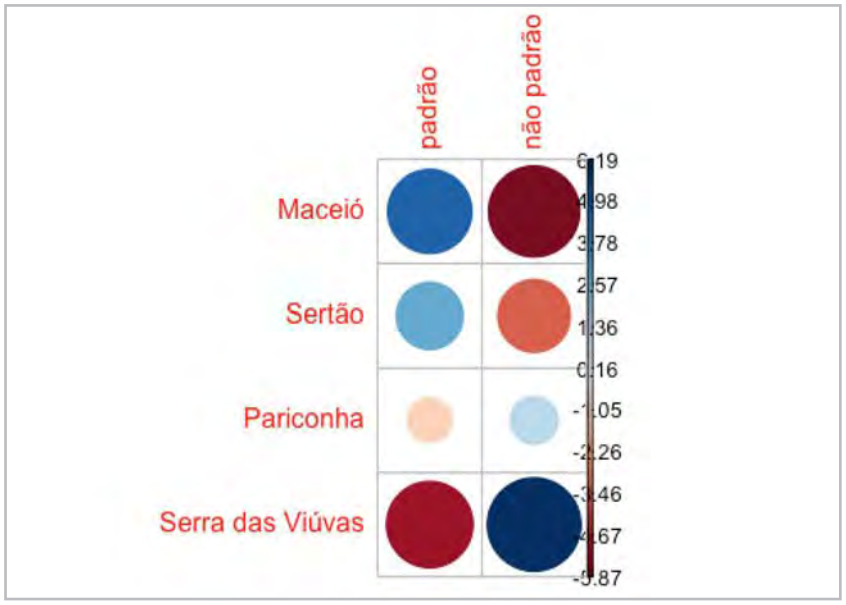


Fonte: Elaboração da autora

Esse conjunto de dados mostra que, à medida que nos afastamos da capital Maceió, há um aumento no uso da variante não padrão, revelando que, em comunidades rurais, o percentual de uso da variante não padrão aumenta consideravelmente. Bastos e Araújo (2025) argumentam que pesquisas sociolinguísticas sobre a variação da concordância verbal têm revelado que os caminhos da mudança em relação às restrições linguísticas são os mesmos, mas comunidades rurais afro-brasileiras tendem a apresentar resultados distantes da norma padrão, com restrições socioculturais sendo relevantes, pois explicam o perfil sócio-histórico dessas comunidades.

Ao analisarmos o quanto cada amostra contribui para o escore do teste de $\chi 2$, geramos a matriz de correlação, conforme figura 4. Os dados mostram uma associação negativa entre a variante não padrão e as comunidades Maceió e Alto Sertão, mas uma associação positiva entre a variante não padrão e a zona rural de Pariconha e a comunidade quilombola Serra das Viúvas. Esse achado pode estar relacionado ao perfil social dessas comunidades, como perfil de urbanização e acesso à escolarização. Souza (2021) argumenta que o alto percentual de uso da variante não padrão na comunidade quilombola Serra das Viúvas está associado ao perfil interiorano da comunidade, bem como a falta de escolarização dos participantes da pesquisa.

Figura 4: Matriz de correlação quanto à concordância verbal com *nós*



Fonte: Elaboração da autora

Em relação ao nível de escolarização, observamos que, nas comunidades Maceió (Vitório, 2017), Alto Sertão (Silva; Vitório, 2020) e zona rural de Pariconha (Silva, 2021), conforme quadro 1, houve o controle da variável escolaridade, mas, na zona rural de Pariconha, há também uma associação negativa com a variante padrão, embora em menor grau se comparada à comunidade quilombola Serra das Viúvas. Esse dado reforça a hipótese de que o processo de urbanização da comunidade, que está relacionado também ao acesso à escolarização, interfere na relação da concordância verbal com o pronome *nós*, com a variante não padrão sendo mais favorecida em variedades rurais, conforme proposto na figura 5.

Figura 5: Contínuo de urbanização para concordância verbal com *nós* na fala alagoana

Nós estuda	*Nós estudamos*
---------->	
variedades rurais	variedades urbanas

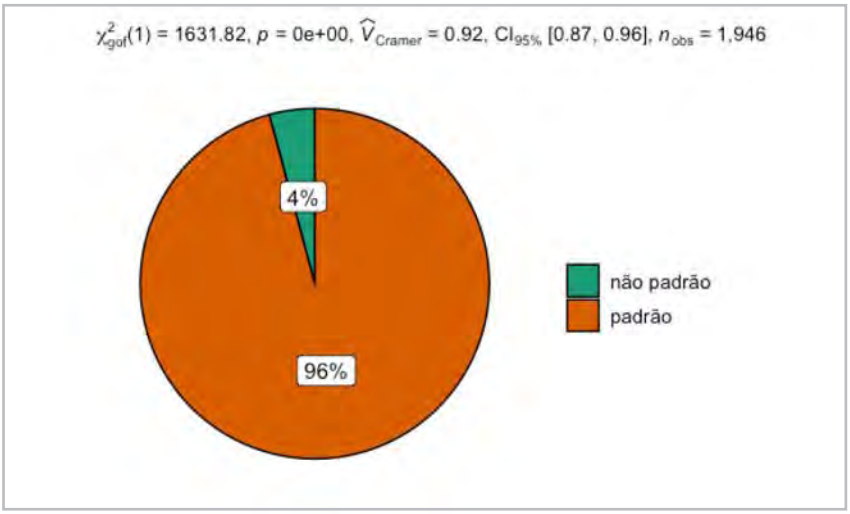
Fonte: Elaboração da autora

Os dados obtidos indicam que, na fala alagoana, o contínuo de urbanização condiciona o uso da concordância verbal com o pronome *nós*, revelando que, em variedades rurais, há um uso maior da variante não padrão, mas o seu percentual de uso diminui em variedades urbanas. Associado ao processo de urbanização e ao acesso à escola está a avaliação social dessa variável. Sendo visto como um processo variável que tende a receber, em comunidades urbanizadas, uma avaliação social negativa, a variante não padrão é um traço linguístico avaliado negativamente. Essa avaliação também tende a inibir o uso da variante não padrão.

Quanto à variação com a concordância verbal com o pronome *a gente*, observamos que também é variável, podendo ser realizada tanto com o morfema zero, variante padrão, como *a gente estuda*, quanto com o moferma *-mos*, variante não padrão, como *a gente estudamos*. No entanto, nas variedades brasileiras, conforme Rubio (2015), esse contexto linguístico variável, diferentemente da variedade europeia, tende a se comportar mais como uma regra semi-categórica, conforme Labov (2003), e o uso da variante não padrão, como *a gente organizamos a manifestação*, é vista como um processo linguístico de hipercorreção (Bagno, 2011).

Para a nossa sistematização, consideramos as comunidades Maceió (Vitório, 2017), Alto Sertão (Rodrigues, 2018), zona rural de Pariconha (Silva, 2022) e comunidade quilombola Serra das Viúva (Souza; Vitório, 2021) e observamos, conforme gráfico 7, que, na fala alagoana, seguindo o padrão das descrições sociolinguísticas, estamos diante de uma regra semi-categórica. Os dados mostram percentuais de 4% (n = 82) para a variante não padrão contra 96% (n = 1864) para a variante padrão, com essa distribuição geral sendo estatisticamente significativa ($\chi^2(1, N = 1946) = 1631.82$, $p = 0.001$), ou seja, não aleatória.

Gráfico 7: Distribuição da concordância verbal com *a gente* na fala alagoana



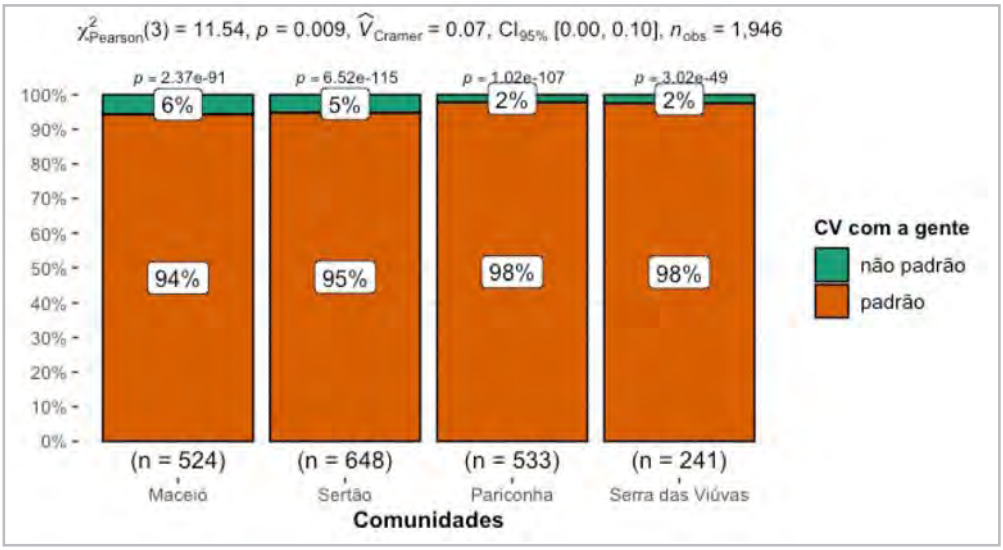
Fonte: Elaboração da autora

Esses dados mostram que, na variedade alagoana, *a gente*, variante preferida na posição de sujeito, tende a ser realizada com o verbo na terceira pessoal do singular – morfema zero, como *a gente estuda nessa escola*. No entanto, computamos um percentual de 4% de *a gente* com o verbo na primeira pessoal do plural, como *a gente estudamos nessa escola*. Por ser caracterizado como um processo de hipercorreção, o uso da variante não padrão, como *a gente assinamos o documento*, tende a ocorrer em situações mais monitoradas de fala e está relacionada ao processo de insegurança linguística frente às pressões normativas.

> A hipercorreção é um fenômeno de linguagem muito comum entre pessoas que deram conta da existência de "outro falar" muito mais prestigiado do que o seu. Essas pessoas também desejam ser usuárias dessa forma prestigiada, do "falar mais correto". Para tal, esforçam-se em "corrigir" sua fala e acabam incorrendo no erro de corrigi-la demasiadamente (Bortoni; Alves, 2014, p. 131).

Ao analisarmos a distribuição geral dos dados em relação à variável comunidade, observamos, conforme gráfico 8, que o efeito da comunidade é estatisticamente significativo ($\chi^2(3, N = 1946) = 11.54$, $p = 0.009$), com associação fraca ($V^2 = 0.07$), bem como que há a realização da variante não padrão, como *a gente gostamos de doce*, em todas as comunidades, com percentuais de 6% (n = 30) em Maceió, 5% (n = 34) no Alto Sertão, 2% (n = 6) na zona rural de Pariconha e 2% (n =12) na comunidade quilombola Serra das Viúvas. Nos termos laboviano, apenas em Maceió e no alto Sertão estamos diante de uma regra variável.

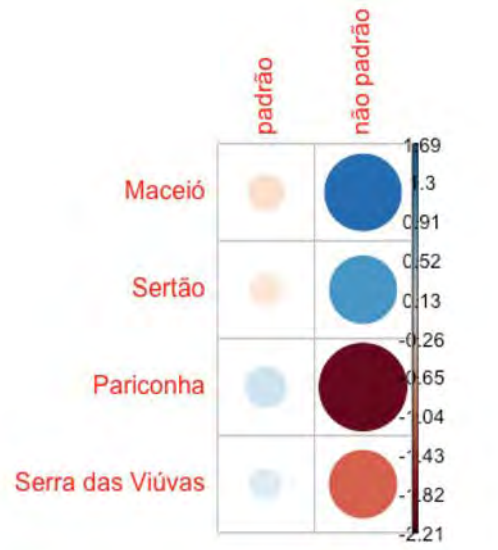
Gráfico 8: Distribuição da concordância verbal com *a gente* relação à variável comunidade



Fonte: Elaboração da autora

Em Maceió e no Alto Sertão, Vitório (2017) e Rodrigues (2018), respectivamente, mostram que essa variação é fortemente condicionada pela variável escolaridade, com a variante não padrão sendo mais frequente entre os falantes dos ensinos fundamental e médio. Ao analisarmos o quanto cada amostra contribui para o escore do teste de $\chi 2$, geramos a matriz de correlação, conforme figura 6, que mostra uma associação positiva entre a variante não padrão e as comunidades Maceió e Alto Sertão, mas uma associação negativa entre a variante não padrão e a zona rural de Pariconha e a comunidade quilombola Serra da Viúvas.

Figura 6: Matriz de correlação quanto à concordância verbal com *a gente*



Fonte: Elaboração da autora

Esses dados mostram que, diferentemente da concordância verbal com o pronome *nós* que tende a indicar que a variante não padrão, como *nós gosta de chocolate*, é mais frequente em variedades rurais, o que pode ser associado ao perfil social da comunidade, principalmente ao acesso à norma de prestígio social, a concordância verbal com o pronome *a gente* mostra que a variante não padrão, como *a gente gostamos de chocolate*, é mais frequente em comunidades urbanas, conforme proposto na figura 7. Esse contínuo de urbanização mostra que a variante *a gente estudamos* está mais relacionada a variedades urbanas.

Figura 7: Contínuo de urbanização para concordância verbal com *a gente* na fala alagoana

A gente estuda	*A gente estudamos*
---------->	
variedades rurais	variedades urbanas

Fonte: Elaboração da autora

Esse dado pode ser relacionado ao processo de hipercorreção que ocorre no uso de construções como *a gente estudamos muito*, *a gente chegamos cedo na aula*, fazendo com que quanto mais acesso à norma de prestígio social, como a marcação de plural, que, em comunidades mais urbanizadas, tende a ser avaliada positivamente, haja mais uso dessa variante. No entanto, é importante destacar que, apesar de ser mais frequente em variedades urbanas, caracterizando-se como uma regra variável, pesquisas sociolinguísticas mostram uma avaliação social negativa para essa variante nessas variedades linguísticas (Vitório, 2025).

4. Conclusão

Sistematizamos resultados de pesquisas sociolinguísticas sobre a representação da primeira pessoa do plural na variedade alagoana, considerando as variáveis *nós* e *a gent*e na posição de sujeito, *nós* e *a gente* na posição de não sujeito, concordância verbal com *nós* e concordância verbal com *a gente*. Para tanto, adotamos métodos de análise de estatística descritiva e inferencial, com o objetivo de responder as seguintes questões: qual o comportamento da representação da primeira pessoal do plural na fala alagoana e se a comunidade é uma variável significativa nesses processos linguísticos variáveis. Os dados mostram que estamos diante de variáveis linguísticas que sofrem interferência da comunidade.

No que diz respeito à variação *nós* e *a gente*, observamos que, na posição de sujeito, *a gente* é a variante preferida na fala alagoana, sendo também a forma mais utilizada nas comunidades analisadas. Também observamos que o seu percentual de uso diminui em comunidades mais rurais, indicando que há uma inserção mais tardia de *a gente* nessas comunidades. Na posição de não sujeito, observamos que as formas padrão do paradigma do *nós* predominam, mas há a inserção de *a gente* nesse contexto sintático, principalmente na capital alagoana – Maceió. Esses dados nos levam a argumentar que, em variedades rurais, maior é o percentual de uso de *nós*, com variedades urbanas sendo mais inovadoras.

Quanto ao processo de concordância verbal com os pronomes *nós* e *a gente*, observamos que as variantes padrão, *nós estamos aqui* e *a gente está aqui*, predominam na fala alagoana, mas com o perfil social da comunidade atuando de forma diferente. Em relação à concordância verbal com o pronome *nós*, observamos que a variante padrão é mais frequente em variedades urbanas, com variedades rurais favorecendo mais um uso da variante não padrão, ao passa que a concordância verbal com o pronome *a gente* mostra que a variante não padrão é mais frequente em comunidades urbanas, com comunidades rurais utilizando mais a variante padrão. Esse dado pode estar associado também com a avaliação social dessas variantes nas comunidades.

Referências

AGOSTINHO, Silvana; COELHO, Izete. Concordância de 1ª pessoa do plural na escrita escolar**.** *In*: ZILLES, Ana; FARACO, Carlos. (org.). *Pedagogia da variação linguística*: língua, diversidade e ensino. São Paulo. Parábola Editorial, 2015, pp. 79-121.

ARAÚJO, Silvana; ALMEIDA, Rosiane. A forma possessiva da gente em comunidades rurais do semiárido baiano. *In*: ALMEIDA, Norma; CARNEIRO, Zenaide (org.). *Variação linguística no semiárido baiano*. Feira de Santana: UEFS Editora, 2014. pp. 137- 173.

BAGNO, Marcos. *Gramática pedagógica do português brasileiro*. São Paulo: Parábola, 2011.

BASTOS, Juliete; ARAÚJO, Silvana. A concordância verbal na comunidade rural afro-brasileira: Mussuca-Laranjeiras-SE. *In*: ARAÚJO, Silvana; SILVA, Manoel; ARAÚJO, Jean. (org.). *Lusofonia afro-brasileira*: questões sócio-históricas e linguísticas. Campinas, SP: Editora Pontes, 2025, pp. 221-256.

BORTONI-RICARDO, Stella Maris. *Educação em língua materna*: a sociolinguística na sala de aula. São Paulo: Parábola Editorial, 2004.

BORTONI-RICARDO, Stella Maris. *Nós cheguemu na escola, e agora?* sociolinguística & educação. São Paulo: Parábola Editorial, 2005.

BORTONI, Marcia; ALVES, Scheyla. O fenômeno da hipercorreção. *In*: BORTONI RICARDO, Stella Maris; SOUSA, Rosineide; FREITAS, Vera; MACHADO, Veruska. (org.). *Por que a escola não ensina gramática assim?* São Paulo: Parábola Editorial, 2014, pp. 129-159.

FEITOSA, Jailma; VITÓRIO, Elyne. Variação nós/a gente no sertão alagoano: restrição e avaliação. *A Cor das Letras*, Feira de Santana, v. 19, n. 2, pp. 199-211, 2018.

FOEGER, Camila. *A primeira pessoa do plural no português falado em Santa Leopoldina*. 2014. Dissertação (Mestrado em Linguística) – Universidade Federal do Espírito Santo, Vitória, 2014.

FREITAG, Raquel. *Como fazer meta-análise com dados sociolinguísticos?* [*S. l.*]: GitHub, 2021. Disponível em: https://rkofreitag.github.io/meta.html. Acesso em: 15 out. 2024.

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

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

MAIA, Francisca. *A variação "nós" / "a gente" no dialeto mineiro*: investigando a transição. 2003. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2003.

MENDES, Rute. *O perfil da alternância do sujeito "nós" e "a gente" em Santo Antônio de Jesus*: um recorte no português popular do interior da Bahia. 2007. Dissertação (Mestrado em Letras e Linguística) – Faculdade de Letras, Universidade Federal da Bahia, Salvador, 2007.

OLIVEIRA, Almir; NASCIMENTO, Ruth. A alternância pronominal entre *nós* e *a gente* em Palmeira dos Índios. *In*: SANTOS, Maria Francisca; OLIVEIRA, Almir. (org.). *Linguagem, uso e ensino*. Arapiraca: EDUNEAL, 2017, pp. 15-29.

OMENA, Nelise. A referência à primeira pessoa do discurso no plural. *In*: OLIVEIRA e SILVA, Giselle; SCHERRE, Marta. (org.). *Padrões sociolinguísticos*: estudos de fenômenos variáveis do português falado do Rio de Janeiro. Rio de Janeiro: UFRJ Editora, 1996.

PINTO, Letícia; BERLINCK, Rosane. "O que que nói vamos falá?": significados sociais na variação/mudança da expressão de 1ª pessoa plural em duas comunidades rurbanas mineiras. *Organon*, Porto Alegre, v. 37, n. 73, pp. 36-58, jan./jun. 2022.

RAMOS, Conceição; BEZERRA, José; ROCHA, Maria de Fátima. Do nosso cotidiano ou do cotidiano da gente? Um estudo da alternância nós/a gente no português do Maranhão. *Revista Signum*. Londrina, v. 12, n. 1, pp. 279-292, jul. 2009.

R CORE TEAM. *R*: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2022. Disponível em: https://www.r-project.org. Acesso em: 15 out. 2024.

RUBIO, Cássio. A importância da metodologia no estudo da alternância pronominal e da concordância verbal de primeira pessoa do plural. *Cuadernos de la Alfal*, n. 7, pp. 90-106, 2015.

RODRIGUES, Fernanda. *Concordância verbal com o pronome a gente no sertão alagoano***.** Trabalho de Conclusão de Curso (Graduação em Letras Português) – Universidade Federal de Alagoas, Maceió, 2018.

SAMPAIO, Rosana; MANCINI, Marisa. Estudos de revisão sistemática: um guia para a síntese criteriosa da evidência científica. *Revista Brasileira de Fisioterapia*. São Paulo, v. 11, pp. 83-89, jan./fev. 2007.

SANTANA, Mardiny; VITÓRIO, Elyne. *A inserção de a gente em contextos sintáticos de complementação e adjunção na zona rural de Pariconha – AL*. 2022. 15 f. Relatório Final de Pesquisa PIBIC (Graduação em Letras Português) – Universidade Federal de Alagoas, 2022.

SILVA, Anilton. *Concordância verbal com o pronome nós na zona rural de Pariconha – AL*. Dissertação (Mestrado em Linguística e Literatura) – Universidade Federal de Alagoas, Maceió, 2021.

SILVA, Layane. *A variação pronominal de primeira pessoa do plural na zona rural de Pariconha – AL*. Dissertação (Mestrado em Linguística e Literatura) – Universidade Federal de Alagoas, Maceió, 2022.

SILVA. Anilton; VITÓRIO. Elyne. A concordância verbal com o pronome nós no Sertão alagoano. *Revista Porto das Letras*, v. 6, n. 1. pp. 50-73. 2020.

SILVA, Manoel; VITÓRIO, Elyne. A implementação de a gente não sujeito no sertão alagoano. *Revista A Cor das Letras*, v. 19, pp. 184-198, 2018.

SIQUEIRA, Manoel; SOUSA, Marta. O paradigma pronominal na fala de estudantes sergipanos. *Travessias Interativas*, São Cristóvão (SE), v. 14, n. 31, pp. 83-102, 2024.

SOUZA, Maria Helena. *A Variação nós e a gente na posição de sujeito na comunidade quilombola Serra das Viúvas / Água Branca – AL*. 2020. Dissertação (Mestrado em Linguística e Literatura) – Universidade Federal de Alagoas, Maceió, 2020.

SOUZA, Maria Helena. A concordância verbal com o pronome nós na comunidade quilombola Serra das Viúvas/Água Branca – AL. *Uniletras*, Ponta Grossa, v. 43, pp. 1-13, 2021.

SOUZA, Maria Helena; VITÓRIO, Elyne. Variação nós e a gente na posição de sujeito na comunidade quilombola Serra das Viúvas. *Letras em Revista*, Teresina, v. 12, n. 1, pp. 121-138, 2021.

TAMANINE, Andréa. *Curitiba da gente*: um estudo sobre a variação pronominal nós/a gente e a gramaticalização de a gente na cidade de Curitiba – PR. 2010. 222 f. Tese (Doutorado em Letras) – Setor de Ciências Humanas, Letras e Artes, Universidade Federal do Paraná, Curitiba, 2010.

VIANNA, Juliana; LOPES, Célia. A competição entre nós e a gente nas funções de complemento e adjunto: desvendando outras portas de entrada para o pronome inovador. *Caligrama*, Belo Horizonte, v. 17, n. 2, pp. 137-161, 2012.

VIANNA, Juliana; LOPES, Célia. Variação dos pronomes “nós” e “a gente”. *In.* MARTINS, Marco; ABRAÇADO, Jussara. *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015.

VITÓRIO, Elyne. A variação *nós* e *a gente* na posição de sujeito na fala de crianças da cidade de Maceió/AL. *Revista (Con)Textos Linguísticos*, Vitória-ES, v. 9, n. 14, p. 126-141, 2015.

VITÓRIO, Elyne. Variação *nós* e *a gente* na fala culta da cidade de Maceió/AL. *Interdisciplinar*, Sergipe, v. 24, pp. 159-172, 2016.

VITÓRIO, Elyne. A realização dos pronomes *nós* e *a gente* na função de sujeito e nas funções de complemento e adjunto na cidade de Maceió/AL. *Letrônica*, Porto Alegre, v. 10, n. 1, pp. 122-138, 2017.

VITÓRIO, Elyne. A Língua Usada no Sertão Alagoano: constituição da amostra. *In*: OLIVEIRA, Almir; PAULA, Aldir. (org.). *Interfaces Sociolinguísticas*: análises variacionistas em Alagoas e Pernambuco. Arapiraca: Eduneal, 2020.

VITÓRIO, Elyne. A difusão de *a gente* sujeito na variedade alagoana: um estudo de meta-análise. *Revista Diadorim*, v. 26, n. 2, pp. 1-21, 2024.

VITÓRIO, Elyne. Monitoramento estilístico e avaliação de nós e a gente: o que pensam estudantes universitários do alto sertão alagoano. *In*: OLIVEIRA, Almir; SANTOS, Dariana; LIMA, Jeylla. *O português brasileiro falado em Alagoas*. Arapiraca: Eduneal, 2025, pp. 151-172.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72078

Crenças e atitudes da elite letrada frente à variação nas estratégias de relativização

Beliefs and attitudes of literate elite regarding variation in relativization strategies

Dante Lucchesi[1]

RESUMO

São analisados os resultados de dois testes de avaliação subjetiva de variáveis linguísticas aplicados a 36 indivíduos com ao menos a graduação completa, distribuídos pelos dois sexos e três faixas etárias das cidades de Salvador e do Rio de Janeiro. No primeiro teste, o participante deveria avaliar frases contendo variáveis morfossintáticas do português brasileiro, como típicas de uma linguagem: formal demais, formal, informal ou incorreta. No segundo teste, o participante deveria escolher, em cada frase apresentada, uma variante que fosse adequada para ser usada em uma situação de comunicação formal, na primeira parte, e numa situação de comunicação informal, na segunda parte. O participante também deveria rejeitar alguma forma que considerasse imprópria para o uso naquela situação. Foram analisados apenas os resultados das frases que focalizavam as três estratégias de relativização do português brasileiro: padrão, cortadora e resumptiva. Os resultados revelaram que os participantes, todos membros da elite letrada do país, valorizam bastante a variante prescrita pela tradição gramatical, apesar dessa estratégia estar praticamente ausente de sua linguagem corrente, o que revela a força da ideologia normativista. A estratégia cortadora, no geral, não sofreu uma avaliação negativa, o que aconteceu com a estratégia resumptiva. Por outro lado, o estudo também identificou dois perfis que distinguiram os participantes do teste: um conservador, mais alinhado ao padrão normativo, e um mais progressista, não alinhado ao padrão normativo. Por fim, o estudo revelou as características sociais de cada um desses dois perfis.

PALAVRAS-CHAVE: Elite letrada. Orações relativas. Crenças e atitudes. Norma padrão. Tradição gramatical.

ABSTRACT

This study analyzes the results of two subjective evaluation tests concerning linguistic variables, applied to 36 individuals with at least a completed undergraduate degree. The participants were distributed by both sexes and three age groups in the cities of Salvador and Rio de Janeiro. In the first test, participants were asked to evaluate sentences containing morphosyntactic variables of Brazilian Portuguese, classifying them as typical of a language that is: “too formal,” “formal,” “informal,” or “incorrect.” In the second test, participants were required to choose, for each presented sentence, a variant suitable for use in a formal communication situation (first part) and in an informal communication situation (second part). They were also asked to reject any form they deemed improper for use in that situation. Only the results from sentences focusing on the three relativization strategies in Brazilian Portuguese were analyzed: the standard strategy, the chopping strategy, and the resumptive pronoun strategy. The results revealed that the participants, all members of the country’s literate elite, highly value the variant prescribed by grammatical tradition, despite this strategy being virtually absent from their current language use. This highlights the strength of prescriptive ideology. The chopping strategy, in general, did not receive a negative evaluation, unlike the resumptive pronoun strategy. On the

[1] Professor Permanente do Programa de Pós-Graduação em Língua e Cultura da Universidade Federal da Bahia (PPGLinC-UFBA), Bolsista de Produtividade em Pesquisa, Nível 1B, CNPq (proc. nº 310919/2023-1). dlucra@hotmail.com, https://orcid.org/0000-0002-8058-2658.

Recebido em 12/01/2026 - Aceito em 08/06/2026

other hand, the study also identified two distinct profiles among the test participants: a conservative one, more aligned with the normative standard, and a more progressive one, not aligned with the normative standard. Finally, the study revealed the social characteristics associated with each of the two profiles.
KEYWORDS: Literate elite. Relative clauses. Language beliefs and attitudes. Standard norm. Grammatical tradition.

1. Introdução

O modelo da Sociolinguística Variacionista foi introduzido no Brasil por Anthony Naro, no final da década de 1970, constituindo o PEUL – *Programa de estudos sobre o uso da língua*, na UFRJ. Com os fundamentos teóricos e metodológicos da Sociolinguística Variacionista (Weinreich; Labov; Herzog, 2006 [1968]; Labov 2008 [1972]), o PEUL dedicou-se, desde o início da década de 1980, à análise dos condicionamentos linguísticos e sociais dos principais fenômenos variáveis do português brasileiro, no plano da fonologia e da morfossintaxe (Paiva; Gomes, 2016) e tornou-se o modelo para praticamente todos os programas de pesquisa sociolinguística no Brasil, como o VARSUL – *Variação linguística urbana da Região Sul*[2] e o VALPB – *Variação linguística no Estado da Paraíba*[3], que desenvolveram análises e constituíram amostras de fala baseadas nos mesmos parâmetros do PEUL. Assim, nas últimas décadas do século passado, a pesquisa sociolinguística no Brasil concentrou-se no que Weinreich, Labov e Herzog (2006 [1968]) denominaram *problema do encaixamento*, ou seja, como o fenômeno variável se encaixa na matriz mais ampla da estrutura linguística e da estrutura social, sendo poucos os estudos empíricos dedicados ao chamado *problema da avaliação*, ou seja, como os utentes da língua percebem e reagem diante das formas linguísticas em variação.

No entanto, o interesse e os estudos que focalizam o problema da avaliação têm crescido ao longo deste século, descortinando um campo fértil para o entendimento de outros fatores que determinam a implementação da mudança linguística e se situam no plano das subjetividades e das crenças difundidos na sociedade acerca da língua. No bojo desse processo, o Projeto *Vertentes do Português no Estado da Bahia*[4], em sua 4ª Etapa, dedicada ao estudo da norma culta na cidade de Salvador, aplicou testes de avaliação subjetiva de variantes linguísticas e adequação à situação de fala, em um conjunto de 24 soteropolitanos com ao menos o curso superior completo, para reunir dados sobre as atitudes e crenças dos membros da elite letrada em relação a diversos fenômenos variáveis do português brasileiro, no plano da morfossintaxe. A análise das respostas desses testes sobre frases contendo variantes de orações relativas constitui o objeto deste artigo, que pretende contribuir, não apenas com a compreensão do processo de mudança que afeta esse aspecto da gramática do português brasileiro, mas também com a produção de subsídios para a necessária atualização da norma de referência linguística do Brasil (Lucchesi, 2015).

---

[2] *Cf*. https://varsul.org.br/index.php?_lng=br. Acesso em: 21/07/2026.

[3] *Cf*. https://projetovalpb.com.br/. Acesso em: 21/07/2026.

[4] *Cf*. https://vertentes.ufba.br. Acesso em: 21/07/2026.

Para cumprir esses objetivos, este artigo se estrutura da seguinte maneira. A seção 2 descreve os instrumentos metodológicos empregados nesta análise. A seção 3 contém um breve panorama dos estudos sobre as orações relativas no português brasileiro. A seção 4 apresenta a análise dos resultados obtidos na aplicação dos testes. Nas considerações finais, são feitas uma síntese e uma análise qualitativa dos resultados e são delineados possíveis desdobramentos desta pesquisa.

2. Os instrumentos usados nesta pesquisa

Depois de dedicar sua pesquisa empírica, em suas três primeiras fases, ao português popular, o Projeto Vertentes do Português no Estado da Bahia, sediado na UFBA, dedica a quarta etapa de suas pesquisas aos padrões de comportamento linguístico da elite letrada da cidade de Salvador.[5] Para levar a cabo essa pesquisa, foram gravadas, entre os anos de 2022 e 2024, 24 entrevistas de tipo de sociolinguístico, com 40 a 60 minutos de duração, com soteropolitanos, distribuídos pelos dois sexos e por três faixas etárias: de 25 a 35 anos; de 45 a 55 anos; e de 65 a 75 anos. Além das variáveis sexo e idade, foram controladas as seguintes variáveis sociais, conquanto não tenham sido consideradas na estratificação da amostra: etnia, opondo brancos a pretos e pardos; nível de escolaridade, distinguindo os que só cursaram a graduação dos que cursaram uma pós-graduação stricto sensu; e nível de escolaridade dos pais, considerando pais com curso superior e sem curso superior. A distribuição dos participantes por essas três variáveis é a seguinte: 17 participantes brancos, 6 participantes pretos e um pardo; 10 participantes cursaram somente a graduação, 2 fizeram uma especialização, 3 concluíram o mestrado e 9, o doutorado; 10 participantes tinham pais com escolaridade elevada, enquanto 14 tinham pais com escolaridade média ou baixa.[6]

Ao lado das entrevistas, também foram gravadas ou recebidas elocuções formais desses mesmos participantes, com duração de 20 a 30 minutos. Isso possibilitará análises de variação estilística, o que também fornece evidências empíricas para descrever a valoração subjetiva dos falantes das variantes linguísticas.

Para aferir as crenças e atitudes dos falantes frente a varáveis do nível da morfossintaxe, foram aplicados dois testes com os 24 participantes da pesquisa.

O primeiro teste, de *Avaliação Direta das Variantes Linguísticas* (doravante *Teste de Avaliação*), foi respondido em um suporte escrito, no qual o participante deveria ler incialmente a seguinte instrução:

> As pessoas constroem as frases que usam para se comunicar de acordo com o nível de formalidade da situação em que se encontram e buscam seguir os preceitos da correção gramatical nas situações mais formais de comunicação. Nas situações de informalidade,

[5] Os membros da elite letrada são os indivíduos com um alto grau de escolaridade, correspondendo, assim, à definição de "falante culto", do Projeto NURC, ou seja, o indivíduo com ao menos um curso superior completo.

[6] Entre os falantes das faixas etárias de 25 a 35 anos e de 45 a 55 anos, foram considerados filhos de pais com escolaridade elevada, quando o pai e a mãe tinham ao menos o curso superior completo. Entre os mais velhos, de 65 a 75, foram considerados filhos de pais com escolaridade elevada, quando ao menos um genitor tinha o curso superior completo, e o outro o ensino médio concluído. Os demais casos foram avaliados como de escolaridade média e baixa, que foram reunidos na tabulação dos resultados em função da análise dos resultados iniciais da análise quantitativa (*cf*. seção 4).

> como um bate papo com os amigos, as pessoas não se preocupam muito com as regras da gramática, mas, mesmo nessas situações, há frases que são rejeitadas por conterem erros gramaticais muito graves. Com base nessas premissas, classifique as frases deste teste em uma dessas categorias: **formal demais**, para frases que soam pedantes por usarem regras muito rebuscadas da gramática; **formal,** para frases que devem ser usadas em situações de comunicação formal, de acordo com as regras da Gramática**; informal**, para frases usadas em situações de descontração, que não seguem necessariamente as regras da gramática, mas não contêm nenhum erro grave; e **incorreta,** para frases que contêm erros gramaticais graves, que são notados mesmo nas situações de informalidade. Em seguida, destaque o elemento da frase que o levou a fazer essa avaliação.

A seguir, o participante deveria classificar, segundo as opções apresentadas na introdução, 36 frases que instanciavam diversas variáveis morfossintáticas, dentre as quais a forma das orações relativas. A primeira frase desse teste contendo uma oração relativa é a seguinte, com as opções oferecidas ao participante:

1. Maria é uma pessoa que ela procura agradar todo mundo sempre.

Em relação à gramática, e não ao vocabulário, essa frase é típica de uma linguagem:

( ) formal demais ( ) formal ( ) informal ( ) incorreta.

O segundo teste, de *Adequação à situação de fala* (doravante *Teste de Adequação*), foi introduzido pelo seguinte texto:

> Não falamos da mesma maneira em todas as situações, porque estamos sempre, de forma mais ou menos consciente, ajustando nossa linguagem, em função da situação em que nos encontramos, da pessoa com quem estamos falando e mesmo do assunto sobre o qual estamos falando. Considerando esse princípio, leia as frases a seguir e escolha a opção de linguagem mais adequada, em cada situação. E, se houver, risque a frase que você nunca usaria, indicando a razão para isso.

Após a leitura desse texto, o participante era instado a escolher uma variante de uma variável da morfossintaxe em um conjunto de 40 frases, divididas igualmente em duas situações de fala: na primeira, o participante deveria se imaginar fazendo uma exposição em uma reunião de trabalho, ou seja, uma situação de fala formal; na segunda, o participante deveria se imaginar em uma mesa de bar, conversando descontraidamente com seus amigos, ou seja, uma situação de fala informal. A primeira frase apresentada ao falante, na situação formal, foi esta:

Você está fazendo uma exposição sobre um tema técnico em uma reunião formal de trabalho. Marque com um X a frase que você usaria em cada grupo abaixo e risque a frase que você não usaria (só se houver alguma).

( ) Lavoratex é o produto que vamos iniciar a campanha.
( ) Lavoratex é o produto com o qual vamos iniciar a campanha.
( ) Lavoratex é o produto que vamos iniciar a campanha com ele.

Por que você não falaria assim?
( ) Porque é muito informal.
( ) Porque é gramaticalmente incorreto.
( ) Porque é formal demais.

Nos dois testes, buscou-se colocar o participante diante de todas as estratégias de relativização possíveis no português brasileiro, para documentar empiricamente a reação dos falantes da elite letrada perante todas as formas em variação. Os resultados da tabulação desses testes serão apresentados e analisados na seção 4 deste artigo. Para além dos 24 testes aplicados no âmbito do Projeto Vertentes, na cidade de Salvador, serão também usados os resultados desses mesmos testes, que foram aplicados, no ano de 2023, com 12 membros da elite letrada da cidade do Rio de Janeiro, no âmbito do Projeto Padrões Urbanos Cultos da Cidade do Rio de Janeiro (PUC-RJ), desenvolvido na Universidade Federal Fluminense (UFF). Portanto, a base empírica desta pesquisa é constituída pelos resultados de 36 testes de avaliação e adequação, aplicados com 24 soteropolitanos e 12 cariocas.

3. A variação na forma das orações relativas e as prescrições gramaticais

Em seu estudo clássico, Fernando Tarallo (1983) adotou a taxonomia de três estratégias de relativização que tem sido seguida por todos os estudos variacionistas desde então. A chamada *estratégia padrão* (prescrita pela tradição gramatical) engloba três possiblidades: na primeira; a oração relativa (OR) é introduzida pelo relativizador *que*, ou pelo pronome relativo *o qual*, regido por uma preposição; na segunda, a OR é introduzida por uma palavra interrogativa (*onde*, *quem*, *quando* etc.), regida ou não por uma preposição; na terceira, a OR é introduzida pelo pronome relativo *cujo* e suas formas flexionadas (*cuja*, *cujos*, *cujas*), que se referem ao caso genitivo. Na chamada *estratégia cortadora*, a preposição que rege o relativizador *que* na estratégia padrão é suprimida. Por fim, na estratégia resumptiva, um pronome lembrete preenche a lacuna que existe nas estratégias padrão e cortadora, e a OR é sempre introduzida pelo relativizador *que*. A variável estratégia de relativização só tem essa configuração ternária nas ORs extraídas de posições sintáticas preposicionadas. Nas ORs extraídas das posições de sujeito e objeto direto, só haveria duas variantes: a estratégia com pronome lembrete, e uma estratégia com a OR introduzida normalmente pelo relativizador *que* e muito esporadicamente pelo pronome relativo *o qual* com suas flexões.

Com base nos resultados quantitativos da análise de sua amostra de fala e nos resultados de seus pioneiros dos testes de avaliação subjetiva, Tarallo (1983) chegou à conclusão de que a estratégia padrão estaria caindo em desuso, sendo substituída pela estratégia cortadora, que no geral não seria avaliada negativamente pelos falantes, enquanto a estratégia resumptiva teria seu emprego muito limitado pela reação negativa dos falantes ao seu uso.

As gramáticas mais tradicionais, elaboradas em meados do século passado, como a de Rocha Lima (1960) e Cegalla (1977), simplesmente ignoravam a variação nas estratégias de relativização.

As gramáticas com uma abordagem mais atualizada e que buscam incorporar os avanços da pesquisa linguística, já se referem à variação na forma das ORs. Azeredo (2011, p. 317), por exemplo, reconhece que:

> Na fala espontânea, nas variedades populares da língua e na escrita de pessoas com baixa escolarização, [...] praticamente só se emprega a forma *que*, que tende a perder a condição de forma substituta – pronome anafórico – de um antecedente em proveito da função puramente conectiva. Daí variações como *O vestido **que** ela gostou/O vestido **de que** (do qual) ela gostou, Tenho um amigo **que** o pai (dele) é marceneiro/ Tenho um amigo **cujo** pai é marceneiro, O caminho **que** eu vim é tranquilo/O caminho **pelo qual** (por onde) eu vim é tranquilo*.[7]

E reconhece ainda que "se trata de um uso que já migrou definitivamente para a escrita menos formal, como a crônica jornalística e da literatura memorialística", dando exemplos, inclusive, de relativas resumptivas em crônicas de jornais de grande circulação, como *O Globo* (Azeredo, 2011, p. 317).

Já outros gramáticos, como Evanildo Bechara, apesar de procurarem incorporar os avanços da análise linguística e reconhecerem a variação, mantêm uma atitude prescritiva, reforçando o uso da estratégia padrão, como se pode ver na seguinte passagem:

> Constitui impropriedade o emprego do pronome relativo precedido de preposição *de* como adjunto adnominal, em lugar de *cujo*. Assim, se evita dizer "o escritor *de que* todos conhecemos o livro", "as pessoas *de quem* reconheceis os privilégios". Em boa linguagem diremos "o escritor *cujo* livro conhecemos", "as pessoas *cujos* privilégios reconheceis".[8] (Bechara, 2001, p. 486)

Nas últimas décadas, linguistas também têm elaborado gramáticas, que assumem um caráter descritivo (Castilho, 2010), ou mesmo pedagógico (Bagno, 2012). Castilho (2010), em linhas gerais, reproduz a avaliação de Tarallo (1983): "Na língua falada, *que* é uma espécie de pronome relativo universal, que está ocupando os espaços dos outros", de modo que "o conjunto dos pronomes relativos vem sofrendo séria restrição no PB" (p. 366 e 367). Castilho (2010, p. 367) destaca "o desaparecimento de *cujo* [...], mesmo no padrão culto, variedade em que esse pronome é substituído por *de que*"; e chama a atenção para "o uso crescente do advérbio *onde*, que substituiria a estrutura padrão *em que*", sendo usado, tanto com antecedentes que se referem a lugar, quanto com os que se referem a tempo.

Bagno (2012, p. 168-169) também proclama o desaparecimento do relativo *cujo* na língua falada e afirma que "mesmo na escrita monitorada, [esse] pronome tende a desaparecer", apresentando exemplos correlatos de relativas cortadoras e resumptivas, extraídos de textos jornalísticos e literários publicados da Internet. Essa situação se estenderia a todos os pronomes relativos, que estariam cada vez mais restritos "aos gêneros escritos mais monitorados", mas "mesmo aí já encontramos indícios que anunciam a provável extinção futura desses pronomes" (Bagno, 2012, p. 900).

---

[7] Grifos do original.

[8] Itálicos do original.

Tratando das "atitudes dos falantes", Bagno (2012, p. 914) converge para as conclusões de Tarallo (1983), embora não apresente dados empíricos que fundamentem suas afirmações:

> Para os falantes urbanos mais letrados, a relativa padrão é pouco aceita porque parece ser "certa demais", ao passo que a relativa copiadora, identificada (sem razão, porém) com os falantes menos instruídos, é menos aceita porque parece "errada demais". Assim, para manter uma posição equidistante entre esses dois extremos, o brasileiro culto opta por empregar a relativa cortadora.

Embora defenda o ensino da estratégia padrão, Bagno (2012, p. 910-911) pondera que "no caso específico do *cujo*, a existência de uma forma lexical própria torna mais fácil (mas nem por isso mais exitoso) o ensino do seu uso". Já quanto às relativas padrão com um pronome regido por uma preposição, a situação seria "um pouco mais complexa", porque "as formas que elas podem assumir em combinação com as preposições são muito numerosas". E conclui: "o ensino da estratégia padrão, repito, é um grande desafio pedagógico – e de êxito mais do que duvidoso" (Bagno, 2012, p. 914).

Os resultados da pesquisa de Corrêa (1998), com estudantes do ensino fundamental e médio, parecem ir ao encontro das afirmações de Bagno:

> Se uma das finalidades do ensino institucional é veicular a língua culta, podemos, com estes dados, concluir que isso não vem ocorrendo com as relativas. Tanto nos dados orais como nos escritos, os informantes não escolarizados e os alunos da 1ª a 8ª série produzem quase que exclusivamente relativas cortadoras. (Corrêa, 1998, p. 82)

Contudo, seus resultados com alunos do ensino médio indicam que a escola estimula o uso da estratégia padrão, que passa "de 7% na séria inicial para 44% na final, enquanto a estratégia vernacular [cortadora] decresceu de 93% para 56%, um resultado que confirma a importância do ensino de 2º grau para aquisição da relativa preposicionada" (Corrêa, 1998, p. 90). Em contrapartida, os dados obtidos com a aplicação de um teste com 27 professores do ensino médio foram muito sintomáticos. No teste são apresentadas ao professor 10 frases contendo ORs cortadoras e copiadoras e se lhe pergunta se essas frases são aceitáveis na língua escrita. Conquanto tenham rejeitado totalmente duas frases do conjunto de dez (uma cortadora e uma copiadora), outras duas frases com uma relativa resumptiva com a retomada do antecedente pelo pronome possesivo – *Xadrez é um jogo que nunca pude aprender suas regras* e *Língua extinta é aquela que não possuímos prova da sua existência* – foram aceitas respectivamente por 9 e 12 dos participantes (ou seja, 33% e 44% do total), enquanto uma resumptiva propriamente dita (*Uma boa causa é aquela que você faz parte dela*) foi aceita por 5 participantes (ou seja, 18,5%) do total (Corrêa, 1998, p. 102-106). Já as demais cortadoras tiveram uma aceitação entre 11 e 26% do total. Além disso, Corrêa observou que: "a resposta negativa não garantiu a versão padrão da relativa. Para vários professores bastou trocar *que* por *o qual*. Tão pouco foi garantia de que a preposição apareceria com o pronome relativo". A autora também reportou o seguinte depoimento de um professor: "só quem possui um certo conhecimento de gramática é que pode identificar as falhas, visto que os erros são imperceptíveis nas orações adjetivas" (Corrêa, 1998, p. 105).

Nesse sentido, deve-se ter em conta que, em 1998, um importante documento do Ministério da Educação, os *Parâmetros Curriculares Nacionais*, já reconhecia como legítimo o emprego da estratégia cortadora, no bojo de uma condenação a um ensino dogmático baseado na ideia de que só haveria uma única variedade de língua legítima e aceitável, sendo as demais impuras e corrompidas. Desde então o ensino de língua portuguesa tem privilegiado, cada vez mais, além do ensino da literatura, a leitura e a produção de textos e o reconhecimento e classificação dos gêneros textuais, em detrimento do ensino da norma padrão e da análise da estrutura gramatical das frases da língua. Deve-se frisar também que a prova de português do Exame Nacional do Ensino Médio, Enem – o principal canal para acesso dos estudantes brasileiros ao ensino superior, nomeadamente às universidades públicas (as mais concorridas) –, não contém questões que cobrem o domínio da norma padrão, nem de análise gramatical. Diante desse quadro, uma pesquisa empírica deve observar não somente a atitude e crenças dos falantes diante das variantes linguísticas, mas a própria capacidade desses falantes de reconhecer a variante padrão, mesmo no caso de falantes altamente escolarizados, como é caso desta pesquisa.

4. Análise dos resultados

Com base no que foi discutido na seção anterior, a análise dos resultados dos testes desta pesquisa verificará as seguintes hipóteses acerca das crenças e atitudes dos falantes ditos cultos frente às estratégias de relativização:

(i) a estratégia padrão é avaliada como excessivamente formal;

(ii) a estratégia cortadora não sofre avaliação explícita;

(iii) a estratégia resumptiva é avaliada negativamente.

Com esse direcionamento, serão analisados, na primeira subseção, os resultados do *Teste de Avaliação*, ao passo que, na segunda subseção, serão apresentados e avaliados os resultados do *Teste de Adequação*. E na última subseção, serão comparados os juízos dos participantes com o seu desempenho linguístico real.

4.1. As orações relativas nos Testes de Avaliação

O *teste de avaliação direta* das variantes linguísticas continha 08 frases com ORs, sendo três com a estratégia padrão, três cortadoras e duas com o pronome resumptivo. O participante deveria julgar cada frase como representante de uma linguagem: *formal demais, formal, informal* ou *incorreta*. Como definido no texto introdutório do teste, as frases de uma linguagem *formal demais* são aquelas que "soam pedantes por usarem regras muito rebuscadas da gramática". As frases *formais* são aquelas "que devem ser usadas em situações de comunicação formal, de acordo com as regras da Gramática". As *informais* são "usadas em situações de descontração, que não seguem necessariamente as regras

da gramática, mas não contêm nenhum erro grave". E as *incorretas* "contêm erros gramaticais graves, que são notados mesmo nas situações de informalidade".

As três frases contendo a estratégia padrão são as seguintes:

Padrão de adjunto adverbial de tempo: Na época ***em que*** *tudo começou*, ele era muito novo.

Padrão de adjunto adverbial de instrumento: A chave ***com a qual*** *você está querendo apertar o parafuso* está ruim.

Padrão de genitivo: O trabalho ***cujo*** *tema era a intolerância religiosa* foi o escolhido.

É preciso observar inicialmente que 19 participantes tiveram dificuldade em identificar a OR de ADV de tempo como a variante padrão: 18 avaliaram essa frase como *informal* e um, como *incorreta*. Por um lado, pode-se pensar em desconhecimento da forma padrão (*cf*. final da seção anterior); por outro, em um desvio interpretativo[9]. Além disso, a não identificação dessa forma padrão pode ser atribuída ao fato de essa variante ser introduzida pelo relativizador *que* regido pela preposição gramatical *em*, sendo, portanto, muito parecida com a variante cortadora, que é introduzida apenas pelo relativizador *que* sem a preposição. O fato de nenhum participante ter classificado essa frase como *formal demais* converge para essa explicação. Assim, um pouco menos da metade dos participantes classificou essa frase como típica da linguagem *formal* (47,2% do total).

Nas ORs extraídas das posições de ADV de instrumento e GEN, não houve esse tipo de problema, e os julgamentos se dividiram entre *formal* e *formal demais*. Quando o participante classifica uma frase como *formal demais*, pode-se interpretar que ele tem uma visão crítica do padrão, pois a instrução do teste orientava para classificar assim as "frases que soam pedantes por usarem regras muito rebuscadas da gramática". Por outro lado, ao classificar uma frase como *formal*, o falante reconhece essa variante como apropriada para ser usada na escrita e em situações de fala formal, como definido na instrução do teste reportada acima.

Considerando esses parâmetros, os participantes foram mais críticos em relação a OR de ADV de instrumento, com 14% do total classificando essa frase como formal demais, enquanto a OR de GEN foi avaliada dessa forma por apenas um pouco menos de 6% do total. Assim, ao que parece, os falantes da elite letrada vêm como uma forma mais empolada a OR o pronome relativo *o qual*, do que o *cujo*. Esse resultado não deixa de surpreender, em face da posição de linguistas que apontam o *cujo* como uma forma ultrapassada, obsoleta ou artificial (*cf*. seção anterior). Não parece ser esse o juízo dos falantes altamente escolarizados, devendo-se ressaltar, sobretudo, que, na grande maioria dos julgamentos, a variante padrão é considerada apropriada para ser usada nas situações de formalidade, tanto na fala, quanto na escrita, mesmo que isso não se coadune com seu uso concreto da língua, pelo menos na fala (*cf*. subseção 4.3), pois somente em 6,5% dos julgamentos a estratégia padrão foi avaliada como formal demais.

[9] O participante pode não ter lido a frase com atenção ou pode ter sido influenciado pelo léxico da frase ou por outro aspecto gramatical.

Focalizando a minoria de participantes que tiveram uma atitude mais crítica em relação ao padrão normativo, classificando a estratégia padrão nos dois casos como *formal demais,* os resultados percentuais revelaram que: as mulheres foram mais críticas, com 9,3% delas avaliando a estratégia padrão como *formal demais*, contra 3,7% dos homens; os pretos e pardos, com 10% de formal demais, praticamente o dobro do percentual entre os brancos; os participantes que cursaram apenas a graduação, com 8,9% do total rejeitando o padrão, contra apenas 4,8% dos que cursaram o mestrado e o doutorado; e os falantes com pais com um nível de escolaridade alto, em face dos participantes filhos de pais com nível de escolaridade baixo ou médio, com os mesmos percentuais da variável anterior. A idade e a cidade não se mostraram significativas. Portanto, em relação ao uso da estratégia padrão, os homens, brancos, com pós-graduação e com pais com um nível de escolaridade mais baixo assumiram uma atitude mais conservadora, avaliando mais a estratégia padrão como apropriada para as situações de comunicação formal.[10]

As três frases contendo ORs com a estratégia cortadora no Teste de Avaliação são as seguintes:

Cortadora de genitivo: O computador *que o HD queimou* já foi substituído.

Cortadora de adjunto adverbial de companhia: Muitas vezes, você não conhece a pessoa *que você convive todo dia.*

Cortadora de adjunto adverbial de lugar: No lugar que eles estavam, chovia demais.

A análise das ORs cortadoras buscou verificar se essa variante não era rejeitada pelos falantes altamente escolarizados como uma forma imprópria para o uso, mesmo em situações informais de comunicação. Assim, a expectativa foi a de que os participantes avaliassem essas frases como *informais,* e, com efeito, um pouco mais de ¾ dos participantes avaliaram essas frases dessa forma (77,8% do total), e menos de ¼ as avaliaram como *incorretas* (22,2% do total). As ORs extraídas das posições de adjunto adnominal (genitivo) e de adjunto adverbial de companhia foram as mais rejeitadas, por um pouco mais de ¼ dos participantes (27% e 28,6% do total, respectivamente), enquanto a OR de ADV de lugar foi rejeitada por apenas 11,1% do total.

Considerando o perfil social das avaliações, não houve diferença entre os homens e mulheres; a aceitação das cortadoras em ambos foi idêntica à frequência geral (77,8% do total). A diferença entre brancos e pretos e pardos também não foi significativa. No tocante à faixa etária, as cortadoras foram mais aceitas pelos mais jovens, que só as rejeitaram em pouco mais de 15% dos julgamentos, havendo uma crescente rejeição dessa variante na medida em que se passou para as faixas dos falantes mais velhos, que se mostraram, como era de se esperar, mais conservadores (16,7%, na faixa de 25 a 35 anos; 22,2%, na faixa de 45 a 55 anos; 27,8%, na faixa de 65 a 75 anos). As cortadoras foram mais aceitas pelos que tinham pós-graduação, em quase 85% dos julgamentos (84,1% do total), contra quase 70% de aceitação entre os que tinham graduação (68,9% do total); bem como entre os participantes com pais com nível mais alto de escolaridade, que só rejeitaram essa variante em 15,6% dos julgamentos,

[10] A distribuição dos participantes pelas variáveis etnia, escolaridade do participante e escolaridade dos seus pais é apresentada no primeiro parágrafo da seção 2.

contra 26,9% de rejeição entre os participantes com pais com nível de escolaridade mais baixo. Em Salvador, as cortadoras também foram mais aceitas, com apenas 16,7% de julgamentos dessa variante como incorreta, enquanto um em cada três cariocas rejeitaram essa variante (33,3% do total). Portanto, em relação ao uso da estratégia cortadora, os participantes jovens, com pós-graduação, com pais com alto nível de escolaridade e soteropolitanos se mostraram mais flexíveis, enquanto os mais velhos, sem pós-graduação, com pais com nível mais baixo de escolaridade e cariocas se mostraram mais conservadores.

A estratégia resumptiva foi representada no teste por estas duas frases:

Relativa resumptiva de sujeito: Maria é uma pessoa *que* ***ela*** *procura agradar todo mundo sempre*.

Relativa resumptiva de adjunto adverbial de companhia: Eu também não gosto daquele rapaz *que você discutiu* ***com ele*** *na saída da Faculdade*.

No caso das frases com pronome lembrete, foi testada a hipótese de que os falantes da elite letrada avaliam negativamente essa construção, classificando-a como incorreta. Por outro lado, a avaliação dessas frases apenas como informais, não sendo vistas como um erro grave de gramática, indicaria uma postura mais flexível, de maior aceitação da variação linguística. E o cômputo geral dos julgamentos revelou um certo equilíbrio, com 55,6% rejeitando o uso do lembrete como incorreto, contra 44,4% que julgou a resumptiva apenas como típica da linguagem informal; o que não deixa de ser surpreendente porque contradiz a ideia de que essa variante seria amplamente rejeitada pelos falantes altamente escolarizados (*cf.* seção anterior). No que concerne à posição de relativização, a OR de ADV de companhia foi um pouco mais rejeitada do que a de sujeito (com 58,3% avaliando aquela como incorreta, contra 52,3% que avaliaram esta última assim), o que também surpreende, em face da ideia geral de que os pronomes lembretes são mais empregados nas posições de relativização menos acessíveis. Contudo, a maior aceitação da OR de sujeito, no caso da frase do teste, pode ser atribuída à predicação da oração matriz, que pode ser classificada como tipificadora (‘Maria é o tipo de pessoa’), um dos contextos que favorecem a estratégia resumptiva (Lucchesi, Santos, 2025).

Quanto ao perfil social dos participantes que tiveram uma atitude mais conservadora, condenando as resumptivas como incorretas, os resultados foram os seguintes. As mulheres rejeitaram mais o pronome lembrete do que os homens (63,9% contra 47,2%), já a diferença geracional não foi significativa, enquanto os brancos se mostraram mais uma vez mais conservadores que pretos e pardos, rejeitando as resumptivas em mais de 60% dos julgamentos (61,5% do total), contra apenas 40% dos pretos e pardos. Os que só cursaram a graduação foram mais conservadores do que os que cursaram a pós-graduação, com 70% de rejeição do pronome lembrete entre aqueles contra apenas 45,2% entre estes. Com os mesmos percentuais, as resumptivas foram mais condenadas pelos que eram filhos de pais com alta escolaridade, do que pelos que tinham pais com escolaridade baixa ou intermediária. Por fim, os baianos se mostraram um pouco mais conservadores do que os cariocas, com 58% de condenação das resumptivas contra 50% dos últimos. Portanto, tiveram uma atitude

mais flexível, avaliando a estratégia resumptiva apenas como informal, os homens, os pretos e pardos, os com pós-graduação, os que tinham pais com escolaridade baixa ou média, e os cariocas. Consequentemente, foram mais conservadores as mulheres, os brancos, os que só cursaram apenas a graduação, com pais com alta escolaridade e os soteropolitanos.

Computando os resultados das três estratégias de relativização, os brancos se mostraram mais conservadores do que os pretos e pardos, no julgamento das três estratégias. Já entre os sexos, as mulheres se mostraram mais críticas em relação à estratégia padrão, mas condenaram mais o uso do pronome lembrete. A diferença geracional só foi significativa em relação à estratégia cortadora, com os mais velhos rejeitando mais essa variante. Quanto à escolaridade, os com pós-graduação foram mais flexíveis em relação ao emprego das cortadoras e resumptivas, ao passo que os que só cursaram a graduação rejeitaram mais a estratégia padrão como formal demais. Já os participantes cujos pais tinham alta escolaridade avaliaram mais a estratégia padrão como formal demais, aceitaram mais as cortadoras, porém foram mais conservadores em relação às resumptivas. Em relação à cidade, os cariocas condenaram um pouco menos o pronome lembrete do que os baianos, mas estes aceitaram mais o uso das cortadoras. Assim, o panorama geral obtido com a aplicação do teste de avaliação direta não permitiu traçar um perfil social muito nítido da atitude e crenças dos falantes da elite letrada no que se refere à variação na forma das ORs, o que pode ser devido a dificuldades de leitura e interpretação das frases na hora de responder ao teste. Já o teste de adequação, no qual as variantes de cada estratégia são reunidas em cada frase forneceu um quadro mais nítido.

4.2. As orações relativas nos Testes de Adequação

Como dito na seção 2, o falante deveria escolher, no teste de adequação, uma estratégia de relativização, considerando duas situações distintas, uma formal (uma reunião de trabalho) e uma informal (uma conversa de mesa de bar). O falante também poderia rejeitar uma variante como inadequada. Como foi facultado ao participante escolher mais de uma variante como adequada à situação, bem como rejeitar mais de uma variante ou não rejeitar nenhuma, os resultados quantitativos nem sempre terão um total de 36, o número de participantes, mas terão o número total de escolhas e rejeições, que poderá ser maior ou menor que 36.

4.2.1. Crenças e atitudes dos falantes em uma situação de comunicação formal

Considerando incialmente a situação de uma discussão técnica em uma reunião de trabalho, a primeira frase do teste foi:

( ) Lavoratex é o produto que vamos iniciar a campanha.
( ) Lavoratex é o produto com o qual vamos iniciar a campanha.
( ) Lavoratex é o produto que vamos iniciar a campanha com ele.

Praticamente 4/5 dos participantes (80% do total) escolheram a segunda opção, da estratégia padrão introduzida pelo pronome relativo *o qual* regido pela preposição *com* (em função da ligação desse pronome com a posição de ADV de instrumento na OR), só 1/5 dos participantes (20% do total) escolheu a primeira opção, da variante cortadora, e nenhum escolheu a resumptiva, que era a terceira opção. Portanto, assim como no teste de avaliação direta, a maioria dos participantes expressou uma conformidade com o padrão normativo, considerando a estratégia padrão a mais adequada para ser usada nas situações de comunicação formal, mesmo na oralidade; reforçando, assim, uma forte dicotomia entre crença e comportamento real (*cf.* subseção 4.3. adiante).

Considerando o perfil social dos que tiveram uma atitude mais flexível, a escolha pela cortadora decresceu, como era de esperar, na medida em que se passou da geração dos participantes mais novos para os mais velhos: a cortadora correspondeu a 1/3 da escolha entre os falantes de 25 a 35 anos (33,3%), caindo para 16,7% da escolha dos que tinham entre 45 e 55 anos; e apenas 8,3, com os de 65 a 75 anos. As mulheres também se mostraram mais flexíveis, escolhendo a cortadora em 22,2% dos julgamentos, contra 16,7% dos homens. A opção pela cortadora foi quase quatro vezes maior entre pretos e pardos, com 40% contra apenas 11,5% entre os brancos. Os que cursaram apenas a graduação também escolheram mais a cortadora, com 26,7% contra apenas 14,3% dos que cursaram uma pós-graduação. Já a escolaridade dos pais não foi relevante, com praticamente o mesmo percentual, tanto para os com pais com alta, quanto com os pais com baixa ou média escolaridade (20 e 19%, respectivamente). Por fim, todos os participantes do Rio de Janeiro escolheram a relativa padrão, enquanto a cortadora foi a opção em quase 30% das escolhas em Salvador (29,2% do total). Considerando o que foi discutido no final da seção 3, especialmente no que concerne ao estudo de Corrêa (1998), a questão que se coloca é: a opção pela cortadora resulta de uma visão crítica em relação ao padrão ou de um desconhecimento da prescrição gramatical? Em face disso, uma inferência mais neutra seria: os homens, mais velhos, brancos, com pós-graduação, no Rio de Janeiro, estão mais alinhados com o padrão normativo, enquanto os mais jovens, as mulheres, pretos e pardos, apenas com a graduação, em Salvador, ignoram ou não adotam conscientemente esse padrão.

Os participantes também deveriam indicar uma ou mais variantes que consideravam inadequada para ser usada em uma situação de comunicação formal. Como previsto, a grande maioria rejeitou a OR resumptiva, 32 de 36 participantes (89% do total), sendo que, desses 32, 26 avaliaram essa construção como gramaticalmente incorreta (81% do total), e só 6 afirmaram que não a usariam apenas por ser muito informal para uma reunião de trabalho; resultado que reforça a ideia de rejeição explicita do pronome lembrete entre os falantes com escolarização muito elevada. Além desses, dois participantes não rejeitaram qualquer variante, mas uma participante rejeitou a variante padrão com *o qual* como excessivamente formal e outro rejeitou a cortadora como gramaticalmente incorreta. É interessante observar o perfil desses dois últimos participantes. A que rejeitou a variante padrão como excessivamente formal é mulher, entre 25 e 35 anos de idade, branca, só cursou a graduação, seus pais têm baixa ou média escolaridade e é de Salvador. Já o que rejeitou a cortadora como incorreta

é homem, branco, entre 65 e 75 anos de idade, com pós-graduação, seus pais têm baixa ou média escolaridade, do Rio de Janeiro.

Na segunda frase da situação formal, a OR era extraída da posição de adjunto adverbial de tempo:

( ) Nos anos que a inflação é baixa, o consumo aumenta.
( ) Nos anos em que a inflação é baixa, o consumo aumenta.
( ) Nos anos onde a inflação é baixa, o consumo aumenta.

Nesse caso, um pouco mais de ¾ das escolhas (77%) recaiu na segunda opção, a variante padrão, enquanto menos de ¼ (23%) das escolhas foi para a primeira opção, com a cortadora. A maioria das escolhas pela estratégia padrão veio dos falantes mais velhos (entre 65 e 75 anos), com 93%, enquanto nas faixas etárias mais baixas essa escolha ficou um pouco abaixo de 70%. A diferença entre os sexos não foi significativa, porém os brancos se mostraram mais uma vez mais alinhados ao padrão, com 83% das escolhas pela variante padrão contra 60% entre os pretos e pardos. A escolha pelo padrão também foi maior entre os que cursaram uma pós-graduação, 86% contra 65% entre os que cursaram só a graduação, porém os com pais com baixa ou média escolaridade escolheram mais o padrão do que os com pais com alta escolaridade (87 contra 62,5%). Os cariocas confirmaram sua maior adesão ao padrão: 92% contra 69% dos soteropolitanos.

Essa frase tinha a particularidade de não oferecer a opção da variante resumptiva, colocando em seu lugar a variante introduzida pela palavra interrogativa *onde*. A tradição gramatical circunscreve o emprego do *onde* a antecedentes que se referem a lugar, condenando seu uso com outros sentidos, como a referência temporal, que ocorre nesta frase. O resultado quantitativo revelou que 29 participantes rejeitaram o OR introduzida por *onde*, o que correspondeu a 90% do total de 32 julgamentos (4 participantes não rejeitaram qualquer variante). A grande maioria rejeitou o uso do *onde* como "gramaticalmente incorreto" – foram 25 julgamentos assim (86%do total) contra apenas 4 (14%) que rejeitaram o uso do *onde* por ser "muito informal". Esse resultado indica que as condenações do uso excessivo do *onde* que ocorrem em diversos canais de divulgação de regras do "bom uso" da língua têm repercutido bastante estre os falantes da elite letrada. Em 2 julgamentos a variante cortadora foi rejeitada como "errada". Vale destacar que vieram de falantes brancos e do grupo dos mais velhos. Houve ainda um julgamento da variante padrão como "errada", o que decorreu, ou de lapso interpretativo ao responder o teste, ou de desconhecimento do padrão.

A terceira frase da situação formal continha uma OR extraída da posição de complemento oblíquo:

( ) Essa é a tendência que nos referimos no relatório passado.
( ) Essa é a tendência que nos referimos a ela no relatório passado.
( ) Essa é a tendência à qual nos referimos no relatório passado.
( ) Essa é a tendência a que nos referimos no relatório passado.

Do total de escolhas, 70% se distribuíram entre as duas possiblidades chanceladas pelo padrão normativo, predominando a escolha pela OR *à qual nos referimos no relatório passado*, com 45% das escolhas contra apenas 25% que escolheram a variante *a que nos referimos no relatório passado*, o que revela que o pronome *o qual* impressiona mais como forma de prestígio do que o relativizador *que* preposicionado. Por outro lado, foi nessa frase que a variante cortadora obteve seu maior percentual de escolhas na situação de formalidade: 30%. Quanto ao perfil social, a escolha pela cortadora foi um pouco maior entre os homens (33% contra 28% entre as mulheres), a faixa etária não fugiu as expectativas, com a escolha pela cortadora caindo bruscamente quando se passou da faixa dos falantes mais jovens para os mais velhos (50%, na faixa de 25 a 35 anos; 42% na faixa de 45 a 55 anos; e nenhuma escolha pelas cortadoras na faixa de 65 a 75 anos: 0%). A diferença por etnia também foi notável: 60% de escolhas pela cortadora entre pretos e pardos contra apenas 20% entre os brancos. As cortadoras também predominaram ligeiramente entre os que só cursaram a graduação e entre os que com pais com alta escolaridade, com a mesma diferença de 33 contra 28% dos com pós-graduação e com pais com média e baixa escolaridade.

Nessa frase com a OR de oblíquo, apenas 31 julgamentos rejeitaram alguma variante, a grande maioria (27 de 31, ou 87%) rejeitou a variante com pronome lembrete, dos quais 23 (85%) a consideraram "errada", e apenas 4 (15%) a consideraram apenas muito informal para ser usada em uma reunião de trabalho. Em dois julgamentos, a variante com *que* preposicionado foi rejeitada como errada, o que pode ser visto como desvio interpretativo ou desconhecimento do padrão, e a variante com *o qual* foi rejeitada como formal demais em um julgamento, de uma mulher, preta ou parda, entre 25 e 35 anos, com apenas a graduação, pais com alta escolaridade, de Salvador.

A quarta e última frase da situação formal continha um OR de genitivo:

( ) O princípio é este: o produto que a embalagem dele é colorida vende mais.
( ) O princípio é este: o produto cuja embalagem é colorida vende mais.
( ) O princípio é este: o produto que a embalagem é colorida vende mais.

Exatamente ¾ dos participantes (75%) escolheram a variante com o pronome *cujo*, contra apenas ¼ que escolheram a cortadora (25%), a terceira opção. A opção pelo *cujo* foi um pouco maior entre os homens (78% contra 72% das mulheres) e, mais uma vez, bem maior entre os mais velhos – 92% dos entre 65 e 71 anos, contra 2/3 (66,6%) dos entre 45 e 55 anos e entre 25 e 35 anos – e entre os brancos – 92% contra apenas 30% de opção pelo *cujo* entre pretos e pardos. O *cujo* também foi mais escolhido pelos com pós-graduação e com pais com alta escolaridade, respectivamente, 86 versus 60% e 87 versus 66,6%, e pelos participantes do Rio de Janeiro, com 83% de opção pelo *cujo*, contra 71% entre os soteropolitanos.

Quanto à rejeição, quase todos os 32 julgamentos que rejeitaram alguma variante escolheram a OR com pronome lembrete, e só um julgamento rejeitou a cortadora como muito informal. Esse julgamento veio de um homem, branco, com idade entre 65 e 75 anos, filho de pais com alto grau

de escolarização, porém tinha apenas a graduação e era de Salvador. Dos 31 julgamentos rejeitaram a OR resumptiva, 27 a julgaram como gramaticalmente incorreta, ou seja 87% do total, enquanto 4 (13%) a julgaram apenas informal demais para uma reunião de trabalho.

Reunindo os julgamentos das frases com as ORs extraídas das posições de adjunto adverbial de instrumento, complemento oblíquo e genitivo,[11] os participantes escolheram a estratégia padrão em ¾ dos julgamentos, ou seja, em 74% do total, sendo que 42% das escolhas recaíram na variante como pronome relativo *o qual* e 33% escolheram o relativizador *que* regido por uma preposição. Esse resultado indica que, no sistema de crenças da maioria dos falantes da elite letrada em relação à língua, a estratégia padrão é mais apropriada para ser usada nas situações de comunicação formal, mesmo na oralidade; e o emprego do pronome relativo *o qual* é preferido em detrimento do relativizador *que*. No que concerne à posição e relativização, a estratégia padrão foi mais escolhida com a OR extraída da posição de ADV de instrumento e introduzida pelo pronome relativo *o qual* (*Lavoratex é o produto com o qual vamos iniciar a campanha*.), com 81% do total das escolhas.

A preferência pelo pronome relativo *o qual* em relação ao relativizador *que* ficou clara na OR extraída de complemento oblíquo, em que foram apresentadas as duas opções: *Essa é a tendência* ***à qual*** */* ***a que*** *nos referimos no relatório passado*. Conquanto tenha sido aquela em que a estratégia padrão tenha sido menos escolhida (em 69% dos julgamentos), a opção por *o qual* predominou com 44% das escolhas contra 25% que escolheram a OR introduzida pelo relativizador *que* regido pela preposição *a*. A opção pela OR introduzida pelo pronome relativo *cujo* foi idêntica à média geral de escolhas pela estratégia padrão: 75% dos julgamentos. Ou seja, apesar de ser considerado anacrônico ou extinto pelos linguistas (*cf*. seção anterior), o *cujo* ainda goza de bastante prestígio no sistema de crenças dos falantes da elite letrada em relação aos usos linguísticos.

No perfil social da minoria de 25% dos participantes que assumiram uma atitude não alinhada ao padrão normativo, escolhendo a estratégia cortadora em uma situação de comunicação formal, o sexo não se mostrou uma variável relevante, com praticamente o mesmo percentual de escolha entre homens e mulheres. Já em relação à geração, confirmou-se o comportamento mais conservador dos mais velhos (entre 65 e 75 anos), que escolheram a cortadora em somente 5,6% dos julgamentos, contra 30% na faixa de 45 e 55 anos e 39% na faixa de 25 a 35 anos. A diferença de atitude em função da etnia foi como sempre notável, pretos e pardos optaram pela cortadora em mais da metade dos julgamentos (57%), enquanto os brancos se mostraram bastante conservadores, só fazendo essa opção em 13% dos julgamentos. A opção pela cortadora também foi muito maior entre os que só tinham a graduação, 33% do total contra apenas 19% entre os que cursaram a pós-graduação; entretanto foi apenas um pouco maior entre os que tinham pais com baixa ou média escolaridade, 27% contra 22% dos filhos de pais com escolaridade elevada. Já os participantes de Salvador foram muito mais desviantes do padrão, optando pela cortadora em 33% dos julgamentos contra apenas 8% entre os cariocas.

[11] A terceira frase, com uma OR extraída da posição de adjunto adverbial de tempo, não foi computada nessa síntese, porque, ao contrário das demais, tinha uma variante com *onde* e não tinha a variante resumptiva.

4.2.2. Crenças e atitudes dos falantes em uma situação de comunicação informal

Na segunda parte do teste de adequação, os participantes deveriam se imaginar conversando descontraidamente com os amigos em uma mesa de bar e escolher a variante mais apropriada para uma situação de comunicação informal. Deveriam indicar também uma variante que não usaria, seja por ser *muito formal*, seja por ser *gramaticalmente incorreta*.

Na primeira frase dessa segunda situação, a OR era extraída da posição de ADV de companhia:

> ( ) Então eu encontrei uns caras que eu tinha estudado com eles na Escola Técnica.
> ( ) Então eu encontrei uns caras que eu tinha estudado na Escola Técnica.
> ( ) Então eu encontrei uns caras com quem eu tinha estudado na Escola Técnica.

A expectativa era a de que o participante escolhesse a segunda opção, com uma OR cortadora, por ser uma estratégia informal e não ser avaliada negativamente, como ocorre com a estratégia resumptiva, representada na primeira opção. Já a terceira frase, com a estratégia padrão, deveria ser considerada muito formal para ser usada em uma conversa de mesa de bar. Porém, as opções dos participantes foram bem divididas. De um total de 39 escolhas,[12] 15 foram pela cortadora (38,5%), 13 pela resumptiva (33,3%) e 11 pela estratégia padrão (28,2%). Não obstante ser a mais escolhida, a opção pela estratégia cortadora não foi tão majoritária quanto se esperava. O fato de o número de escolhas pela resumptiva ser bem próximo ao da cortadora revela que essa estratégia está com um nível de aceitação, para a comunicação informal, maior do que o que se tem dito normalmente na literatura. Alguns fatores estruturais favoreceram a opção pelo pronome lembrete nessa frase: uma posição de relativização pouco acessível, situada no interior do sintagma verbal (SV), e o antecedente ter o traço semântico [+humano]. Por fim, o elevado número de opções pela estratégia padrão revela que, embora não corresponda ao uso linguístico concreto dos participantes, essa variante é bem valorizada no seu sistema de crenças em relação ao uso da língua.

As características sociais dos participantes que escolheram a estratégia padrão correspondem, no geral, ao perfil do falante conservador: homem (40% versus 16% entre as mulheres), mais velho (46%, contra 31% na faixa intermediária, de 45 a 55 anos, e 8% apenas na faixa de 25 a 35 anos), branco (31% contra 20% entre pretos e pardos), com pós-graduação (32% contra 23,5% dos só com graduação) e carioca (36% contra 24% entre os baianos); destoando apenas uma maior preferência entre os de pais com baixa e média escolaridade (33,3% contra 20% entre os filhos de pais com alta escolaridade). Já as características dos que optaram por usar o pronome lembrete não se enquadraram no que tem se configurado nesta análise como o perfil do falante progressista. Correspondeu a esse perfil apenas uma maior escolha entre as mulheres (42% contra 25% dos homens). Optaram mais pela resumptiva: os brancos (com 38% contra 20% dos pretos e pardos), os da geração intermediária e mais

[12] Como dito anteriormente, o participante poderia escolher mais de uma forma que usaria em cada situação, daí haver um número de escolhas (39) maior do que o de participantes (36).

velhos (com 46% e 38,5%, respectivamente, contra 15,5% dos mais jovens), os com pós-graduação (com 36% contra 29% dos só com graduação) e com pais com escolaridade elevada (40% contra 29% dos com pais com baixa e média escolaridade). A diferença entre as cidades não foi significativa: 36% no Rio de Janeiro contra 32% em Salvador. A falta de um perfil social nítido na escolha do pronome lembrete reforça a visão de que essa estratégia de relativização é mais condicionada por fatores estruturais do que por fatores sociais (Lucchesi; Santos, 2025).

Contradizendo um pouco o resultado da forma escolhida, dos 30 julgamentos que rejeitaram alguma frase,[13] mais da metade (16, ou 53,3%) rejeitaram a OR com pronome lembrete, considerando-a imprópria para ser usada mesmo em uma situação de grande informalidade, por conter um erro gramatical grave. Esse resultado vai ao encontro da ideia corrente de que essa variante recebe uma avaliação explicitamente negativa entre os falantes da elite letrada. Ainda mais conservadores foram 3 julgamentos (10% do total) que rejeitaram a relativa cortadora por também a considerarem *gramaticalmente incorreta*. Por outro lado, 5 rejeitaram a variante padrão como formal demais para ser usada numa conversa de mesa de bar. Mas, surpreendeu haver 6 julgamentos que rejeitaram a OR padrão por a considerarem gramaticalmente incorreta, o que tem duas explicações possíveis: desvio de atenção na hora de responder o teste ou desconhecimento do padrão.

Entre os 3 julgamentos que rejeitaram a cortadora, predominaram as características sociais do perfil do falante conservador: todos brancos, da geração dos falantes mais velhos, majoritariamente homens e com pós-graduação (2 contra 1), porém a maioria tinha pais com baixa e média escolaridade e era de Salvador. Mas, surpreendentemente, os que rejeitaram a OR padrão como muito formal eram todos brancos, em sua grande maioria com pós-graduação e do Rio de Janeiro (4 contra 1); ou seja, no geral seriam as características do falante com perfil conservador, mais alinhado ao padrão normativo.

Na segunda frase do teste, a OR foi extraída da posição de adjunto adverbial de situação:

( ) Você viu aquele filme que o menino fala com os mortos?
( ) Você viu aquele filme que o menino fala com os mortos nele?
( ) Você viu aquele filme onde o menino fala com os mortos?
( ) Você viu aquele filme em que o menino fala com os mortos?

Os participantes escolheram 45 possibilidades de uso, das quais praticamente a metade foi pela relativa cortadora (23 escolhas, 51% do total), a primeira opção, em seguida escolheram a padrão, a última opção (19 escolhas, 42% do total), e somente um participante escolheu a resumptiva, a segunda opção, e dois a relativa introduzida por *onde*, a terceira opção. Mais uma vez, a estratégia padrão foi bastante escolhida, o que só reforça a ideia de que essa variante é muito valorizada no imaginário dos falantes, mesmo não correspondendo em absoluto ao seu uso concreto da língua. Por outro lado, os membros da elite letrada, mais uma vez, demonstraram uma atitude refratária ao uso do *onde* como relativizador.

[13] Como já foi dito, o participante poderia não rejeitar nenhuma alternativa, daí o número menor de rejeição (30) em relação ao número total de participantes.

Quanto às características sociais dos que optaram pela cortadora, predominaram as que têm definido aqui o perfil do falante mais progressista, ou menos alinhado ao padrão normativo: foram majoritariamente mulheres (com 58% contra 43% dos homens), mais jovens (com 67%, entre os com 25 a 35 anos; 53%, entre 45 e 55 anos, e 33,3%, entre 65 e 75 anos), pretos e pardos (com 64% contra 47% dos brancos) e baianos (53% contra 47% dos cariocas); já os com pais com alta escolaridade também optaram mais pela cortadora do que os cujos pais tinham escolaridade baixa ou intermediária (58% contra 46%); e o nível de escolaridade do participante não se mostrou relevante nesse caso. Já os que escolheram a variante padrão corroboraram o que se tem definido aqui como perfil do falante conservador: homem (57% contra 29% entre as mulheres), mais velho (60%, entre 65 e 75 anos; 47%, entre 45 e 55 anos; 20%, entre 25 e 35 anos), branco (50% contra 18% dos pretos e pardos), com pós-graduação (50% contra 32% dos só com graduação) e do Rio de Janeiro (47% contra 40% dos soteropolitanos), conquanto essa diferença entre as cidades não tenha sido muito expressiva. E a escolaridade dos pais não teve qualquer influência nesse caso.

O resultado da forma rejeitada nesta frase ratificou a avaliação explicitamente negativa do uso do pronome lembrete. Essa variante foi rejeitada em mais da metade dos julgamentos (23 em 40, ou 57,5% do total). Alguns fatores estruturais contribuíram para a rejeição do lembrete nessa frase, especialmente o antecedente ser [-humano]. Merece destaque também o fato de a OR introduzida por *onde* ser a outra variante bem rejeitada, com 15 recusas em um total de 40 (37,5%), o que confirma a resistência ao emprego dessa palavra interrogativa como relativizador. E mais uma vez observa-se uma contradição entre crença e uso concreto da língua, pois tem sido crescente o emprego do *onde* na fala corrente, o que tem despertado a ira dos puristas. Por fim, só um participante rejeitou a OR padrão como muito formal e outro, a cortadora como gramaticalmente incorreta; o primeiro foi um homem, jovem, preto ou pardo, com pós-graduação, com pais com escolaridade baixa ou média e de Salvador; a outra é uma mulher, da geração mais velha, branca, com pós-graduação, com pais com escolaridade elevada e de Salvador; o que no geral corrobora, respectivamente, a definição dos perfis progressista e conservador entre os falantes da elite letrada em relação ao padrão normativo.

A terceira frase do teste continha uma OR de sujeito, mas a predicação da oração matriz pode ser definida aqui como *apresentativa*, ou seja, é aquela que introduz um referente novo como antecedente da OR, o qual é, portanto, discursivamente destacado, um contexto que favoreceria o emprego do pronome lembrete.

( ) Eu conheço um cara que é igualzinho a você.
( ) Eu conheço um cara que ele é igualzinho a você.

Apesar disso e de se tratar de uma situação bem informal de comunicação, praticamente 90% das escolhas foram pela primeira opção (33 de 37 escolhas), e só quatro participantes escolheram a estratégia resumptiva, ou seja, pouco mais de 10% do total (10,8%). Isso reforça o caráter minoritário do emprego do pronome lembrete, mesmo nos contextos estruturais que lhe são

favoráveis – nesse caso, o antecedente [+humano], além da predicação da oração matriz. O perfil social dos que optaram pela resumptiva não foi muito nítido: quatro mulheres cujos pais tinham escolaridade baixa ou média, mais concentradas na geração intermediária (duas contra uma na geração mais velha e na mais nova), majoritariamente brancas e só com a graduação (três contra uma, nos dois casos), mas distribuídas igualmente entre as duas cidades (duas em cada uma). Esse resultado reforça mais uma vez a ideia de que o uso do pronome lembrete é menos condicionado por fatores sociais, do que por fatores estruturais.

Quanto à forma que os participantes rejeitaram, quase a metade do total de 36 participantes (ou seja, 16, ou 44%) não rejeitou nenhuma das duas opções, o que não corrobora a ideia de que o uso do lembrete é muito rejeitado pelos falantes muito escolarizados. Dos 20 que rejeitaram uma forma, só um participante rejeitou a primeira opção (*eu conheço um cara que é igualzinho a você*) como formal demais, o que deve ser atribuído a um problema de interpretação. Entre os 19 que rejeitaram a OR com pronome lembrete, predominaram os homens (11 contra 9 mulheres), os mais velhos (11 com 65 a 75 anos, 6 com 45 a 55 anos, e só 2 com 25 a 35 anos), os brancos (15 contra apenas 4 pretos e pardos), com pós-graduação (14 contra 5 com graduação apenas); com pais tinham que escolaridade baixa ou média (11 contra 8 com pais de escolaridade elevada); e de Salvador (14 contra apenas 5 do Rio de Janeiro).

Na última frase que continha um OR no teste, essa OR era extraída da posição de adjunto adnominal (genitivo)

( ) Eu tinha uma namorada que o irmão dela era autista.
( ) Eu tinha uma namorada que o irmão era autista.
( ) Eu tinha uma namorada cujo irmão era autista.

Em se tratando de uma situação informal, a expectativa era que o participante escolhesse a segunda opção, a cortadora, ou mesmo a primeira opção, resumptiva, porque a posição de genitivo favorece essa estratégia de relativização por ser a posição menos acessível, considerando a clássica Hierarquia de Acessibilidade (HA), proposta por Keenan e Comrie (1977). Com efeito a cortadora foi a opção em 70% das escolhas (28 em um total de 40), e só 20% das escolhas recaiu na estratégia resumptiva, confirmando seu caráter minoritário, mesmo nos contextos estruturais que lhe são favoráveis. Por outro lado, deve-se registrar que em 10% dos julgamentos (4 em 40) a opção foi pela OR padrão introduzida pelo pronome *cujo*, o que confirma duas coisas: a valorização da variante padrão, mesmo com esse pronome considerado extinto pelos linguistas, e a contradição entre crença e desempenho dos falantes, que nunca usariam essa frase com *cujo* numa conversa descontraída de mesa de bar.

Todos os quatro participantes que escolheram a OR padrão com *cujo* eram homens brancos, três tinham cursado a pós-graduação e eram cariocas e filhos de pais com alta escolaridade, porém dois eram da geração dos mais novos, contra um da geração intermediária e um dos mais velhos. Assim, no

geral, confirmaram-se as características sociais do perfil conservador. Do lado oposto, a escolha pela OR resumptiva foi mais frequente com as mulheres (28,6% contra 10,5% entre os homens), entre os mais jovens e de meia idade (21,4% e 28,6%, respectivamente, contra apenas 8,3% dos mais velhos), entre os que cursaram apenas a graduação (23,5% contra 17,4% dos com pós-graduação), cujos pais tinham escolaridade baixa ou intermediária (29,2% contra 6,2% dos com pais com alta escolaridade) e em Salvador (22,2% contra 15,4% no Rio de Janeiro); porém, foi mais frequente entre os brancos, com 21,4% das escolhas contra 16,7% entre os pretos e pardos.

4.3. Crenças e atitudes *versus* uso da língua

Os resultados dos testes de *avaliação* e *adequação* revelaram que as formas prescritas pela tradição normativa para a construção das ORs gozam de bastante prestígio no sistema de crenças dos membros da elite letrada, que reputam ser a estratégia padrão a mais adequada para ser usada, até mesmo em situações mais informais de comunicação, embora nesse caso essa escolha tenha sido minoritária. Porém, uma análise variacionista feita com base em entrevistas do tipo sociolinguístico com os doze participantes do Rio de Janeiro que responderam os testes aqui apresentados revelou uma grande contradição entre essas crenças e o seu desempenho linguístico real (Santos, 2025).

Nas doze entrevistas, foram encontradas 962 ORs, sendo que praticamente ¾ delas (74,7% do total) eram ORs de sujeito e objeto direto, o que vai ao encontro da previsão da HA de Keenan e Comrie (1977) de que essas posições são as mais acessíveis à relativização. Nas ORs extraídas das posições sintáticas preposicionadas e menos acessíveis, quase 20% era de ORs cortadoras (19,6%) e só 3,2% era de ORs padrão e 2,4% de resumptivas. Deve-se destacar que, em termos percentuais, o uso da estratégia padrão praticamente se equivale ao uso do pronome lembrete, uma estratégia minoritária supostamente rejeitada pelos falantes altamente escolarizados.

Isolando as ORs extraídas de posições de relativização adverbiais e preposicionadas, nas quais ocorre a oposição entre as estratégias padrão e cortadora, esta última é amplamente majoritária, com 85,9% do total de ocorrências (189 em 220), contra uma frequência de 14,1% da estratégia padrão (31 em 220). Das 31 ORs padrão, quase 1/3 foi introduzida pelo relativizador *onde/aonde* (8 ocorrências de *onde* e 2 de *aonde*), que só é superado pelo relativizador universal *que*, empregado em 13 ocorrências de ORs padrão da amostra. Foram encontrados ainda os seguintes pronomes relativos: *quando*, em três ocorrências, *quem* e *como*,[14] em duas ocorrências cada, e apenas uma ocorrência de *o qual*, pronome tão valorizado no sistema de crenças da elite letrada, enquanto o *onde*, que é tão rejeitado nos testes é o pronome relativo mais usado no desempenho real desses falantes. Reunidos todos os pronomes relativos propriamente ditos (excluindo o relativizador universal *que*), eles só são empregados em 1,9% do total de ORs encontradas na amostra (18 em um total de 960 ORs). Deve-se destacar ainda o caso do *cujo* – tão valorizado nos testes de avaliação e adequação –, que não é usado sequer uma vez em um tempo estimado de oito horas de fala coloquial.

[14] Em construções do tipo: *a forma / a maneira como...*

Com esses dados, resta evidente a grande contradição entre o sistema de crenças dos falantes altamente escolarizados em relação às estratégias de relativização disponíveis em língua portuguesa e seu desempenho linguístico. A valorização da estratégia padrão de relativização e dos pronomes relativos, especialmente o *cujo* e *o qual*, nos testes contrasta com o reduzidíssimo emprego dessa estratégia de relativização na fala corrente e com a ausência de emprego *cujo*. Por fim, ironicamente, o pronome relativo *onde*, mal avaliado nos testes, é o mais empregado na fala.

5. Considerações finais

Antes de apresentar uma síntese dos resultados alcançados neste estudo, cabe uma avaliação da metodologia empregada. O teste de adequação a uma situação de fala se mostrou mais eficaz para elicitar as crenças e atitudes dos falantes do que o teste de avaliação direta das variantes linguísticas; os resultados daquele foram mais congruentes do que os deste, produzindo um quadro mais nítido, não apenas dos juízos e atitudes dos falantes frente ao fenômeno linguístico em foco, como também do perfil social desses julgamentos e atitudes. Portanto, não obstante as limitações e restrições da metodologia empregada, o estudo aqui apresentado alcançou resultados significativos.

Considerando incialmente as crenças e atitudes dos falantes da elite letrada em relação à variação na forma das orações relativas, os resultados revelaram que as prescrições da tradição gramatical têm muita força no sistema de crenças dos falantes altamente escolarizados, que valorizam muito a estratégia padrão de relativização, embora praticamente não a usem em sua fala corrente. Em pelo menos 70% dos julgamentos, os participantes desta pesquisa consideraram a relativa padrão como a mais apropriada para ser empregada nas situações mais formais de comunicação, tanto na escrita, quanto na fala. O prestígio da variante padrão pode ser medido pelo fato de essa forma ser escolhida em até 40% das opções numa hipotética situação bem informal de comunicação. Conquanto essa opção esteja em flagrante contradição com o comportamento linguístico real do participante, ela expressa o grande valor simbólico da OR padrão em seu sistema de crenças sobre o uso da língua. Chamou atenção o valor que os participantes emprestaram a formas, como os pronomes relativos *cujo* e *o qual*, que estão praticamente ausentes de sua fala corrente. Por outro lado, identificou-se uma resistência ao emprego do *onde* como partícula relativizadora. Isso pode ser atribuído às condenações dos puristas ao crescente uso dessa palavra gramatical. E, ironicamente, os falantes altamente escolarizados refutam em seu imaginário o único relativizador morfologicamente marcado que ainda usam em sua comunicação cotidiana. Assim, um dos resultados mais importantes deste estudo foi a confirmação da afirmação de Labov (2008 [1972]) de que a forma como as pessoas pensam que falam é muito diferente da forma como elas efetivamente falam.

No que concerne às ORs cortadoras, no geral se confirmou a hipótese reiterada desde Tarallo (1983) de que essa estratégia de relativização não recebe uma avaliação negativa. No teste de avaliação direta, a estratégia cortadora foi avaliada como típica da linguagem informal por quase 80% dos participantes (77,8% do total), e só foi avaliada como incorreta em pouco mais de 20% das

respostas (22,2% do total). Na situação formal do teste de adequação, a opção pela cortadora ficou entre 20 e 30% (*circa*). Já na situação informal, a opção pela cortadora atingiu a marca de 70% das escolhas na posição de genitivo, sendo a opção em um pouco mais da metade das escolhas (51,1%) na posição de ADV de situação, e só caindo para menos da metade na OR extraída da posição de ADV de companhia, mas mesmo aí foi a mais escolhida (38,5%), porque as demais escolhas se dividiram entre a estratégia resumptiva (33,3%) e padrão (28,25). Portanto, já em relação à estratégia cortadora, a avaliação dos falantes se aproxima mais de seu desempenho linguístico real.

Por fim, esta análise buscou testar a hipótese de que a estratégia resumptiva era explicitamente rejeitada pelos falantes da elite letrada. O resultado do teste de avaliação direta não confirmou plenamente essa hipótese, pois o uso do pronome lembrete só foi rejeitado como "um erro gramatical grave", por pouco mais da metade dos participantes (55,6%), ou seja, quase a metade (44,4%) avaliou que as frases com pronome lembrete são usadas "em situações de descontração, que não seguem necessariamente as regras da gramática, mas não contêm nenhum erro grave". Já na situação formal do teste de adequação, a resumptiva naturalmente não foi escolhida por nenhum participante e foi rejeitada em um nível de 67 a 97% dos julgamentos, sendo essa rejeição justificada por ser uma forma gramaticalmente incorreta, em um nível de 81 a 87% dos julgamentos. Esses resultados já favorecem a hipótese de que o uso do pronome lembrete é avaliado negativamente pelos falantes altamente escolarizados de forma explícita, porém essa afirmação deve ser modalizada para as situações de uso formal da língua, pois parte dos participantes do teste escolheu usar a resumptiva na situação de comunicação informal – o uso do pronome lembrete chegou a ser a forma escolhida numa conversa descontraída em 1/3 das respostas (33,3% do total). Por outro lado, o pronome lembrete foi rejeitado, nessa situação informal, em mais da metade dos julgamentos, nas ORs extraídas das posições de ADV de companhia e de situação (em 53,3% e 57,5% do total, respectivamente); sendo que sua rejeição caiu bruscamente na OR de genitivo (24% do total), por se tratar da posição sintática menos acessível à relativização, o que confirma a visão de essa estratégia de relativização é fortemente condicionada por fatores estruturais, mais do que por fatores sociais. De qualquer forma, as diferenças na forma como os participantes avaliaram as estratégias de relativização revelou que os falantes da elite letrada se dividem em dois perfis distintos, no que se refere ao seu alinhamento às prescrições gramaticais.

Assim, um dos mais importantes achados desta pesquisa foi a identificação de duas atitudes distintas em relação ao padrão normativo tradicional entre os falantes da elite letrada, distinguindo aqueles indivíduos com um perfil mais conservador, que reproduzem mais o padrão normativo em suas crenças e atitudes, e aqueloutros, com um perfil mais progressista, que divergem mais do padrão normativo e têm uma atitude e um comportamento mais flexível, aceitando e empregando formas não aceitas pela tradição gramatical, apesar de não ser possível identificar sempre se as respostas divergentes resultavam de uma atitude crítica ou simplesmente do desconhecimento do padrão normativo.

É razoável pensar que essa divisão nas atitudes e crenças dos falantes letrados está relacionada ao contexto mais amplo da disputa política e ideológica que se trava na sociedade. Nas últimas décadas, movimentos que lutam contra o preconceito e a discriminação têm ocupado cada vez mais espaço na cena política, buscando construir uma sociedade mais tolerante e inclusiva, na qual haja aceitação e respeito à diversidade comportamental, de gênero, de opção religiosa etc. Esses movimentos, denominados *movimentos identitários*, convergem para uma ideologia progressista e de esquerda. Em contrapartida, a direita e a extrema direita têm investido numa pauta de costume conservadora como uma forma de galvanizar uma base popular de apoio, atacando, inclusive, o que denominam "ideologia de gênero" (Boulos, 2025, p. 26-27). Essa disputa já se refletiu diretamente na língua, com o que se denominou *linguagem neutra* ou *linguagem inclusiva*, a proposição por parte de alguns grupos minoritários de formas de um gênero neutro, como *amigue* e *elo*, para se referir a pessoas não binárias, ou *todes*, para evitar a supremacia do gênero masculino como "gênero universal" (Barbosa Filho; Othero, 2022; Freitag, 2024). Tal inciativa tem sido fortemente combatida pelos grupos conservadores de direita e extrema direita, inclusive com parlamentares que propõem leis, proibindo o uso dessa "linguagem neutra".

Mas o desenvolvimento de uma consciência social mais aberta à diversidade cultural e comportamental pode se refletir no plano da ideologia linguística em outros níveis; por exemplo, na contraposição à ideologia discricionária do normativismo tradicional, que só reconhece uma variedade da língua como correta e condena as demais variedades, sobretudo a variedade popular, como degradadas ou inferiores. Estudos desenvolvidos na Europa têm identificado esse movimento que leva ao enfraquecimento da hegemonia da ideologia normativista (Kristiansen; Grondelaers, 2013; Monteagudo, 2024).

Não obstante a visão hegemônica de língua na sociedade ainda ser plasmada pela secular tradição gramatical – o que se reflete, por exemplo, no grande valor que as formas das ORs padrão assumem no imaginário da elite letrada –, pode-se vislumbrar uma mudança de atitude dos falantes em relação à língua, pela influência desses movimentos de defesa da diversidade e da inclusão social, bem como em função de uma mudança na orientação do ensino de língua portuguesa que vem acontecendo em muitas escolas desde o final do século passado (*cf.* o final da seção 3). Assim, a aceitação de relativas cortadoras em situações formais e do pronome lembrete em situações de comunicação informal, nas respostas aos testes desta pesquisa podem ser atribuídas ao desenvolvimento dessa atitude mais progressista de questionamento a uma tradição gramatical discricionária. E o perfil social das atitudes dos participantes desta análise é muito significativo e coerente com o que está acontecendo no contexto social mais amplo das disputas simbólicas e ideológicas que se travam na sociedade.

Dentre as características sociais dos participantes que assumiram um perfil mais progressistas, menos alinhado às prescrições normativistas nesta pesquisa, se destacaram o sexo, a idade, a etnia e a cidade do participante. No contexto mais geral, as mulheres têm assumido uma atitude política mais progressista e mais sensível à diversidade cultural e comportamental do que os homens. Nas

pesquisas eleitorais das últimas eleições, bem polarizadas, entre um candidato da extrema direita, com características fascistas, e um candidato progressista de centro-esquerda, o candidato progressista sempre foi o mais escolhido pelas mulheres, enquanto o candidato da extrema direita predominou entre os homens.[15] A extrema direita tem se associado às seitas neopentecostais, que pregam a submissão da mulher ao homem no arquétipo da família tradicional. Em contrapartida, o feminismo e demais movimentos progressistas tem combatido a ideologia machista, que fomenta a misoginia e o feminicídio. Todo esse contexto mais amplo leva às mulheres, em média, a ter um comportamento mais inclusivo, sensível e aberto à diversidade comportamental, o que naturalmente se reflete, não apenas na sua forma de avaliar as formas da língua, mas também no seu próprio comportamento linguístico, pois análises realizadas sobre a amostra desses falantes do Rio de Janeiro têm observado que as mulheres estão liderando as mudanças para fora do padrão normativo (Santos, 2025; Castelló, 2025). Em contrapartida, os homens que buscam manter seu poder simbólico na sociedade tendem a reproduzir as representações hegemônicas do status quo, inclusive o padrão da tradição normativa, como forma de fortalecer sua posição de poder.

O racismo estrutural da sociedade brasileira afeta a vida de todos os pretos e pardos cotidianamente, o que tende a despertar a sensibilidade e a consciência para a questão do preconceito e da discriminação. Essa consciência se aprofunda com o engajamento em movimentos de luta contra a discriminação racial e a valorização da cultura e ancestralidade negras, o que acaba por se refletir em sua opção eleitoral [16,] bem como em suas crenças e atitudes em relação à língua. Nesse sentido, foi muito consistente a diferença observada nesta pesquisa. Enquanto os brancos assumiram quase sempre uma postura mais conservadora de reprodução e valorização do padrão normativo, pretos e pardos tiveram uma postura mais progressista, aceitando e optando mais pelas formas não padrão. Assim, seja por desconhecimento das prescrições gramaticais, seja por assumirem uma postura mais crítica e sensível a diversidade linguística, os pretos e pardos se mostraram bem menos influenciados pela ideologia normativa do que os brancos.

Como esperado, os jovens assumiram uma postura muito mais inovadora do que os mais velhos. É uma característica universal dos jovens serem mais abertos à mudança e às novidades. É natural que isso se reflita no plano das atitudes e crenças em relação à língua. Contudo, deve-se ter em mente também a possibilidade de um menor domínio das prescrições gramaticais, seja por problemas no sistema de ensino, seja em decorrência da mudança nas orientações pedagógicas, que têm deixado de privilegiar o ensino da norma padrão como principal objetivo do ensino de língua portuguesa, nas últimas décadas.

---

[15] *Cf*. Boulos (2025, p. 27): "No Brasil, em 2022, segundo pesquisa do Datafolha na véspera do segundo turno, 52% das mulheres declararam voto em Lula, contra 41% em Bolsonaro. Entre os homens, Bolsonaro levou vantagem de 48% a 46%. O corte racial é ainda mais nítido: Lula venceu Bolsonaro por 60% a 34% entre os pretos, e Bolsonaro venceu por 54% a 40% entre os brancos."

[16] Ver nota 12.

Também foi significativa a diferença de atitude e avaliação entre os participantes do Rio de Janeiro e de Salvador. Aqueles reproduziram muito mais o padrão normativo em seus julgamentos do que estes. Nas últimas eleições, o Nordeste é a região que tem se destacado por uma votação mais progressista, inclusive foi decisiva para a eleição do presidente Lula da Silva, em 2022. Pode-se pensar que essa postura mais progressista se tenha refletido na atitude dos falantes de Salvador perante a língua. Por outro lado, pode-se pensar que os moradores do Rio de Janeiro teriam, em média, um maior domínio da norma de referência linguística, por viverem em uma cidade mais desenvolvida socioeconomicamente. Outros fatores associados a história social dessas duas cidades serão considerados nos desdobramentos futuros desta pesquisa.

Em princípio, a questão do domínio do padrão estaria diretamente ligada às variáveis nível de escolaridade do participante, bem como dos seus pais. Porém, essas variáveis não apresentaram uma distribuição muito nítida. Os participantes com pós-graduação stricto sensu se mostraram um pouco mais conservadores do que os que só cursaram a graduação. O trabalho de redação de dissertação e/ou tese, bem como de artigos acadêmicos, pode desenvolver uma atitude mais ciosa em relação ao padrão normativo. Por outro lado, os falantes com pais com escolaridade baixa ou intermediária se mostraram um pouco mais apegados ao padrão normativo do que os que os com pais com elevada escolaridade, o que parece contraditório com o resultado anterior. Entretanto, pode-se pensar que muitos dos que têm pais com baixa escolaridade são os primeiros de sua família a entrar na universidade, o que foi possibilitado pelas políticas de ação afirmativa implementadas nos governos do Partido dos Trabalhadores, neste século. É frequente que os indivíduos que ascendem socialmente busquem assimilar os símbolos das classes sociais mais elevadas e de prestígio social, como forma de afirmar sua nova condição de classe. Isso pode explicar a maior preocupação em escolher as formas do padrão normativo dos filhos de pais com baixa escolaridade.

De qualquer forma, seja por um maior ou menor conhecimento, seja por uma atitude mais ou menos crítica em relação ao padrão normativo, ou ainda por outra motivação, um dos mais importantes achados desta pesquisa foi identificar dois perfis dos falantes na elite letrada: um que reproduz mais as prescrições gramaticais, e outro que não reproduz o padrão. Há séculos a visão hegemônica de língua na sociedade é plasmada pela tradição gramatical e difundida pela escola e pelos meios de comunicação de massa. Essa visão que preconiza o uso de uma única variedade da língua, julgando as demais como impróprias ou corrompidas, naturalmente teria mais força entre os falantes com um nível de escolaridade muito elevado. Contudo, os avanços no plano das mentalidades observados nas últimas décadas, pela atuação de movimentos que lutam pela aceitação e em defesa da diversidade cultural e comportamental estaria se refletindo no plano das atitudes e crenças dos falantes em relação à língua, promovendo uma aceitação mais ou menos consciente de formas que, apesar de não serem acolhidas pela tradição gramatical, são legitimadas por seu uso corrente.

Nesse sentido, coloca-se em questão a concepção de Labov (2008 [1972]) de *comunidade de fala*, baseada na homogeneidade da avaliação subjetiva dos falantes das formas variantes em uso na

língua. Este estudo revelou que, mesmo no seio da elite letrada, onde se esperaria uma convergência em torno do padrão normativo, não há essa homogeneidade, encontrando-se dois perfis diferenciados. Além disso, nas classes socialmente marginalizadas, que não têm acesso aos direitos da cidadania, como a escolarização, a assimilação do padrão normativo é muito menor, ou mesmo mínima, o que levou Lucchesi (2015) a propor, como um dos parâmetros que configuram polarização sociolinguística do Brasil, exatamente a diferença na forma diferenciada como os falantes da elite social e os das classes mais marginalizadas avaliam as variantes linguísticas. Assim, o modelo de Lucchesi e o de Milroy e Milroy (1992), que se baseiam nas contradições entre as classes sociais e na disputa política e ideológica que se trava na sociedade, se mostram mais adequados para produzir uma representação social da língua do que o modelo de Labov (2008 [1972]), que se baseia numa visão hierarquizada das classes sociais e no consenso em torno da ideologia dominante, especialmente no plano subjetivo das crenças e atitudes dos falantes em relação aos usos da língua.

Referências

AZEREDO, José Carlos de. *Gramática Houaiss da língua portuguesa*. São Paulo: Publifolha, 2011.

BAGNO, Marcos. *Gramática pedagógica do português brasileiro*. São Paulo: Parábola Editorial, 2012.

BARBOSA Filho, Fábio; Othero, Gabriel. *Linguagem "neutra":* língua e gênero em debate. São Paulo: Parábola, 2022.

BECHARA, Evanildo. *Moderna Gramática Portuguesa*. 73 ed. revista e ampliada. Rio de Janeiro: Lucerna, 2001.

BOULOS, Guilherme. *Pra onde vai a esquerda?* Avaré, SP: Contracorrente, 2025.

CASTELLÓ, Vera Lourenço. *A variação nas construções existências na norma culta do Rio de Janeiro*. 2025. 126 f. Dissertação (Mestrado em Estudos de Linguagem) – Instituto de Letras, Universidade Federal Fluminense, Niterói, 2025.

CEGALLA, Domingos P. *Novíssima Gramática da Língua Portuguesa*. 17. ed. São Paulo: Companhia Editora Nacional, 1977.

CORRÊA, Vilma. *Oração relativa: o que se fala e o que se aprende no português do Brasil*. 1998. 174 f. Tese de Doutorado, Campinas: Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, 1998.

FREITAG, Raquel. *Não existe linguagem neutra!* gênero na sociedade e na gramática do português brasileiro. São Paulo: Contexto, 2024.

KEENAN, Edward L.; COMRIE, Bernard. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, n.8, pp. 63-99, 1977.

KRISTIANSEN, Tore; GRONDELAERS, Stefan (eds.). *Language (De)standardisation in Late Modern Europe*: Experimental Studies. Oslo: Novus Press, 2013.

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

LUCCHESI, Dante. *Língua e Sociedade Partidas:* a polarização sociolinguística do Brasil. São Paulo: Contexto, 2015.

LUCCHESI, Dante; Santos, Michelle. As orações relativas resumptivas na norma culta da cidade do Rio de Janeiro em cinquenta anos (1973-2023). *D.E.L.T.A.*, São Paulo, v. 41, n. 4, pp. 1-26, jul./dez. 2025. DOI: https://doi.org/10.1590/1678-460x202541474060. Acesso em: 02 jan.2026.

MILROY, Lesley; MILROY, James. Social network and social class: toward an integrated sociolinguistic model. *Language in Society*, n. 21, pp. 1-26, 1992.

MONTEAGUDO, Henrique (ed.). *A estandarización das linguas da Península Ibérica:* procesos, problemas e novos horizontes. Santiago de Compostela: Consello da Cultura Galega, 2024.

PAIVA; Maria da Conceição de; GOMES, Christina Abreu. Grupo PEUL: passado, presente e futuro de uma agenda de pesquisa. *Cadernos de Estudos Linguísticos*, Campinas, pp. 503-519 - set./dez. 2016. Disponível em: https://periodicos.sbu.unicamp.br/ojs/index.php/cel/article/view/8647614. Acesso em: 12 jan.2026.

ROCHA LIMA, Carlos Henrique da. *Gramática Normativa da Língua Portuguesa*. 5. ed. Rio de Janeiro: Briguiet, 1960.

SANTOS, Michelle Silva dos. *As estratégias de relativização na norma urbana culta do Rio de Janeiro*: um estudo em tempo real em três sincronias (1970, 1990, 2020). 2025. 196 f. Tese (Doutorado em Estudos de Linguagem) – Instituto de Letras, Universidade Federal Fluminense, Niterói, 2025.

TARALLO, Fernando. *Relativization strategies in Brazilian Portuguese*. 1983. 295 f. Tese (Doutorado em Linguística) – Universidade da Pensilvânia, Pensilvânia, 1983.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança linguí*stica. São Paulo: Parábola, 2006 [1968].

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72247

O sotaque das capitais da região sudeste do Brasil: um estudo de avaliação

The accents of the capital cities of Brazil's southeastern region: an evaluation study

Leila Maria Tesch[1]

RESUMO

Este estudo investiga a avaliação dos sotaques das capitais da Região Sudeste do Brasil — Belo Horizonte, Rio de Janeiro, São Paulo e Vitória — à luz dos pressupostos teórico-metodológicos da Sociolinguística, com foco nos estudos de percepção e avaliação linguística. Parte-se da concepção de que o sotaque é um fenômeno socialmente construído, associado a identidades, representações sociais e valores simbólicos, e de que as percepções dos ouvintes são sistemáticas, variáveis e ordenadas. Para tanto, foi elaborado um formulário *on-line*, aplicado via *Google Forms*, no qual participantes brasileiros, maiores de 18 anos, de diferentes níveis de escolaridade e de ambos os sexos, avaliaram o modo de falar de pessoas das capitais do Sudeste em quatro dimensões: prestígio, beleza, correção e agradabilidade, totalizando 1.401 respostas válidas. Os resultados indicam que os sotaques do Rio de Janeiro e de São Paulo foram mais frequentemente associados ao prestígio; o de Belo Horizonte destacou-se nos critérios beleza e agradabilidade; e o de Vitória foi percebido como a variedade mais "correta". Tais resultados evidenciam que as avaliações não são homogêneas e refletem estereótipos linguísticos e sociais historicamente construídos, influenciados tanto por fatores linguísticos quanto pelo perfil social dos ouvintes. Ao discutir esses padrões, o estudo reforça a importância do respeito linguístico, especialmente no que se refere às variedades linguísticas típicas de cada região, reconhecendo-as como formas legítimas de expressão do português brasileiro. Dessa forma, a pesquisa contribui para o debate sobre atitudes linguísticas, variação e identidade no contexto urbano contemporâneo.

PALAVRAS-CHAVE: Sociolinguística. Estudo de percepção. Sotaque. Identidade. Região Sudeste.

ABSTRACT

This study investigates the evaluation of accents from the capital cities of Brazil's Southeast Region—Belo Horizonte, Rio de Janeiro, São Paulo, and Vitória—within the theoretical and methodological framework of Sociolinguistics, with an emphasis on studies of linguistic perception and evaluation. It is grounded in the assumption that accent is a socially constructed phenomenon, associated with identities, social representations, and symbolic values, and that listeners' perceptions are systematic, variable, and ordered. To this end, an online questionnaire was developed and administered via Google Forms, in which Brazilian participants aged 18 or older, from different educational backgrounds and of both sexes, evaluated the speech of speakers from the Southeast capitals across four dimensions: prestige, beauty, correctness, and pleasantness, totaling 1,401 valid responses. The results indicate that the accents of Rio de Janeiro and São Paulo were most frequently associated with prestige; Belo Horizonte stood out in the dimensions of beauty and pleasantness; and Vitória was perceived as the most "correct" variety. These findings show that linguistic evaluations are not homogeneous and reflect historically constructed linguistic and social stereotypes, influenced by both linguistic factors and the listeners' social profiles. By discussing these patterns, the study highlights the importance of linguistic respect, particularly with regard to the regional varieties typical of each area, recognizing them as legitimate forms of expression of Brazilian Portuguese. Thus, the research contributes to the debate on language attitudes, variation, and identity in the contemporary urban context.

[1] Universidade Federal do Espírito Santo (UFES), leilatesch@gmail.com, https://orcid.org/0000-0003-3919-1230.

Recebido em 20/01/2026 - Aceito em 10/06/2026

KEYWORDS: Sociolinguistics. Perception studies. Accent. Identity. Southeast region of Brazil.

1. Introdução

Este artigo tem como objeto de estudo a avaliação perceptiva das variedades linguísticas associadas às capitais da Região Sudeste do Brasil — Vitória/ES, Rio de Janeiro/RJ, São Paulo/SP e Belo Horizonte/MG — a partir de julgamentos de ouvintes sobre diferentes dimensões valorativas do falar regional. Fundamentada nos pressupostos teórico-metodológicos da Sociolinguística, com especial ênfase nos estudos de percepção, a pesquisa parte da compreensão de que as variedades linguísticas não apenas diferem estruturalmente, mas são socialmente avaliadas e indexam identidades, relações de poder e hierarquias simbólicas. Para tanto, foram coletados dados por meio de um instrumento de pesquisa aplicado a participantes brasileiros de diferentes perfis sociais, os quais foram convidados a identificar e avaliar os sotaques das capitais analisadas segundo os critérios de prestígio, correção, beleza e agradabilidade, além de responder a questões de caráter projetivo sobre essas quatro variedades linguísticas. A análise dos resultados busca evidenciar como representações sociais historicamente construídas sobre essas capitais se refletem nas percepções e avaliações atribuídas às suas variedades linguísticas.

A discussão sobre as variedades linguísticas associadas às capitais do Sudeste dialoga diretamente com a tradição dos estudos dialetológicos no Brasil, em especial com a proposta de divisão dialetal apresentada por Nascentes (1953), em seu livro intitulado *O linguajar carioca*. Ao tratar da constituição dos falares brasileiros, Nascentes observa

> o falar brasileiro, apesar de sua relativa uniformidade, apresenta variações bem características. (...) Tal foi o que se deu em nosso país; a enorme extensão territorial sem fáceis comunicações interiores quebrou a unidade do falar, fragmentando-o em subfalares." (Nascentes, 1953, p 17)

Para o autor, tal fragmentação não se explica apenas por fatores geográficos, mas também pelo "modo diferente de povoamento das diversas regiões" (Nascentes, 1953, p. 19), aspecto fundamental para compreender a configuração histórica e social das variedades regionais. Ainda assim, Nascentes ressalta que essa diversidade não compromete a unidade linguística do português no Brasil, uma vez que "o falante do norte do país não tem a menor dificuldade em entender o falante do sul, embora ocorram diferenças na fonética, na sintaxe e no léxico" (Nascentes, 1953, p. 8). A divisão dialetal proposta pelo autor, representada no mapa (imagem 1), constitui, portanto, um marco teórico importante para este estudo, na medida em que evidencia a coexistência entre unidade e diversidade linguística no território brasileiro, permitindo situar as capitais do Sudeste em um quadro mais amplo de variação regional que continua a informar, direta ou indiretamente, as percepções e avaliações sociolinguísticas contemporâneas.

Imagem 1: Divisão dialetal do Brasil proposta por Antenor Nascentes. Há seis subfalares: Amazônico, Nordestino, Baiano, Mineiro, Fluminense, Sulista, além de um território indefinido.



Fonte: (Nascentes, 1953, p. 19)

Cumpre, no entanto, relativizar o alcance explicativo da proposta de Antenor Nascentes à luz de transformações socioculturais e teórico-metodológicas mais recentes. Conforme apontam abordagens contemporâneas da Sociolinguística (Votre; Rocaratti, 2008; Mollica; Braga, 2017; Oliveira e Silva; Scherre, 1996) e da Dialetologia (Cardoso, 2014; Cardoso, 2010; Brandão, 2005; Aguilera, 2007), a configuração da diversidade linguística brasileira sofreu alterações significativas ao longo do século XX e início do XXI, em decorrência de fatores como a interiorização do país após a fundação de Brasília, a ampliação dos fluxos migratórios internos e a intensa difusão dos meios de comunicação de massa e digitais. Tais processos contribuíram para uma maior integração entre regiões historicamente mais isoladas, relativizando explicações baseadas predominantemente na fragmentação geográfica. Além disso, estudos recentes têm demonstrado que as delimitações dialetais propostas por Nascentes não correspondem, de modo direto, às percepções e avaliações dos próprios falantes, que tendem a organizar a variação linguística a partir de outros critérios identitários, sociais e ideológicos. Nesse contexto, destacam-se as pesquisas desenvolvidas pelo grupo Programa de Estudos sobre o Uso da Língua (Peul), liderado por Anthony Julius Naro, que, ancoradas na tradição variacionista laboviana, têm contribuído de forma decisiva para a descrição empírica das variedades do português brasileiro, evidenciando a sistematicidade da variação e a atuação de fatores linguísticos e sociais na configuração dos usos (Naro, 1981; Naro; Scherre, 2007; Scherre; Naro, 1993). A partir dessa iniciativa pioneira, diversos outros grupos de pesquisa foram constituídos no país, ampliando o escopo dos estudos variacionistas, entre os quais o Varsul, o ValPB e o PortVix, que têm contribuído significativamente para a consolidação e expansão da Sociolinguística Variacionista no Brasil.

Para o estudo apresentado neste artigo, torna-se necessário explicitar o conceito de *sotaque* adotado. A opção pelo uso desse termo justifica-se por seu maior grau de familiaridade junto aos participantes, o que favorece uma compreensão mais imediata da proposta da pesquisa e, sobretudo,

uma maior aderência entre os instrumentos de coleta e as formas pelas quais os próprios falantes conceituam e avaliam a variação linguística. Nesse sentido, *sotaque* é aqui compreendido como uma categoria êmica, isto é, como uma noção mobilizada pelos próprios participantes para descrever diferenças entre variedades linguísticas.

Ainda que, em descrições mais tradicionais, o termo esteja frequentemente associado a traços fonéticos e fonológicos — segmentais e suprassegmentais —, os dados analisados revelam que, para os participantes, o conceito de *sotaque* extrapola o domínio estritamente fônico, abrangendo também aspectos léxicos e morfossintáticos associados a usos recorrentes em contextos regionais específicos. Nessa direção, Leite e Callou (2010, p. 51) propõem a noção de "sotaques sintáticos", ao considerarem que o nível morfossintático também pode ser afetado, em maior ou menor grau. Nesse sentido, os respondentes mencionam, por exemplo, expressões típicas de determinadas localidades, além do emprego diferencial de diversos fenômenos morfossintáticos, como as formas pronominais, com a variação entre *tu* e *você* na referência à segunda pessoa, como elementos constitutivos do que reconhecem e avaliam como *sotaque*. Essa ampliação do conceito, ancorada na percepção dos falantes, reforça a natureza socialmente construída das variedades linguísticas e orienta as análises desenvolvidas ao longo deste artigo.

Partindo dessa concepção ampliada de sotaque e do reconhecimento de que as avaliações linguísticas estão intrinsecamente relacionadas a representações sociais mais amplas, faz-se relevante considerar os contextos histórico-demográficos das capitais investigadas. As trajetórias de formação urbana, os diferentes papéis político-administrativos e econômicos assumidos por Rio de Janeiro, São Paulo, Belo Horizonte e Vitória, bem como suas dinâmicas socioculturais contemporâneas, constituem elementos centrais para a compreensão das percepções e atitudes linguísticas observadas neste estudo.

Imagem 2: Localização da região Sudeste no Brasil e delimitação do *locus* da pesquisa: capitais de Minas Gerais, Rio de Janeiro, São Paulo e Espírito Santo



Fonte: Disponível em: https://www.todamateria.com.br/estados-do-sudeste/. Acesso em: 28 de abril de 2026

A caracterização histórico-demográfica das capitais da Região Sudeste - Rio de Janeiro, São Paulo, Belo Horizonte e Vitória - constitui um passo fundamental para a compreensão dos processos de avaliação linguística investigados neste estudo. As trajetórias de formação urbana, os distintos papéis político-administrativos e econômicos assumidos por essas cidades ao longo da história, bem como suas dimensões populacionais e dinâmicas socioculturais contemporâneas, configuram contextos específicos de produção e circulação das variedades linguísticas nelas associadas. Nesse sentido, as percepções e atitudes em relação aos sotaques dessas capitais não podem ser dissociadas de representações sociais mais amplas, historicamente construídas, que articulam língua, identidade, poder e pertencimento regional.

A cidade do Rio de Janeiro, fundada em 1565, desempenhou papel central na história política, administrativa e cultural do Brasil, tendo sido capital do país por quase dois séculos, entre o período colonial e a primeira metade do século XX. Essa posição histórica contribuiu para a projeção nacional e internacional da cidade, bem como para a consolidação de práticas culturais amplamente difundidas no imaginário brasileiro. De acordo com o Censo Demográfico de 2022, a cidade do Rio de Janeiro conta com cerca de 6,2 milhões de habitantes, configurando-se como a segunda capital mais populosa do país. Sua trajetória histórica e seu protagonismo político e cultural ajudam a compreender a recorrente associação de seu sotaque a valores de prestígio e visibilidade social.

Fundada em 1554, a cidade de São Paulo passou por um processo de crescimento acelerado sobretudo a partir do século XIX, impulsionado inicialmente pelo ciclo do café e, posteriormente, pela industrialização e pela consolidação do setor financeiro. Segundo dados do Censo de 2022, a capital paulista possui mais de 11 milhões de habitantes, sendo a cidade mais populosa do Brasil e o principal centro econômico do país. Sua formação foi marcada por intensos fluxos migratórios internos e internacionais, o que resultou em uma população altamente heterogênea do ponto de vista social, cultural e linguístico. Esse contexto favorece a construção de representações do falar paulistano associadas à modernidade, ao poder econômico e à diversidade linguística, principalmente ao observar processos de acomodação linguística.

Belo Horizonte foi fundada em 1897 como capital planejada do estado de Minas Gerais, concebida dentro dos ideais republicanos de modernização e racionalização do espaço urbano. Diferentemente das demais capitais analisadas, sua fundação não está vinculada diretamente ao período colonial, mas a um projeto político-administrativo de reorganização estadual. Conforme o Censo de 2022, a capital mineira abriga aproximadamente 2,3 milhões de habitantes, integrando uma das principais regiões metropolitanas do país. Ao longo do século XX, Belo Horizonte consolidou-se como importante polo educacional e cultural, frequentemente associada, no imaginário social, a valores de hospitalidade e cordialidade, elementos que podem influenciar avaliações positivas relacionadas à agradabilidade e à beleza de seu sotaque.

A cidade de Vitória, capital do Espírito Santo, fundada em 1551, possui origem colonial e figura entre as capitais mais antigas do Brasil, tendo desempenhado papel estratégico na organização

político-administrativa da capitania. Segundo o Censo de 2022, Vitória apresenta cerca de 340 mil habitantes, sendo uma das capitais menos populosas do Brasil, embora integre uma região metropolitana significativamente mais ampla. Seu menor porte populacional, aliado a uma posição geográfica intermediária entre grandes centros como Rio de Janeiro e São Paulo, contribui para a construção de representações sociais específicas acerca de sua variedade linguística, sendo, provavelmente, a variedade menos reconhecida e estando associada à percepção de maior neutralidade ou correção linguística.

Dessa forma, as variedades linguísticas associadas às quatro capitais do Sudeste são compreendidas, neste estudo, não apenas como conjuntos de traços estruturais da língua, mas como fenômenos socialmente avaliados, atravessados por ideologias linguísticas e por representações historicamente construídas sobre seus espaços urbanos e seus falantes. As avaliações de prestígio, correção, beleza e agradabilidade atribuídas aos sotaques de Vitória, Rio de Janeiro, São Paulo e Belo Horizonte refletem, em grande medida, as posições simbólicas ocupadas por essas capitais no cenário nacional, bem como as relações de poder que permeiam o imaginário social brasileiro. Ao articular a análise das percepções linguísticas aos contextos histórico-demográficos dessas cidades, o presente estudo busca contribuir para a compreensão das atitudes linguísticas no Brasil contemporâneo, reafirmando a centralidade da diversidade linguística e do respeito às variedades regionais no campo da Sociolinguística.

Por fim, este estudo integra um número temático dedicado ao professor Anthony Julius Naro, cuja trajetória acadêmica foi decisiva para a consolidação da Sociolinguística Variacionista no Brasil. Suas contribuições pioneiras aos estudos da variação e da mudança linguística estabeleceram bases teórico-metodológicas sólidas para a investigação científica das variedades do português brasileiro, em especial aquelas historicamente estigmatizadas. Ao articular língua, sociedade e história, Naro reafirmou o estatuto científico da diversidade linguística e contribuiu de modo significativo para a formação de gerações de linguistas. Nesse sentido, ao investigar as avaliações perceptivas dos sotaques das capitais do Sudeste, este trabalho alinha-se ao legado intelectual do homenageado, ao compreender a variação linguística como fenômeno sistemático, socialmente condicionado e atravessado por ideologias linguísticas.

2. A Sociolinguística e os estudos de percepção/avaliação

Este estudo insere-se no campo da Sociolinguística Variacionista, cuja principal contribuição consiste na compreensão da língua como um sistema heterogêneo e variável, estruturado por condicionamentos linguísticos e sociais. Nesse quadro teórico, as diferentes formas de realização linguística são entendidas como portadoras de significados sociais, indexando identidades, estilos de fala, pertencimentos regionais e posições sociais.

No interior dessa tradição, os estudos de percepção e avaliação linguística assumem papel central ao deslocarem o foco analítico da produção para a recepção da variação. Importa destacar

que, desde os primeiros estudos variacionistas, Labov já incorporava a dimensão avaliativa em suas investigações sobre o inglês[2], ao demonstrar que diferentes variantes linguísticas eram socialmente estratificadas e objeto de julgamentos explícitos e implícitos por parte dos falantes. Ao evidenciar que formas linguísticas acionam avaliações sociais relacionadas a prestígio, correção e status, essas pesquisas inaugurais apontaram que a variação não apenas reflete diferenças estruturais, mas também carrega significados sociais compartilhados. Nesse sentido, a vertente dos estudos de percepção dentro da Sociolinguística investiga não apenas como as formas linguísticas são usadas, mas sobretudo como são percebidas, interpretadas e avaliadas, tornando visíveis as ideologias linguísticas que orientam tais julgamentos.

Os estudos de percepção linguística buscam observar como certos significados sociais se associam a usos linguísticos específicos, partindo da premissa de que a escuta e o processamento da fala são atravessados por expectativas, estereótipos e conhecimentos socioculturais compartilhados. Conforme argumenta Campbell-Kibler (2006), essa abordagem procura compreender como diferentes formas linguísticas são ouvidas, categorizadas e interpretadas pelos membros de uma comunidade, permitindo identificar as relações entre variação linguística e construção de significados sociais. Nessa perspectiva, a percepção não é um processo neutro, mas um fenômeno socialmente situado, no qual elementos linguísticos acionam inferências sobre origem regional, escolaridade, status social e identidade dos falantes.

No contexto brasileiro, marcado por ampla diversidade regional e por fortes desigualdades sociais, os estudos de avaliação linguística revelam-se particularmente relevantes. As percepções sobre sotaques regionais frequentemente se articulam a discursos de prestígio, correção e legitimidade, muitas vezes associados a centros políticos e econômicos, em detrimento de variedades percebidas como periféricas ou não padrão. Desse modo, investigar como os falantes avaliam os sotaques das capitais do Sudeste permite compreender não apenas padrões de julgamento linguístico, mas também as ideologias que sustentam representações sociais sobre língua, região e poder.

Avançando nessa perspectiva, Preston (2013) sistematiza os estudos de percepção e avaliação linguística ao propor um conjunto de questões centrais voltadas à compreensão de como os falantes percebem, categorizam e avaliam a variação linguística. Segundo o autor, investigações nessa área buscam responder, entre outras, às seguintes questões: como os falantes reconhecem e distinguem variedades linguísticas; que características associam a determinados modos de falar; quais valores sociais - como prestígio, correção ou estigmatização - são atribuídos a essas variedades; e de que modo tais avaliações se relacionam a crenças, atitudes e ideologias linguísticas compartilhadas socialmente. Ao enfatizar que as percepções linguísticas constituem formas de conhecimento socialmente construídas sobre a língua, Preston destaca que essas avaliações não se limitam a aspectos estritamente linguísticos, mas se articulam a representações mais amplas sobre grupos sociais, regiões e identidades. Essas contribuições são particularmente relevantes para o presente estudo, uma vez que orientam a análise das avaliações atribuídas aos sotaques das capitais da região sudeste do

[2] Um dos exemplos é o estudo sobre as estruturas negativas em inglês descrito em Labov (2008 [1972], p. 225-233)

Brasil, permitindo compreender como julgamentos linguísticos refletem e reproduzem hierarquias simbólicas no imaginário social.

É a partir desse enquadramento teórico que o presente estudo analisa as avaliações perceptivas atribuídas aos sotaques de Vitória, Rio de Janeiro, São Paulo e Belo Horizonte, buscando evidenciar como a variação linguística é socialmente interpretada e como tais interpretações refletem posições simbólicas ocupadas por essas capitais no imaginário social brasileiro.

3. Procedimentos metodológicos

Os procedimentos metodológicos adotados para a realização desta pesquisa[3] sobre a avaliação dos sotaques das capitais dos estados da região sudeste – mais especificamente, das cidades de Belo Horizonte/MG, Rio de Janeiro/RJ, São Paulo/SP e Vitória/ES[4] – seguiram as seguintes etapas: i) elaboração de formulário *on-line* via *google forms*; ii) aplicação desse questionário e iii) quantificação e análise dos resultados. Para efeito deste estudo, foram consideradas apenas as respostas de participantes maiores de 18 anos de idade, brasileiros, do sexo masculino e feminino e de diversos níveis de escolaridade.

O questionário, intitulado *Sotaques do Brasil,* que explorou diferentes dimensões relacionadas à percepção e avaliação dos sotaques brasileiros, é composto por cinco seções: i) apresentação da pesquisa e acesso ao Termo de Consentimento Livre e Esclarecido (TCLE)[5]; ii) perguntas referentes às informações pessoais[6] e ao perfil sociodemográfico, incluindo local de nascimento, local de residência, idade, gênero e escolaridade; iii) questões relacionadas ao reconhecimento de sotaques brasileiros; iv) questionamentos voltados à avaliação dos sotaques das capitais da região Sudeste do Brasil; e v) perguntas específicas sobre o sotaque capixaba.

Neste artigo, descrevemos e analisamos as questões referentes à quarta seção do formulário, explorando, de modo específico, as seguintes perguntas: i) como você avalia, em relação ao *prestígio*, o modo de falar das pessoas que residem nas capitais dos estados da região Sudeste?; ii) como você avalia, em relação à *beleza*, o modo de falar das pessoas que residem nas capitais dos estados da região Sudeste?; iii) como você avalia, em relação à *correção*, o modo de falar das pessoas que residem nas capitais dos estados da região Sudeste? e iv) como você avalia, em relação à *agradabilidade*, o modo de falar das pessoas que residem nas capitais dos estados da região Sudeste?

---

[3] Esta pesquisa foi submetida ao Comitê de Ética e foi aprovada sob CAAE n° 38296720.1.0000.8142.

[4] Neste artigo, o *locus* de análise foi delimitado às capitais dos estados da região Sudeste — Belo Horizonte/MG, Rio de Janeiro/RJ, São Paulo/SP e Vitória/ES. Tal recorte justifica-se pela dificuldade de associar, de forma homogênea, uma única variedade linguística a cada estado como um todo, dada a reconhecida heterogeneidade interna dessas unidades federativas. Nesse sentido, a referência às capitais mostrou-se metodologicamente mais adequada, uma vez que facilita a identificação, por parte dos participantes, de traços linguísticos socialmente indexados a esses espaços urbanos

[5] Para prosseguir no preenchimento do formulário, o participante deveria obrigatoriamente assinalar a opção "Sim, aceito participar"

[6] Com a ressalva de que nenhuma informação de identificação pessoal seria solicitada, como nome ou qualquer outro dado que permitisse a identificação do participante.

No que se refere à delimitação dessas perguntas, vale destacar que foram elaboradas com base na proposta de Preston (2013), que investigou avaliações perceptuais semelhantes entre respondentes dos estados de Michigan e Alabama, nos Estados Unidos. Em seu estudo, o autor evidenciou resultados relevantes acerca das atitudes e crenças linguísticas dos falantes, especialmente no que diz respeito à percepção e à avaliação de variedades regionais desses dois estados. Inspirada por essa abordagem, a presente pesquisa busca adaptar tais procedimentos ao contexto brasileiro, considerando as especificidades sociolinguísticas da região Sudeste.

Além das questões de avaliação direta dos sotaques das capitais da região Sudeste em relação aos traços de *prestígio*, *beleza*, *correção* e *agradabilidade*, o questionário contemplou duas perguntas adicionais de caráter projetivo, com o objetivo de acessar de forma indireta as atitudes linguísticas dos participantes. A primeira delas foi: *"Em um mundo imaginário, se você pudesse escolher com qual sotaque nasceria, entre as capitais dos estados da região Sudeste, qual você escolheria?"*. A segunda pergunta foi: *"E qual sotaque dessas capitais você não gostaria de ter?"*. Essas questões permitiram observar preferências e rejeições simbólicas associadas aos sotaques analisados, contribuindo para uma compreensão mais ampla das representações sociolinguísticas mobilizadas pelos participantes e complementando os dados obtidos por meio das avaliações explícitas.

O questionário foi divulgado nas redes sociais da pesquisadora via *WhatsApp*, *Telegram*, *Instagram*, *Facebook* e e-*mail*. Também contamos com a divulgação realizada por amigos e por amigos de amigos, nessas mesmas redes sociais, de 17 de março a 20 de abril de 2021.

Ao todo, foram preenchidos 1472 instrumentos. Desses, seis participantes se recusaram a participar, 58 foram excluídos por não atenderem ao critério etário mínimo – por serem menores de 18 anos – e sete foram eliminados por não terem nascido no Brasil. Assim, o *corpus* final analisado é composto por 1.401 questionários válidos, respondidos por participantes com idades entre 18 e 76 anos (com idade média de 33 anos). Do total, 984 identificaram-se como do sexo feminino (70%), 406 do sexo masculino (29%) e 11 participantes preferiram não dizer ou assinalaram a opção outro (menos de 1%). Em relação ao grau de escolaridade, a maioria dos participantes possui pós-graduação (45%) ou graduação (45%), enquanto uma parcela menor tem ensino médio ou ensino fundamental (10%).

Ressalta-se que também foram controladas as informações relativas à cidade e ao estado de nascimento dos participantes, bem como ao local de residência atual. Em razão da diversidade de cidades informadas, apresentamos apenas a distribuição por estados de nascimento, a saber: Acre, 2 participantes; Alagoas, 3; Amapá, 3; Amazonas, 3; Bahia, 68; Ceará, 46; Distrito Federal, 5; Espírito Santo, 746; Goiás, 12; Maranhão, 6; Mato Grosso, 1; Mato Grosso do Sul, 4; Minas Gerais, 169; Pará, 11; Paraíba, 4; Paraná, 20; Pernambuco, 13; Piauí, 1; Rio de Janeiro, 114; Rio Grande do Norte, 5; Rio Grande do Sul, 54; Rondônia, 4; Santa Catarina, 15; Sergipe, 2 e São Paulo, 122. Observa-se que a maior parte dos participantes nasceu no estado do Espírito Santo, correspondendo a aproximadamente 53% do total. Destaca-se, ainda, a participação de respondentes oriundos de praticamente todos os estados brasileiros, à exceção de Roraima e Tocantins.

Em relação aos estados de residência atual, novamente verifica-se a predominância de participantes residentes no estado do Espírito Santo, que correspondem a 55% dos respondentes. Além disso, não houve participação de moradores de apenas dois estados brasileiros: Acre e Tocantins. A distribuição por estado de residência foi a seguinte: Alagoas, 1 participante; Amapá, 1; Amazonas, 2; Bahia, 44; Ceará, 40; Distrito Federal, 5; Espírito Santo, 771; Goiás, 14; Maranhão, 4; Mato Grosso, 1; Mato Grosso do Sul, 2; Minas Gerais, 93; Pará, 7; Paraíba, 1; Paraná, 20; Pernambuco, 5; Piauí, 1; Rio de Janeiro, 46; Rio Grande do Norte, 5; Rio Grande do Sul, 46; Rondônia, 1; Roraima, 1; Santa Catarina, 22, São Paulo, 56 e Sergipe 1. Para nossa surpresa, tivemos a participação de brasileiros que atualmente moram em outros países, a saber: 7 em Portugal; 3 nos Estados Unidos; 3 no Peru; 1 no Reino Unido; 1 na Holanda; 1 na Turquia; 1 na França; 1 no Canadá; 1 na Espanha e 1 na República Tcheca. Essa diversidade de localidades revela-se particularmente relevante para a pesquisa, na medida em que possibilita analisar a avaliação dos sotaques das capitais da região sudeste tanto por brasileiros residentes em diferentes regiões do país quanto por aqueles que vivem no exterior.

Com o objetivo de verificar a avaliação dos sotaques das quatro capitais dos estados do Sudeste em relação aos traços de *prestígio*, *beleza*, *correção* e *agradabilidade*, apresentamos, a seguir, a análise das 1.401 respostas obtidas para essas questões.

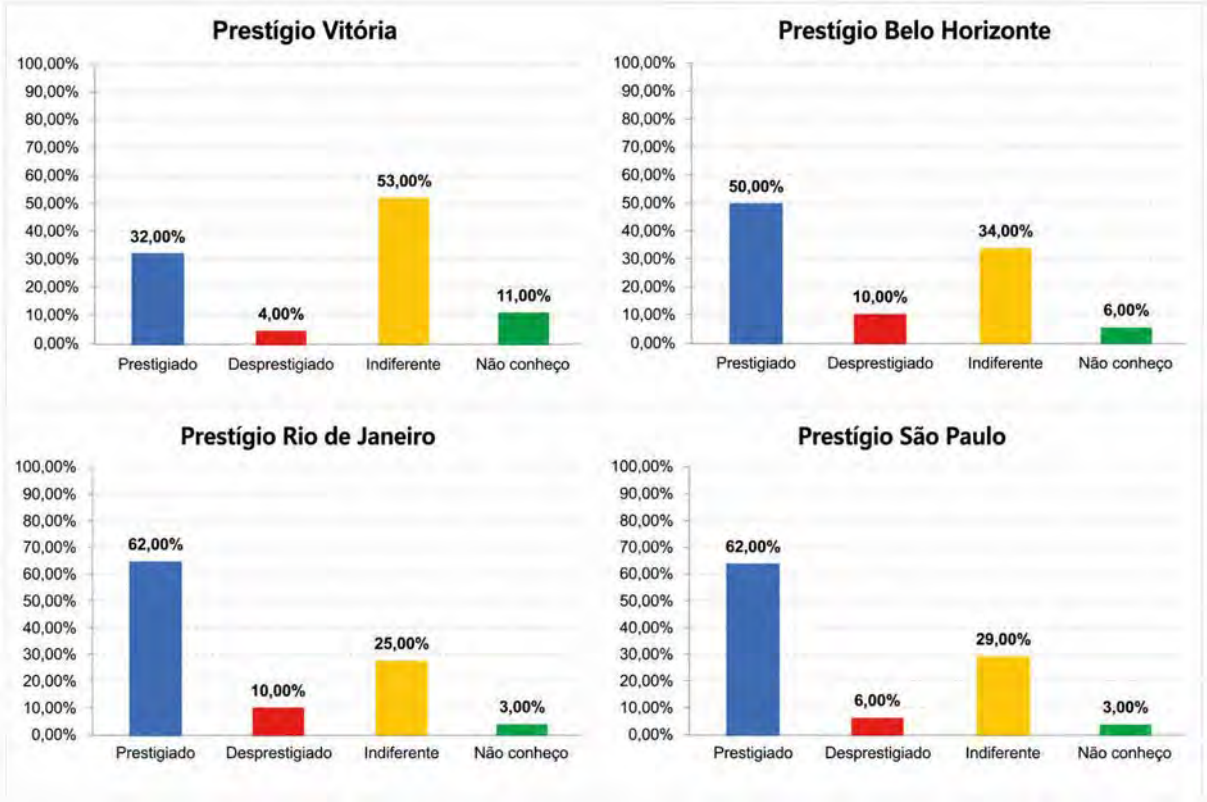
4. Resultados das avaliações dos sotaques das capitais da região sudeste

Nesta seção, apresentamos e discutimos os resultados relativos à avaliação dos sotaques das capitais dos estados da região Sudeste do Brasil – Belo Horizonte, Rio de Janeiro, São Paulo e Vitória – com base nos dados obtidos por meio do questionário *Sotaques do Brasil*.

A análise focaliza, inicialmente, as avaliações explícitas atribuídas pelos participantes aos modos de falar dessas capitais a partir de quatro dimensões: *prestígio*, *beleza*, *correção* e *agradabilidade*. Esses parâmetros permitem observar padrões de valoração sociolinguística e processos de hierarquização dos sotaques no imaginário dos falantes.

Complementarmente, são examinados os resultados de duas questões de natureza projetiva, voltadas à identificação de preferências e rejeições simbólicas associadas às variedades analisadas em um cenário hipotético: a escolha do sotaque com o qual os participantes gostariam de nascer e a indicação daquele que não gostariam de ter. Essas questões possibilitam acessar atitudes linguísticas de forma indireta, ampliando a compreensão dos mecanismos de avaliação mobilizados pelos respondentes.

Gráfico 1: Frequência relativa das avaliações de *prestígio* dos sotaques das capitais da região Sudeste do Brasil



Fonte: Elaboração da autora

Os resultados referentes à avaliação do *prestígio* dos sotaques das capitais da região Sudeste revelam diferenças significativas na forma como os modos de falar de Vitória, Belo Horizonte, Rio de Janeiro e São Paulo são percebidos pelos participantes, evidenciando processos distintos de valoração sociolinguística.

O sotaque de *Vitória* apresenta um padrão singular de avaliação: a categoria *indiferente* concentra a maior proporção de respostas, seguida pela avaliação *prestigiado*. Observa-se, ainda, um percentual relativamente expressivo de participantes que declararam *não conhecer* o sotaque capixaba, o que sugere menor visibilidade nacional desse modo de falar entre as capitais da região sudeste do Brasil. A baixa incidência da categoria *desprestigiado* indica que, embora pouco reconhecido, o sotaque de Vitória não é fortemente associado a julgamentos negativos, mas ocupa uma posição marcada pela neutralidade avaliativa e pela invisibilidade sociolinguística.

No caso de *Belo Horizonte*, verifica-se uma distribuição mais equilibrada entre as categorias *prestigiado* e *indiferente*, com predominância da avaliação positiva. Diferentemente de Vitória, o sotaque belo-horizontino apresenta menor percentual de desconhecimento e uma presença um pouco mais elevada da categoria *desprestigiado*, o que aponta para um maior grau de reconhecimento social e, consequentemente, maior exposição a avaliações polarizadas.

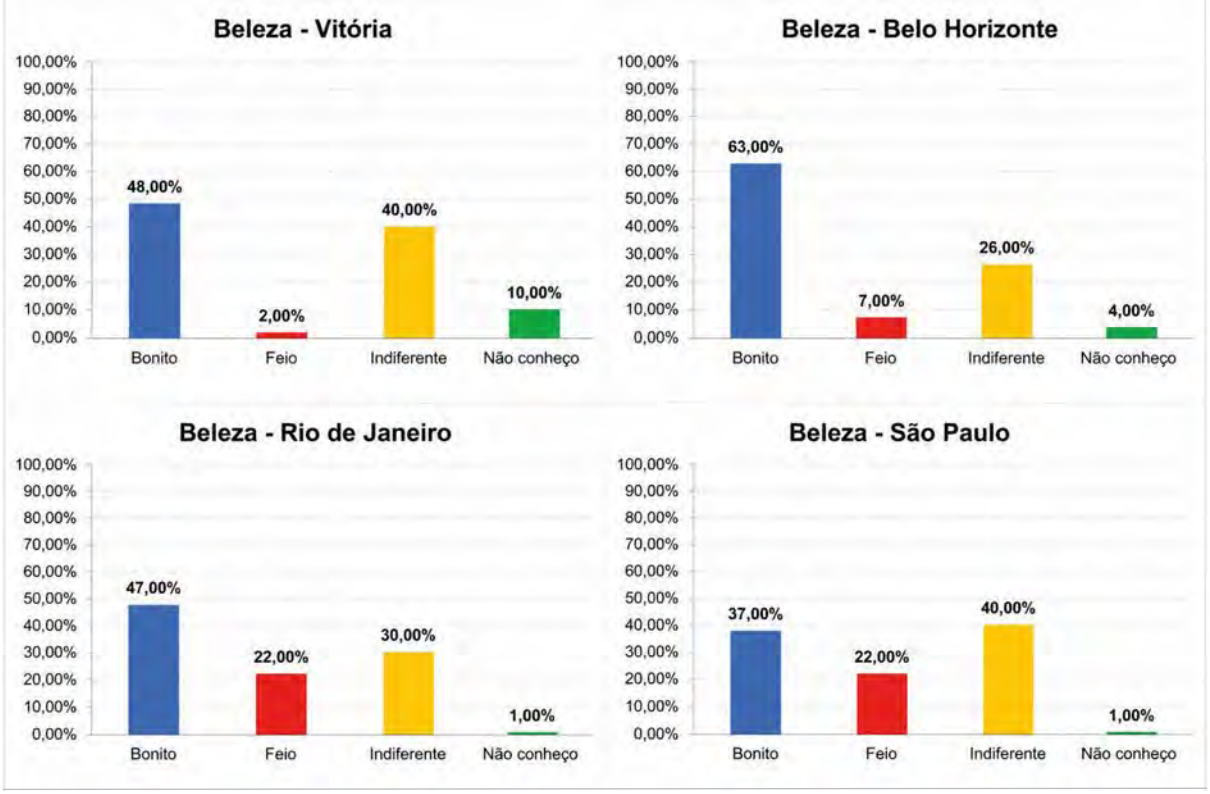
O sotaque do Rio de Janeiro destaca-se como aquele que concentra uma das mais altas taxas de avaliação como prestigiado, superando amplamente as categorias *indiferente* e *desprestigiado*. Esse resultado pode ser interpretado à luz do que discutem Leite e Callou (2010), ao apontarem que a noção de variedade de prestígio no Brasil historicamente se articula em torno de determinados centros urbanos, especialmente Rio de Janeiro, São Paulo e Salvador. No caso específico do Rio de Janeiro, as autoras ressaltam que seu lugar de destaque se deve, prioritariamente, a razões extralinguísticas: o fato de o Rio de Janeiro estar geograficamente no centro de uma polaridade norte/sul, ser centro político há mais tempo, capital da Colônia desde 1763, e ser uma área cuja linguagem culta tende a apresentar menor número de marcas locais e regionais, com uma tendência universalista, dentro do país. (Leite; Callou, 2010, p. 9-10)

Nesse sentido, os dados deste estudo parecem corroborar essa interpretação, ao evidenciarem não apenas o alto índice de avaliações positivas, mas também o reduzido percentual de respostas do tipo *não conheço*, o que indica um amplo reconhecimento e difusão do sotaque carioca em escala nacional. Ainda assim, a presença de avaliações negativas sugere que esse prestígio não se distribui de forma homogênea, coexistindo com processos de estigmatização que revelam a complexidade das hierarquias sociolinguísticas no Brasil.

Situação semelhante é observada em relação ao sotaque de *São Paulo*, que também apresenta elevada concentração de respostas na categoria *prestigiado*, acompanhada por uma parcela significativa de avaliações *indiferente*. O sotaque paulistano, assim como o carioca, é amplamente reconhecido, o que se reflete nos percentuais extremamente baixos da categoria *não conheço*. A menor incidência de avaliações negativas sugere que o prestígio atribuído a esse modo de falar está fortemente associado à centralidade econômica e institucional da cidade.

De modo geral, os dados indicam que os sotaques de Rio de Janeiro e São Paulo ocupam posições de maior prestígio na hierarquia sociolinguística percebida pelos participantes, enquanto Belo Horizonte assume uma posição intermediária e Vitória se caracteriza por um padrão de menor reconhecimento e maior neutralidade avaliativa. Esses resultados reforçam a relação entre prestígio linguístico, visibilidade social e centralidade política, econômica e midiática.

Gráfico 2: Frequência relativa das avaliações de *beleza* dos sotaques das capitais da região Sudeste do Brasil



Fonte: Elaboração da autora

Os dados representados no gráfico 2 revelam que o sotaque de *Belo Horizonte* é o que apresenta a avaliação mais positiva no que se refere ao traço *beleza*, concentrando o maior percentual de respostas na categoria *bonito*. Esse resultado sugere uma percepção amplamente favorável do falar mineiro, frequentemente associado, no imaginário sociolinguístico brasileiro, a valores de suavidade, cordialidade e familiaridade, aspectos que tendem a influenciar julgamentos sobre a língua.

O sotaque de *Vitória* também apresenta uma avaliação predominantemente positiva, com percentual expressivo de respostas na categoria *bonito*, embora acompanhado por uma proporção considerável de avaliações *indiferente*. Esse dado pode ser interpretado à luz da menor saliência social e midiática do sotaque capixaba no cenário nacional, o que contribui para avaliações menos polarizadas e para um maior índice de neutralidade avaliativa.

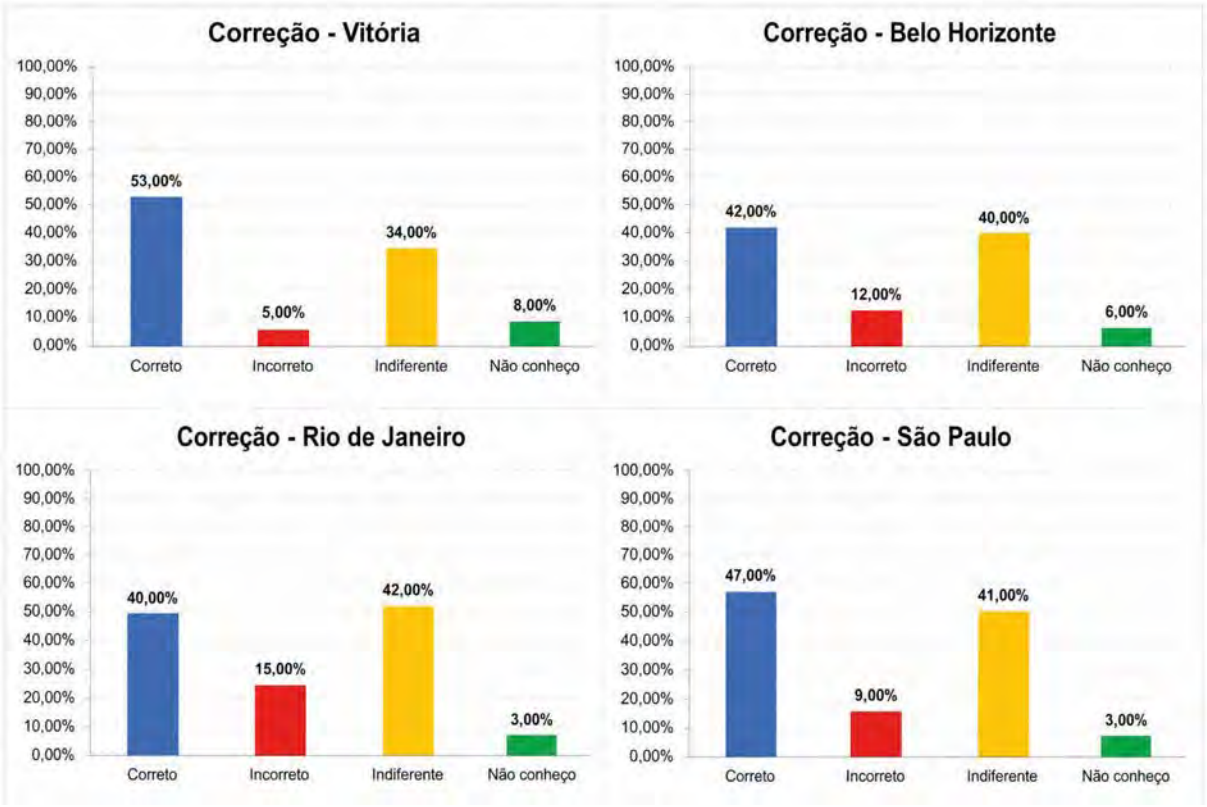
No caso do *Rio de Janeiro*, observa-se uma configuração mais polarizada: embora a categoria *bonito* concentre o maior percentual, há um índice significativo de avaliações *feio*, superior ao observado para Vitória e Belo Horizonte. Esse padrão sugere que o sotaque carioca, por sua forte projeção social e simbólica, tende a suscitar reações avaliativas mais extremadas, refletindo tanto prestígio quanto estigmatização.

Por fim, o sotaque de *São Paulo* apresenta um comportamento distinto, com predomínio da categoria *indiferente*, seguida de percentuais relativamente próximos entre *bonito* e *feio*. Esse resultado

indica uma avaliação menos marcada, o que pode estar relacionado à percepção do sotaque paulistano como mais "neutro" ou "não marcado", sobretudo em função de sua associação a contextos urbanos, institucionais e midiáticos de ampla circulação.

De modo geral, os resultados evidenciam que as avaliações estéticas dos sotaques não se distribuem de forma homogênea entre as capitais analisadas, refletindo processos de avaliação social da variação linguística, nos termos da Teoria da Variação e Mudança Linguística. As diferenças observadas reforçam a ideia de que julgamentos sobre *beleza* linguística estão profundamente imbricados em fatores extralinguísticos, como identidade regional, prestígio social e grau de exposição pública dos falares.

Gráfico 3: Frequência relativa das avaliações de *correção* dos sotaques das capitais da região Sudeste do Brasil



Fonte: Elaboração da autora

Os dados relativos à percepção de *correção* linguística indicam que o sotaque de *Vitória* é aquele que concentra o maior percentual de avaliações na categoria *correto*, acompanhado por um índice relativamente baixo de julgamentos como *incorreto*. Ainda assim, observa-se uma presença significativa da categoria *indiferente*, o que reforça a interpretação de que o sotaque capixaba, embora avaliado positivamente em termos normativos, permanece pouco conhecido no imaginário sociolinguístico.

No caso de *Belo Horizonte*, nota-se uma distribuição mais equilibrada entre as categorias *correto* e *indiferente*, com um percentual mais elevado de avaliações como *incorreto* quando comparado a

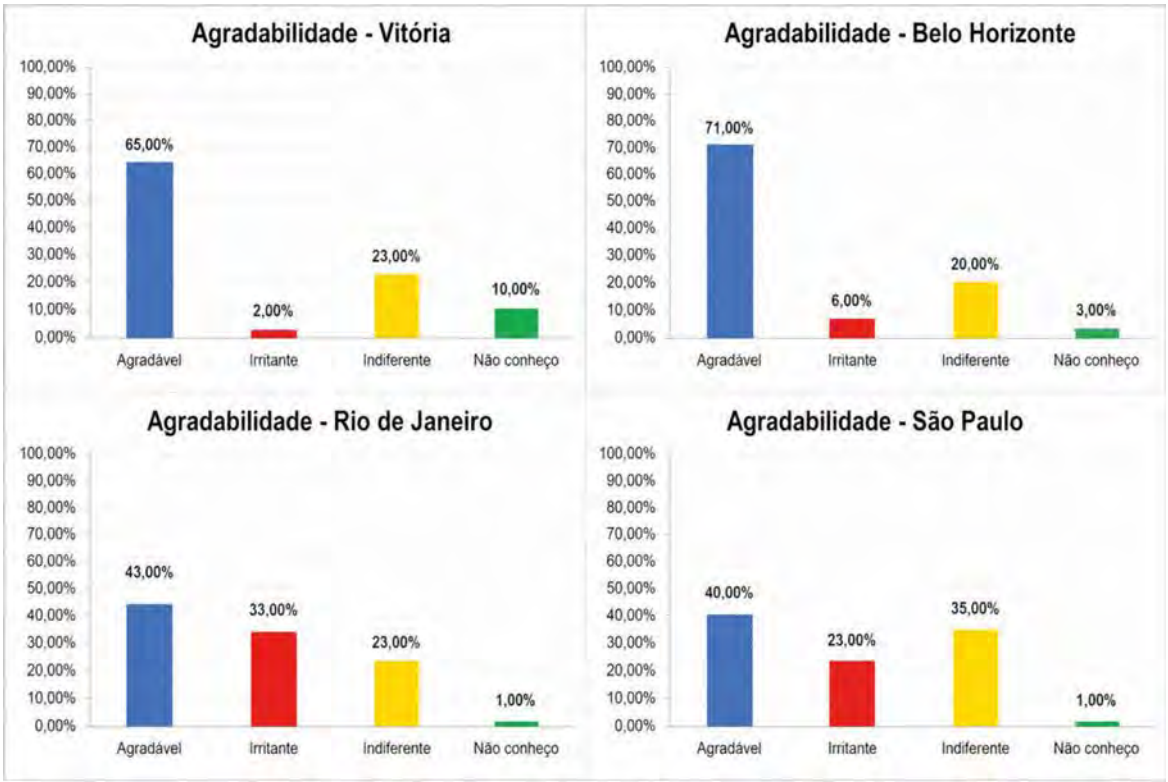
Vitória. Esse padrão sugere que, apesar do reconhecimento positivo frequentemente atribuído ao falar mineiro em termos afetivos e estéticos, sua associação à correção normativa não é tão consensual, revelando tensões entre prestígio encoberto e prestígio aberto.

O sotaque do *Rio de Janeiro* apresenta uma configuração semelhante à de Belo Horizonte, com predominância das categorias *correto* e *indiferente*, mas também com um índice expressivo de avaliações como *incorreto*. Esse resultado pode ser interpretado à luz da forte visibilidade social do sotaque carioca, que, ao mesmo tempo em que goza de prestígio simbólico e midiático, é frequentemente associado a traços estigmatizados do ponto de vista da norma padrão.

Já o sotaque de *São Paulo* revela uma avaliação marcada pela coexistência entre *correto* e *indiferente*, com percentuais intermediários e menor polarização. Esse comportamento pode indicar a percepção do falar paulistano como relativamente próximo da norma de prestígio institucionalizada, sobretudo em função de sua associação a centros econômicos, acadêmicos e midiáticos, o que contribui para sua indexação como linguisticamente adequado ou neutro.

De forma geral, os resultados evidenciam que a noção de *correção* linguística mobilizada pelos participantes desta pesquisa não se ancora exclusivamente em critérios estruturais, mas reflete valores sociais, ideológicos e identitários atribuídos aos diferentes falares regionais. À luz da Teoria da Variação e Mudança Linguística, esses dados reforçam a compreensão de que avaliações de correção constituem um terreno privilegiado para a observação de atitudes linguísticas, nas quais a língua funciona como marcador simbólico de prestígio, poder e pertencimento social.

Gráfico 4: Frequência relativa das avaliações de *agradabilidade* dos sotaques das capitais da região Sudeste do Brasil



Fonte: Elaboração da autora

Os resultados indicam que o sotaque de *Vitória* é avaliado de forma amplamente positiva quanto à *agradabilidade*, concentrando a maior parte das respostas na categoria *agradável* e apresentando índices residuais de avaliação como *irritante*. A presença de um percentual moderado de respostas *indiferente* e *não conheço* reforça, novamente, a interpretação de que o falar capixaba é percebido como pouco saliente, o que favorece julgamentos positivos ou neutros, com baixa rejeição explícita.

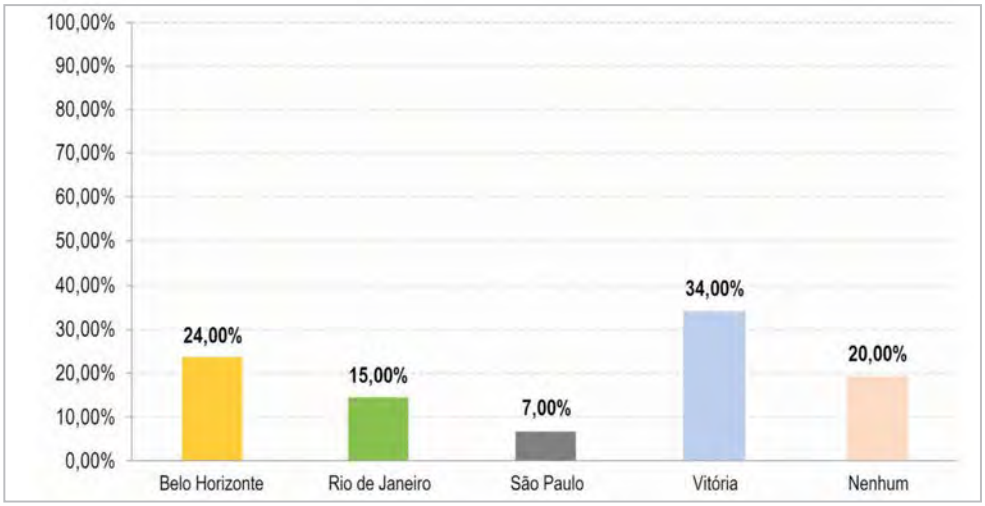
O sotaque de *Belo Horizonte* apresenta o padrão mais favorável entre as capitais analisadas, com o maior percentual de avaliações como *agradável* e baixos índices de respostas *irritante*. Esse resultado dialoga diretamente com a ideia que aponta o falar mineiro como associado a traços de cordialidade, proximidade e simpatia, atributos que se projetam de maneira consistente nas avaliações afetivas.

Em contraste, o sotaque do *Rio de Janeiro* revela uma configuração mais polarizada. Embora a categoria *agradável* ainda concentre parcela expressiva das respostas, observa-se um percentual elevado de avaliações como *irritante*, significativamente superior ao registrado para Vitória e Belo Horizonte. Esse dado sugere que o sotaque carioca, por sua alta visibilidade social e forte indexação identitária, mobiliza reações afetivas mais intensas, oscilando entre apreciação e rejeição.

Por sua vez, o sotaque de *São Paulo* apresenta uma distribuição mais equilibrada entre as categorias *agradável*, *irritante* e *indiferente*, com leve predominância da avaliação *indiferente*. Esse padrão indica uma percepção menos afetivamente marcada, possivelmente associada à imagem do falar paulistano como funcional, urbano e institucional, o que tende a reduzir tanto a idealização quanto a rejeição explícita.

De modo geral, a análise do traço *agradabilidade* evidencia que avaliações de natureza afetiva se mostram particularmente sensíveis a fatores extralinguísticos, como projeção midiática, estereótipos regionais e indexação identitária. À luz da Teoria da Variação e Mudança Linguística, esses resultados confirmam que atitudes em relação aos sotaques constituem um espaço privilegiado para observar como valores sociais são projetados sobre a variação linguística, produzindo padrões avaliativos distintos mesmo entre falares geograficamente próximos.

Gráfico 5: Frequência relativa da escolha do sotaque com o qual os participantes gostariam de nascer entre as capitais da região Sudeste do Brasil



Fonte: Elaboração da autora

Os dados referentes à escolha hipotética do sotaque de nascimento revelam padrões particularmente elucidativos sobre as atitudes linguísticas dos informantes, uma vez que essa pergunta mobiliza dimensões identitárias, afetivas e simbólicas de forma mais direta do que as avaliações anteriores. Observa-se que o sotaque de *Vitória* é o mais frequentemente escolhido, concentrando o maior percentual de respostas. Esse resultado deve ser interpretado à luz do perfil da amostra, uma vez que a maior parte dos respondentes é composta por participantes capixabas, o que tende a favorecer escolhas alinhadas à identidade linguística local.

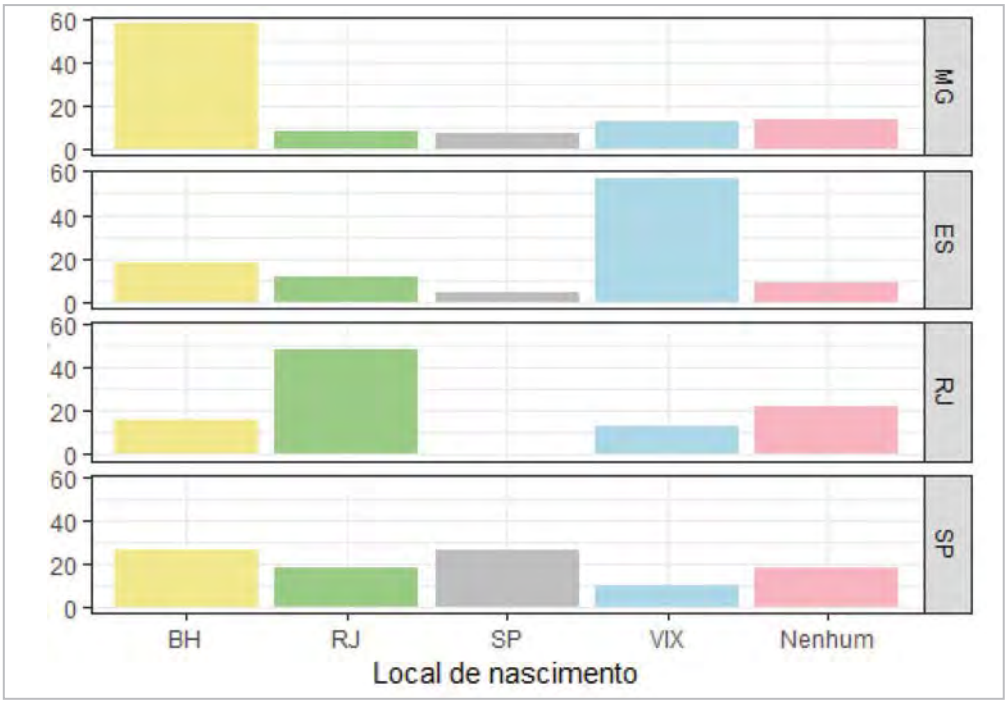
Nesse sentido, a preferência pelo sotaque de Vitória pode ser compreendida como um efeito de autoavaliação positiva e de lealdade linguística, uma vez que os falantes demonstram maior disposição para valorizar e eleger como desejável a variedade com a qual mantêm vínculos identitários e afetivos mais estreitos. Esse dado não invalida, contudo, a leitura positiva do resultado, uma vez que o sotaque capixaba também apresentou avaliações favoráveis ou neutras nos critérios de beleza, correção e agradabilidade, com baixos índices de rejeição explícita.

O sotaque de *Belo Horizonte* ocupa a segunda posição, mantendo um padrão consistente de avaliações positivas ao longo da pesquisa, o que sugere que, mesmo diante do peso identitário da variedade local, o falar mineiro se projeta como uma alternativa socialmente valorizada e afetivamente prestigiada. O sotaque do *Rio de Janeiro* aparece em posição intermediária, refletindo a ambivalência já observada nas análises anteriores, enquanto o sotaque de *São Paulo* apresenta o menor índice de escolha entre as capitais, coerente com sua avaliação predominantemente indiferente.

Por fim, o percentual expressivo da opção *nenhum* indica que parte dos participantes opta por não se identificar plenamente com nenhuma das variedades apresentadas, o que pode sinalizar tanto uma postura crítica em relação aos estereótipos regionais quanto a valorização de uma identidade linguística própria, não necessariamente coincidente com os falares das capitais analisadas.

Em conjunto, os resultados reforçam que a escolha hipotética do sotaque de nascimento é fortemente atravessada por fatores identitários e de pertencimento regional, devendo ser analisada em articulação com o perfil sociogeográfico dos participantes.

Gráfico 6: Distribuição da frequência relativa da escolha do sotaque com o qual os participantes gostariam de nascer entre as capitais da região Sudeste do Brasil e o local de nascimento do participante da pesquisa



Fonte: Elaboração da autora

A distribuição dos dados por local de nascimento dos respondentes evidencia de forma clara o papel da identidade regional e da lealdade linguística nas escolhas relativas ao sotaque desejado. Entre os respondentes nascidos em Minas Gerais (MG), observa-se uma preferência majoritária pelo sotaque de *Belo Horizonte*, o que indica um forte alinhamento entre pertencimento regional e avaliação positiva da variedade local. Esse padrão é consistente com resultados clássicos da sociolinguística das atitudes, segundo os quais falantes tendem a valorizar positivamente a própria variedade como forma de afirmação identitária. Como afirma Freitag (2025, p. 22), "é a partir de nossas experiências que construímos o nosso valor de referência para as indexações sociais que fazemos: quem fala como eu pertence ao meu grupo, quem fala diferente tem sotaque."

Entre os respondentes capixabas (ES), nota-se uma preferência expressiva pelo sotaque de *Vitória*, superando amplamente as demais opções. Esse dado confirma que o resultado global anteriormente observado — no qual Vitória aparece como o sotaque mais escolhido — é fortemente condicionado pela composição da amostra. Ao mesmo tempo, essa preferência pode ser interpretada como um indício de autoavaliação positiva e de fortalecimento simbólico do sotaque capixaba, tradicionalmente descrito na literatura como menos saliente ou menos reconhecido no cenário nacional, como demonstram estudos realizados pelo grupo PortVix.

> os resultados alcançados ao longo de 20 anos de atuação do PortVix[7] contribuíram para a realização desta pesquisa de percepção e os resultados encontrados até o momento em estudos de produção realizados nesse *corpus* convergem com os aqui encontrados, uma vez que a variedade capixaba parece não ter marcas fonológicas, morfológicas, sintáticas e discursivas muito evidentes, por isso sendo muitas vezes caracterizada como um sotaque menos marcado, ou como uma fala mais neutra. (Tesch, 2022, p. 240)

No grupo de respondentes nascidos no Rio de Janeiro (RJ), a escolha do sotaque carioca se destaca de maneira significativa em relação às demais opções, reiterando o padrão de valorização da variedade de origem. Ainda assim, chama atenção a presença relevante da opção *nenhum*, o que sugere uma relação mais ambivalente com o próprio sotaque, coerente com os resultados anteriores que apontaram maior polarização nas avaliações de *agradabilidade* e *correção*.

Já entre os respondentes nascidos em São Paulo (SP), observa-se uma distribuição mais dispersa das escolhas, com destaque para o sotaque paulistano, mas também com percentuais relevantes atribuídos a outras capitais e à opção *nenhum*. Esse comportamento sugere uma identificação menos categórica com a variedade local, o que dialoga com a avaliação predominantemente indiferente do sotaque paulistano nas dimensões analisadas anteriormente. Além disso, esse padrão sugere uma identificação linguística menos concentrada em uma única variedade e pode ser relacionado ao caráter historicamente heterogêneo de São Paulo, marcado pela intensa mobilidade populacional e pelo convívio prolongado com sujeitos provenientes de diferentes regiões do país. A exposição contínua a múltiplos falares tende a relativizar fronteiras dialetais e a favorecer atitudes linguísticas mais flexíveis, nas quais o sotaque assume um papel menos fixo e mais negociável na construção das identidades. Como afirmou Oushiro, ao finalizar sua tese de doutorado, ao destacar a importância de sua pesquisa intitulada *Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo*,

> este trabalho atualiza uma agenda de pesquisas sobre o português da cidade de São Paulo, na qual também se incluem o contato linguístico com italianos, árabes, japoneses, bolivianos, peruanos e brasileiros de toda parte; inúmeras comunidades de práticas que convivem no espaço urbano; os padrões de variação dos falantes que residem nas dezenas de outros municípios que compõem a Região Metropolitana e transitam pela capital cotidianamente. A diversidade sociodemográfica de São Paulo ainda deixa muito a se explorar. (Oushiro, 2015, p. 327)

Ainda em relação a *São Paulo*, destaca-se um resultado particularmente significativo: entre os participantes nascidos no estado de São Paulo, não houve qualquer registro de escolha pelo sotaque do Rio de Janeiro como aquele com o qual gostariam de nascer. Esse resultado sugere a existência de fronteiras simbólicas bem demarcadas entre essas duas variedades urbanas, frequentemente colocadas em relação de comparação no imaginário social brasileiro. Do ponto de vista sociolinguístico, a ausência de identificação com o sotaque carioca por parte dos participantes paulistas pode ser

[7] PortVix é um banco de dados de entrevistas sociolinguísticas realizada com 46 participantes da comunidade de fala de Vitória/ES (Yacovenco *et al.*, 2012)

interpretada como um indício de forte ancoragem identitária ao próprio modo de falar, associado a valores de prestígio, legitimidade e centralidade socioeconômica. Além disso, esse dado pode refletir disputas simbólicas históricas entre São Paulo e Rio de Janeiro, nas quais o sotaque opera como marcador de pertencimento e distinção social. Assim, mais do que uma simples preferência linguística, a não escolha do sotaque carioca pelos paulistas aponta para processos de identificação negativa ou de distanciamento simbólico, nos quais o falar do outro é percebido como pouco desejável para a construção da própria identidade. Esse achado reforça a compreensão de que as atitudes linguísticas não são neutras, mas atravessadas por relações de poder, representações sociais e dinâmicas identitárias que orientam tanto a valorização quanto a rejeição de determinadas variedades linguísticas.

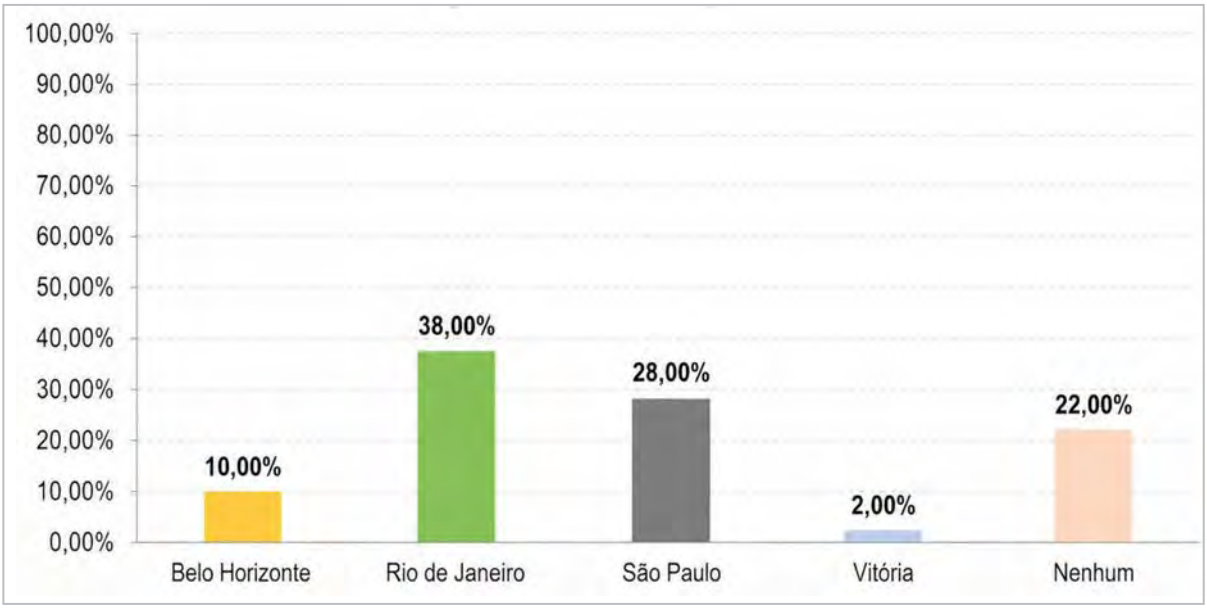
De modo geral, com base nos resultados apresentados no gráfico 6, percebe-se que a escolha hipotética do sotaque desejado está fortemente condicionada pelo local de nascimento, reforçando a centralidade do pertencimento regional na construção das atitudes linguísticas. Esse resultado confirma que avaliações sobre sotaques não refletem apenas julgamentos estéticos ou normativos, mas funcionam como marcadores simbólicos de identidade.

Stuart Hall, ao discutir sobre as culturas nacionais, em seu livro *A identidade cultural na pós-modernidade* afirma que

> no mundo moderno, as culturas nacionais em que nascemos se constituem em uma das principais fontes de identidade cultural. Ao nos definirmos, algumas vezes dizemos que somos ingleses ou galeses ou indianos ou jamaicanos. Obviamente, ao fazer isso estamos falando de forma metafórica. Essas identidades não estão literalmente impressas em nossos genes. Entretanto, nós efetivamente pensamos nelas como se fossem parte de nossa natureza essencial. (Hall, 2006, p. 47)

Esse sentimento de identidade e lealdade a sua comunidade de nascimento se deve, segundo ele, à ideia de que a nação não é apenas uma identidade política, mas algo que produz sentidos, um sistema de representação cultural, uma vez que as pessoas não são apenas cidadãs daquele espaço geográfico, mas se identificam com vários aspectos que formam essa representação cultural e entre elas está a variedade linguística típica daquela região.

Gráfico 7: Frequência relativa da escolha do sotaque com o qual os participantes NÃO gostariam de nascer entre as capitais da região Sudeste do Brasil



Fonte: Elaboração da autora

Os resultados relativos à pergunta *Qual sotaque desses estados você NÃO gostaria de ter?* evidenciam de forma ainda mais clara os processos de estigmatização linguística e de avaliação negativa associados a determinadas variedades regionais. Observa-se que o sotaque do *Rio de Janeiro* concentra o maior percentual de rejeição, seguido pelo sotaque de *São Paulo*, configurando um padrão que contrasta fortemente com os dados de escolha positiva analisados anteriormente.

A rejeição mais elevada do sotaque carioca pode ser interpretada como reflexo de sua forte hipervisibilidade social e midiática, que, embora contribua para sua difusão e reconhecimento, também potencializa a cristalização de estereótipos negativos. Esse resultado dialoga diretamente com as análises dos traços de *correção* (gráfico 3) e *agradabilidade* (gráfico 4), nas quais o falar carioca apresentou índices elevados de avaliações como *incorreto* e *irritante*, sugerindo uma percepção social marcada pela polarização.

O sotaque de *São Paulo*, por sua vez, ocupa a segunda posição entre os mais rejeitados, o que reforça a interpretação de que, apesar de sua associação à norma de prestígio institucional, o falar paulistano não mobiliza identificação afetiva positiva e pode ser percebido como pouco agradável ou excessivamente marcado por valores de ordem funcional, urbana e impessoal. Essa rejeição se articula com o alto índice de avaliações *indiferente* observado nos critérios anteriores, indicando distanciamento mais do que rejeição categórica, mas ainda assim suficiente para afastar sua escolha identitária.

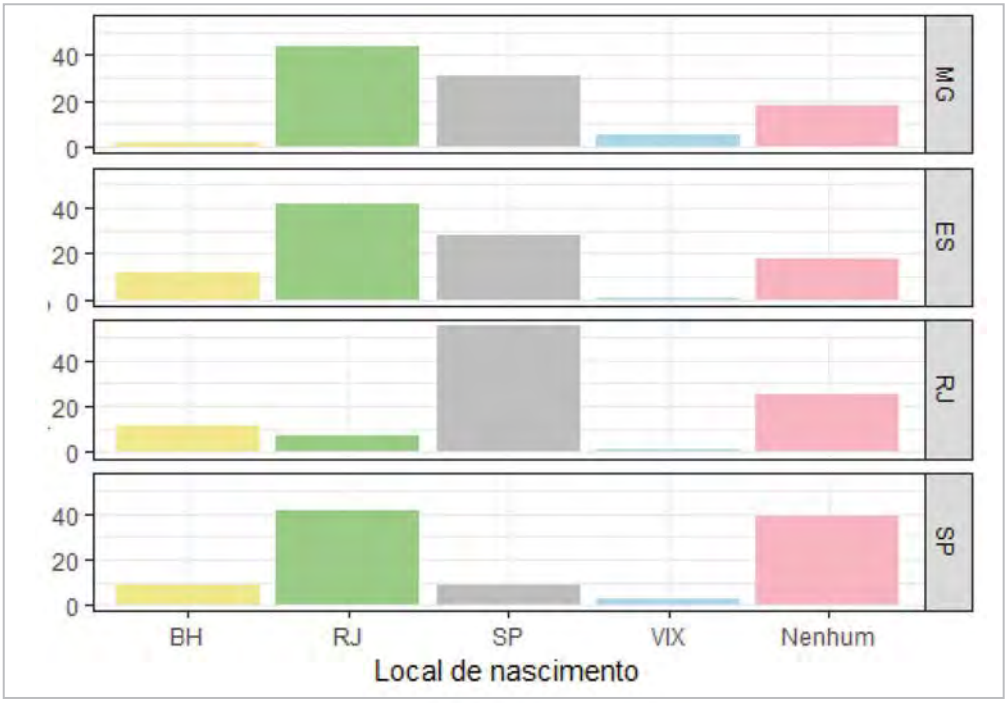
Em contraste, os sotaques de *Belo Horizonte* e, sobretudo, de *Vitória* apresentam índices muito baixos de rejeição. O resultado referente a *Vitória* é particularmente significativo, uma vez que o

sotaque capixaba praticamente não figura como opção rejeitada. Esse dado reforça a interpretação de que se trata de uma variedade percebida como pouco estigmatizada, socialmente mais neutra e amplamente aceitável, sobretudo quando se considera que a maior parte dos respondentes é composta por capixabas, o que favorece processos de lealdade linguística e autoavaliação positiva.

A opção *nenhum*, com percentual expressivo, indica novamente uma postura de recusa à lógica comparativa entre sotaques, sugerindo que parte dos participantes evita associar avaliações negativas a variedades regionais específicas. Esse comportamento pode ser interpretado como indício de maior consciência sociolinguística, resistência a julgamentos estigmatizantes ou valorização de uma identidade linguística própria.

Em conjunto, os dados mostram que as avaliações negativas são distribuídas de forma assimétrica entre as capitais analisadas e se articulam diretamente a fatores extralinguísticos, como exposição midiática, estereótipos regionais e pertencimento identitário. Esses resultados confirmam que atitudes de rejeição são tão reveladoras quanto as de valorização, funcionando como importantes indicadores das hierarquias simbólicas que atravessam a variação linguística no português brasileiro.

Gráfico 8: Distribuição da frequência relativa da escolha do sotaque com o qual os participantes NÃO gostariam de nascer entre as capitais da região Sudeste do Brasil e o local de nascimento do participante da pesquisa



Fonte: Elaboração da autora

Os dados relativos à pergunta *qual sotaque desses estados você NÃO gostaria de ter* revelam padrões claros de rejeição linguística, que se mostram fortemente condicionados tanto pela projeção social dos sotaques quanto pela origem dos respondentes. Considerando o resultado global, observa-se

que o sotaque do *Rio de Janeiro* concentra o maior percentual de rejeição, seguido pelo sotaque de *São Paulo*. Em contraste, o sotaque de *Vitória* apresenta índices residuais de rejeição, configurando-se como o menos recusado entre as capitais analisadas. A opção *nenhum*, por sua vez, aparece com percentual relevante, indicando que parte dos informantes evita explicitamente associar rejeição a uma variedade específica.

Esse padrão geral se torna mais elucidativo quando os dados são analisados a partir do local de nascimento dos respondentes. Entre os participantes nascidos em Minas Gerais (MG) e no Espírito Santo (ES), o sotaque do Rio de Janeiro emerge de forma consistente como o mais rejeitado, seguido pelo sotaque de São Paulo. Esse resultado dialoga com as análises anteriores, nas quais o falar carioca apresentou maior polarização avaliativa, especialmente nos traços de *agradabilidade* e *correção*, sugerindo que sua alta saliência social favorece tanto avaliações positivas quanto rejeições explícitas.

Entre os respondentes nascidos no Rio de Janeiro (RJ), observa-se um movimento inverso: o sotaque paulistano concentra o maior índice de rejeição, enquanto o sotaque carioca apresenta percentuais reduzidos. Esse dado evidencia um efeito claro de lealdade linguística, segundo o qual falantes tendem a proteger simbolicamente a própria variedade, deslocando a rejeição para falares associados a outras regiões percebidas como concorrentes no espaço sociocultural.

Já entre os respondentes nascidos em São Paulo (SP), o sotaque do Rio de Janeiro volta a ocupar a posição de maior rejeição, seguido pela opção *nenhum*. Esse comportamento sugere, por um lado, a presença de estereótipos negativos associados ao falar carioca e, por outro, uma possível resistência dos falantes paulistas em atribuir rejeição direta ao próprio sotaque, optando por uma estratégia avaliativa mais neutra.

De modo geral, os resultados confirmam que a rejeição de um sotaque opera como um índice particularmente sensível de fronteiras identitárias, nas quais pertencimento regional, rivalidades simbólicas e estereótipos socioculturais desempenham papel central. Assim, esses dados reforçam que atitudes linguísticas negativas não são aleatórias, mas sistematicamente organizadas em função de fatores sociais, históricos e identitários, evidenciando que a avaliação da variação linguística constitui, sobretudo, um processo socialmente situado.

5. Considerações finais

Este estudo teve como objetivo investigar as atitudes linguísticas em relação aos sotaques das capitais da região Sudeste do Brasil, considerando tanto o reconhecimento e a identificação dessas variedades quanto sua avaliação a partir dos critérios de *prestígio*, *beleza*, *correção* e *agradabilidade*. Os resultados evidenciam que as percepções dos participantes não se distribuem de forma homogênea entre os falares analisados, refletindo a complexa articulação entre fatores linguísticos, sociais, identitários e ideológicos.

No que diz respeito ao *prestígio*, os sotaques do Rio de Janeiro e de São Paulo se destacam como os mais fortemente associados a esse critério. Tal resultado sugere a centralidade dessas capitais no imaginário sociolinguístico nacional, em grande medida relacionada à sua projeção econômica,

política e midiática. Trata-se de um prestígio predominantemente aberto, vinculado à visibilidade institucional e à circulação ampliada dessas variedades, ainda que, como os dados demonstram, esse reconhecimento não implique avaliações uniformemente positivas nos demais critérios.

Quanto ao critério *beleza*, o sotaque de Belo Horizonte emerge como o mais bem avaliado, padrão que se mantém de forma consistente ao longo da análise. Essa valorização aponta para a existência de um prestígio afetivo, frequentemente associado ao falar mineiro, indexado a traços como suavidade, cordialidade e proximidade interpessoal. De modo convergente, o mesmo sotaque é também o mais bem avaliado em termos de *agradabilidade*, reforçando a hipótese de que avaliações de natureza estética e afetiva tendem a caminhar juntas, sendo particularmente sensíveis a estereótipos positivos socialmente compartilhados.

No que se refere à *correção*, o sotaque de Vitória apresenta os índices mais elevados de avaliação positiva. Esse resultado é especialmente relevante, uma vez que o falar capixaba costuma ser descrito como menos saliente no cenário nacional. Os dados sugerem que essa menor marcação social pode favorecer sua associação à norma linguística valorizada, produzindo julgamentos de adequação e correção, ainda que acompanhados, em alguns casos, por avaliações de indiferença. Tal padrão reforça a ideia de que a correção linguística percebida não se ancora exclusivamente em critérios estruturais, mas em processos sociais de avaliação da variação.

As análises relativas às escolhas hipotéticas - tanto do sotaque que os participantes gostariam de ter quanto daquele que não gostariam - evidenciam de forma contundente o papel do local de nascimento e da identidade regional na construção das atitudes linguísticas. A preferência pela variedade local e, inversamente, a rejeição mais frequente de sotaques associados a avaliações negativas confirmam a presença de fenômenos de lealdade linguística e de autoavaliação positiva. Ao mesmo tempo, a recorrência da opção *nenhum* indica posturas críticas ou distanciadas em relação aos estereótipos regionais, sugerindo que parte dos participantes resiste a classificações rígidas das variedades em termos de valor social.

Em conjunto, os resultados reforçam que as avaliações sobre sotaques constituem um terreno privilegiado para observar a projeção de valores sociais sobre a língua. Nesse sentido, os dados dialogam diretamente com o conceito de Respeito Linguístico, proposto por Scherre (2020), ao evidenciar que julgamentos negativos, estigmatizações ou hierarquizações entre variedades não têm base linguística, mas ligações a questões sociais e ideológicas. Reconhecer a legitimidade e a diversidade dos falares regionais implica promover uma postura de respeito às diferentes formas de falar, compreendendo a variação como constitutiva da língua e da identidade dos falantes.

Por fim, este artigo pretende contribuir para o campo da Sociolinguística ao oferecer evidências empíricas sobre como sotaques da região Sudeste do Brasil são percebidos e avaliados, reafirmando a necessidade de que essas discussões ultrapassem o âmbito acadêmico e alcancem contextos educacionais e sociais mais amplos, de modo a fomentar práticas linguísticas pautadas no respeito, na diversidade e na inclusão.

Referências

AGUILERA, Vanderci de Andrade. *Atlas linguísticos regionais: caminhos e perspectivas*. Londrina: EDUEL, 2007.

BRANDÃO, Silvia Figueiredo. *A geografia linguística no Brasil*. São Paulo: Contexto, 2005.

CAMPBELL-KIBLER, Kathryn. *Listener perceptions of sociolinguistic variables*: the case of (-ing). 2006. 282 f. Dissertation (Doctorate in Linguistics) – Department of Linguistics, Stanford University, Stanford, 2006.

CARDOSO, Suzana Alice Marcelino *et al*. *Atlas Linguístico do Brasil (ALiB).* Londrina: EDUEL, 2014.

CARDOSO, Suzana Alice Marcelino. *Geolinguística*: tradição e modernidade. Salvador: EDUFBA, 2010.

FREITAG, Raquel. *Variação linguística: diversidade e cotidiano*. São Paulo: Contexto, 2025.

HALL, Stuart. *A identidade cultural na pós-modernidade*. Rio de Janeiro: DP&A, 2006.

LABOV, W. *Padrões Sociolinguísticos*. São Paulo: Parábola Editorial, 2008 [1972].

LEITE, Yonne; CALLOU, Dinah. *Como falam os brasileiros*. Rio de Janeiro: Zahar, 2010.

MOLLICA; Maria Cecília; BRAGA, Maria Luiza. *Introdução à Sociolinguística*: o tratamento da variação. São Paulo: Contexto, 2017.

NARO, Anthony Julius. The social and structural dimensions of a syntactic change. *Language*, v. 57, pp. 63-98, 1981.

NARO, Anthony Julius; SCHERRE, Maria Marta Pereira. *Origens do Português Brasileiro*. São Paulo: Parábola Editorial, 2007.

NASCENTES, Antenor. *O linguajar carioca*. Rio de Janeiro: Organização Simões, 1953.

OLIVEIRA E SILVA, Giselle Machline de; SCHERRE, Maria Marta Pereira. *Padrões sociolinguísticos*: análise de fenômenos variáveis do português falado no Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996.

OUSHIRO, L. *Identidade na pluralidade*: avaliação, produção e percepção linguística na cidade de São Paulo. 2015. 394 f. Tese (Doutorado em Letras) – Universidade de São Paulo, Faculdade de Filosofia, Letras e Ciências Humanas, Programa de Pós-Graduação em Semiótica e Linguística Geral, São Paulo, 2015.

PRESTON, Dennis. Language with an attitude. *In*: CHAMBERS, J. K.; SCHILLING, N. (ed.). *The Handbook of Language Variation and Change*. 2. ed. Malden/MA: Wiley-Blackwell, 2013. pp. 157-182.

SCHERRE, Maria Marta Pereira. *Respeito linguístico*: contribuições da sociolinguística variacionista, 2020. Palestra apresentada no ABRALIN ao vivo. Disponível em:https://www.youtube.com/watch?v=W4XqhsiB9I0. Acesso em: 15 abr. 2021.

SCHERRE, Maria Marta Pereira; NARO, Anthony J. Duas dimensões do paralelismo formal na concordância de número no português popular do Brasil. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, São Paulo, v. 9, n. 1, pp. 1-14, 1993.

TESCH, Leila Maria. Sotaque capixaba: um estudo de percepção. *Revista (Con)Textos Linguísticos*, v. 16, n. 34, pp. 225-242, 2022.

VOTRE; Sebastião; RONCARATTI, Cláudia. *Anthony Julius Naro e a linguística no Brasil:* uma homenagem acadêmica. Rio de Janeiro: 7Letras, 2008.

YACOVENCO, Lilian Coutinho; SCHERRE, Maria Marta Pereira; TESCH, Leila Maria; BRAGANÇA, Marcela Langa Lacerda; EVANGELISTA, Elaine M.; MENDONÇA, Alexandre K. de; CALMON, Elba Nusa; CAMPOS JÚNIOR, Heitor da Silva; BARBOSA, Astrid F.; BASÍLIO, Jucilene O. S.; DEOCLÉCIO, Carlos Eduardo; SILVA, Janaína B. da.; BERBERT, Aline F.; BENFICA, Samine de A. Projeto PortVix: a fala de Vitória/ES em cena. *Alfa: Revista de Linguística*, v. 56, pp. 771-806, 2012.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72243

Do papel à consciência: a escrita de cartas como instrumento de reflexão sobre a diversidade linguística

From paper to consciousness: letter writing as a tool for reflection on linguistic diversity

Susana Menezes Araujo[1]

Ormezinda Maria Ribeiro[2]

RESUMO

A variação linguística é uma característica de todas as línguas naturais. Língua e sociedade estão intimamente relacionadas e exercem influência uma sobre a outra. No ensino da língua portuguesa, esse critério deve ser levado em consideração. Nesse sentido, o presente trabalho tem como objetivo conscientizar alunos do Ensino Médio sobre a legitimidade da variação linguística no Brasil e as implicações sociais da discriminação baseada na fala. O eixo teórico que norteia esta pesquisa é a Sociolinguística Educacional. A sua fundamentação é embasada nos trabalhos de Bagno, Stubbs e Gagné (2002), Bagno (2007, 2017), Baronas (2022), Bortoni-Ricardo (2004, 2005), Scherre (2005), Silva (2002), entre outros. A metodologia utilizada é de base qualitativa interpretativista e adota como estratégia a realização de oficinas para estudantes da 2ª e 3ª séries do Ensino Médio de uma escola pública localizada em Ceilândia – DF. A análise dos dados revelou que as discussões, bem como as reflexões realizadas nas oficinas contribuíram para a ampliação do conhecimento dos estudantes sobre a variação linguística, a heterogeneidade linguística e social no Brasil e sobre o preconceito linguístico e sua consequente discriminação, apontando para possíveis mudanças de posturas diante da diversidade linguística brasileira. Além disso, a análise das cartas evidenciou que os discentes foram capazes de adequar sua linguagem ao interlocutor, demonstrando maior consciência sobre a diversidade linguística e respeito às diferenças.

PALAVRAS-CHAVE: Variação linguística. Ensino. Língua portuguesa. Preconceito linguístico. Ensino médio.

ABSTRACT

Linguistic variation is a characteristic of all natural languages. Language and society are intimately related and exerting mutual influence on one another. In the teaching of the Portuguese language, this criterion must be taken into account. In this sense, the present study aims to raise awareness among high school students regarding the legitimacy of linguistic variation in Brazil and the social implications of speech-based discrimination. The theoretical framework guiding this research is Educational Sociolinguistics. Its foundation is based on the works of Bagno, Stubbs e Gagné (2002), Bagno (2007, 2017), Baronas (2022), Bortoni-Ricardo (2004, 2005), Scherre (2005), Silva (2002), among others. The methodology employed is qualitative-interpretative and adopts as a strategy the conduction of workshops for students in the 2nd and 3rd grades of high school at a public school located in Ceilândia – DF. Data analysis revealed that the discussions and reflections carried out in the workshops contributed to expanding the students' knowledge of linguistic variation, linguistic and social heterogeneity in Brazil, and linguistic prejudice and its consequent discrimination, pointing toward possible changes in attitude regarding Brazilian linguistic diversity. Furthermore, the analysis of the letters evidenced that the students were able to adapt their language to the interlocutor, demonstrating greater awareness of linguistic diversity and respect for differences.

[1] Universidade de Brasília (UnB), susanamenezes.a@gmail.com, https://orcid.org/0000-0002-6087-5832.

[2] Universidade de Brasília (UnB), aya.ribeiro@yahoo.com.br, https://orcid.org/0000-0002-5911-3005.

Recebido em 20/01/2026 - Aceito em 13/06/2026

KEYWORDS: Linguistic variation. Teaching. Portuguese language. Linguistic prejudice. High school.

1. Introdução

A variação linguística é um fenômeno próprio das línguas naturais e reflete a sociedade que as utiliza. Nenhuma língua natural está isenta da variação, uma vez que todas elas recebem influências de fatores que as fazem variar, sejam esses fatores de ordem linguística ou social.

Embora seja algo natural, a variação que as línguas apresentam nos usos feitos pelos interagentes é avaliada de formas distintas pelas sociedades. Ao falarmos, deixamos transparecer nossas origens geográfica, social e cultural, nossa escolarização, entre outros fatores. De acordo com a percepção que nossos interlocutores têm sobre nossos usos linguísticos, recebemos avaliações positivas ou negativas. Quando essas avaliações são negativas, somos segregados por sermos considerados indignos de fazermos parte de determinados grupos, isto é, somos discriminados pelo modo como falamos. Isso acontece, segundo Gnerre (1991, p. 6), porque "uma variedade linguística "vale" o que "valem" na sociedade os seus falantes".

Assim, a diversidade inerente à fala é alvo constante de estigmas, visto que a norma culta muitas vezes é utilizada como um mecanismo de segregação social em detrimento de sua função comunicativa. Nesse sentido, Scherre e Naro (2008) argumentam que a investigação da variação linguística transcende a descrição técnica de dados, exigindo um olhar crítico sobre as dinâmicas de poder e a construção da identidade dentro de contextos sociais marcados por profundas disparidades.

O desconhecimento de grande parte da população sobre a questão da variação linguística e seus condicionantes sociais alimenta práticas de discriminação que permeiam diversos espaços coletivos, sendo o ambiente escolar um dos palcos mais críticos dessas tensões. Diante desse cenário, esta pesquisa parte de um questionamento central: qual o papel da escola no trato com a diversidade linguística e de que maneira a Sociolinguística Educacional pode fundamentar práticas pedagógicas capazes de mitigar esse problema? Assim, torna-se imprescindível refletir sobre tal realidade, empregando estratégias que estimulem o estudante a usar a própria língua como suporte para essa reflexão, transcendendo, assim, ao ensino normativo tradicional, com vista a combater o preconceito linguístico no cotidiano escolar. Nessa perspectiva, este trabalho justifica-se pela necessidade de implementar intervenções que sensibilizem os estudantes, tendo como objetivo principal conscientizar alunos do Ensino Médio sobre a legitimidade da variação linguística no Brasil e as implicações sociais da discriminação baseada na fala.

Para alcançarmos esse objetivo, realizamos oficinas destinadas a estudantes da 2ª e 3ª séries do Ensino Médio de uma escola pública situada em Ceilândia – DF. A proposta partiu do reconhecimento de um cenário problemático: o desconhecimento generalizado sobre os mecanismos da variação linguística e a persistência de práticas de discriminação na sociedade. Como hipótese de intervenção, propusemos que, por meio da leitura de textos temáticos e da produção de cartas direcionadas a diferentes interlocutores, os participantes fossem levados a confrontar seus próprios juízos de valor

sobre a língua. Esperávamos, assim, que a reflexão crítica permitisse não apenas a aquisição de novos conhecimentos teóricos, mas uma mudança efetiva de comportamento diante da diversidade linguística brasileira.

Este trabalho está estruturado em quatro seções: a revisão da literatura, a metodologia adotada, a análise e discussão dos dados e as considerações finais.

2. A essência das línguas naturais

As línguas são um fenômeno dinâmico e estão em constante processo de adaptação que reflete a interação existente entre a língua e a sociedade. Elas são "uma **atividade social**, um *trabalho coletivo*, empreendido por todos os seus falantes, cada vez que eles se põem a interagir por meio da fala ou da escrita" (Bagno, 2007, p. 36, destaques do autor). Dado que são fruto de um trabalho coletivo, é natural que apresentem variações, uma vez que seus falantes, diversos que são, ao usarem-nas, transparecem suas diversidades. Dessa maneira, a variação linguística é uma característica de todas as línguas naturais, como a nossa língua materna, por exemplo.

Ao atentarmos para a constituição do povo brasileiro, logo notamos o quão diversificado ele é. Essa diversidade apresenta-se em diversos aspectos, como na aparência física, no modo de vestir, nas culturas, nas práticas de letramentos, no uso da língua, entre outros. Esse último aspecto interessa-nos de modo particular, visto que, no que tange ao uso da língua portuguesa, algumas ideologias estão presentes. Segundo Bagno (2017, p. 200, destaque do autor),

> ideologia linguística se refere ao conjunto de **atitudes** e crenças acerca da língua compartilhadas no interior de uma sociedade e que servem de esteio para certos valores culturais e sociais. As ideologias linguísticas frequentemente servem para racionalizar estruturas e relações sociais existentes, assim como hábitos linguísticos dominantes.

As ideologias linguísticas exercem forte influência no senso comum e, dessa maneira, contribuem substancialmente para os processos de padronização das línguas.

> O senso comum pode ser definido como um corpo de conhecimentos provenientes das experiências e das vivências que orientam o ser humano nas várias ações e situações de sua vida. Ele se constitui de opiniões, valores, crenças e modos de pensar, sentir, relacionar e agir. O senso comum se expressa na linguagem, nas atitudes e nas condutas e é a base do entendimento humano (Minayo, 2012, p. 622).

A língua portuguesa no Brasil apresenta diversas variedades. Essas variedades não são, entre si, melhores ou piores, são apenas diferentes e igualmente válidas no processo de interação, devendo o falante atentar para a adequação ou inadequação de determinado uso linguístico em decorrência da situação sociocomunicativa na qual estiver inserido. Para Haguette:

> A sociedade é constituída por elementos materiais e simbólicos que se entrelaçam e se consubstanciam em realidade concreta, segundo a percepção que deles tenha a população envolvida; ela é, pois, um misto de aspectos objetivos e subjetivos. Esta percepção, entretanto,

> não é homogênea, variando segundo a posição que os indivíduos e os grupos ocupam na estrutura social, ou seja, a partir de sua posição de classe (Haguette, 1987, p. 135).

Contudo, contrariamente à ideia de percepção defendida por Haguette, no que diz respeito à língua portuguesa do Brasil, curiosamente, ainda persiste em algumas esferas da sociedade brasileira, a visão de uma língua homogênea que deve ser, acima de tudo, venerada, respeitada e seguida por todos aqueles que ocupam ou pretendem ocupar uma posição de destaque na sociedade. Quando nos expressamos, por meio da língua, somos imediatamente incluídos em redes, que podem ser de prestígio ou de estigma. Assim, nas práticas sociais mediadas pela língua, as sociedades constituem-se de maneiras variadas e criam categorias positivas ou negativas em relação aos usos linguísticos feitos pela sociedade, legitimando-os ou rechaçando-os. Tal categorização estende-se aos usuários e, a depender das formas como usam a língua, podem ser considerados dignos ou indignos de fazerem parte de determinados grupos. Nesse sentido, "a linguagem torna-se terreno de disputas porque demarca pertencimento, identidades que são produzidas também pelos ouvintes, que categorizam e, muitas vezes, racializam falantes a partir de como falam, contribuindo para as desigualdades sociais" (Jung, 2022, p. 173).

As ideologias linguísticas compartilhadas pelas pessoas as impulsionam a comportarem-se de maneiras distintas diante da diversidade existente e, consequentemente, algumas atitudes são tomadas. "As atitudes são uma espécie de disposição para reagir favorável ou desfavoravelmente a uma situação dialógica e que pode influenciar comportamentos positivos ou negativos quanto à acomodação a um dialeto" (Silva; Gomes, 2020, p. 55). Nesse prisma, o desconhecimento científico das pessoas sobre as regras que determinam o uso da língua portuguesa por seus interactantes as faz produzirem julgamentos de valor sobre algumas formas de usar a língua e, como consequência, incidem atos de violência simbólica contra aqueles que não compartilham dos mesmos usos linguísticos que um grupo seleto da sociedade. Assim sendo, os estudiosos da língua exercem um papel de extrema relevância na disseminação de informações sobre o seu funcionamento. Segundo Scherre,

> Analisar uma língua em sua intimidade é um privilégio. Esmiuçar as entranhas das formas linguísticas e sentir a sistematicidade que envolve línguas, dialetos e variedades, sem julgamento de valor, é de beleza ímpar e só pode fazer bem aos que têm essa possibilidade. Partilhar esse bem constitui mais do que um dever, é uma responsabilidade social, é uma questão de cidadania (Scherre, 2005, p. 10).

Nessa perspectiva, além da importante contribuição dos estudiosos para a informação sobre a língua e a sua sistematicidade, julgamos fundamental o papel da escola no processo de esclarecimento e conscientização da realidade heterogênea brasileira como fator primordial para a existência da diversidade linguística, de modo a oportunizar aos sujeitos pertencentes à comunidade escolar a ampliação dos conhecimentos sobre sua língua materna, suas características, seus usos e funções, possibilitando-lhes oportunidades de usarem variedades linguísticas distintas dentro e fora da sala de aula. Dessa maneira,

> a tarefa educativa da escola, em relação à língua materna, é justamente criar condições para que o educando desenvolva sua competência comunicativa e possa usar, com segurança, os recursos comunicativos que forem necessários para desempenhar-se bem nos contextos sociais em que interage (Bortoni-Ricardo, 2004, p.78).

Se pretendemos suscitar, nas pessoas, mudanças de comportamentos quanto à percepção da heterogeneidade linguística brasileira, o trabalho de conscientização e valoração da diversidade linguística existente deve ser contínuo. Nesse processo, é essencial que as práticas escolares realizadas no sentido de informarem e conscientizarem sobre a variação linguística e, particularmente, sobre o preconceito linguístico, além de acontecerem ininterruptamente, transponham o espaço da sala de aula e envolvam todos os membros pertencentes à comunidade. Nesse contexto, a inserção da Sociolinguística no ambiente escolar torna-se algo imprescindível.

3. A Sociolinguística no ensino da língua portuguesa

Há um ramo da Sociolinguística que se preocupa com a variação da língua no âmbito escolar e com suas implicações no ensino da língua, entre as quais o preconceito linguístico: a Sociolinguística Educacional. Essa vertente foi cunhada pela sociolinguista brasileira Stella Maris Bortoni-Ricardo por meio de pesquisas desenvolvidas na área, e tem contribuído com algumas melhorias no ensino da língua materna nas escolas brasileiras, conforme aponta Bortoni-Ricardo:

> Muito embora os preconceitos linguísticos não tenham desaparecido, a sociolinguística forneceu munição teórica e tecnológica para combatê-los, bem como para que os sistemas escolares começassem a se preocupar com a adequação de seus métodos às peculiaridades linguísticas e culturais de seus alunos que não provinham das camadas dominantes da sociedade (Bortoni-Ricardo, 2005, p. 151).

Os estudiosos dessa área são favoráveis à pedagogia da variação linguística e acreditam que o ensino da língua será mais eficiente se ocorrer pelo viés do uso. Segundo suas concepções, as variedades linguísticas existentes devem ser abordadas no ensino da língua sem julgamentos de valor. Tal abordagem permitirá aos discentes um conhecimento mais consistente sobre a língua e resultará em um ensino linguístico mais significativo na vida deles, visto que os preparará melhor para atuarem nas práticas sociais mediadas pela língua.

> A principal influência dos estudos sociolinguísticos para a educação provém da ênfase veemente na premissa 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 linguisticamente ou culturalmente deficientes. Essa premissa representou uma verdadeira revolução na forma de encarar as variedades ou línguas minoritárias nas escolas (Bortoni-Ricardo, 2005, p. 151).

No entanto, os avanços proporcionados pelos construtos teóricos dessa ciência, no ambiente escolar, ainda não conseguem barrar o preconceito linguístico existente entre os membros das comunidades escolares. Segundo Baronas, as alterações almejadas ainda não se efetivaram em razão do

> distanciamento entre as instituições superiores de ensino e as instituições do ensino básico, a falta de condições de estudo para os professores do ensino básico e a falta de incentivo para os professores do ensino superior para se dedicarem a projetos extensionistas para uma formação continuada com os professores do ensino básico. Diante deste cenário, é fundamental buscar forças para uma efetiva integração entre as pesquisas científicas e a prática escolar (Baronas, 2022, p. 123).

A escola tem como uma de suas tarefas ensinar sobre respeito e formar cidadãos que atuem na sociedade de forma respeitosa. Quando não aceitamos nem respeitamos o outro por ser diferente de nós e, no caso específico do preconceito linguístico, quando não acolhemos alguém e o discriminamos por apresentar uma variedade linguística diferente da nossa, estamos sendo preconceituosos e, junto com o preconceito, vem a discriminação que fere, magoa e silencia. Comportamentos de intolerância a determinadas variedades linguísticas, no espaço escolar, trazem implicações sérias às vítimas do preconceito linguístico, dentre as quais: o medo, o sofrimento, a vergonha, o sentimento de inferioridade, a dificuldade de expressar-se, de fazer amizades e de estabelecer relações dentro e fora da sala de aula, o desinteresse pelas aulas ou pela escola, a resistência em interagir nas aulas, seja por espontânea vontade ou quando solicitadas pelo professor, e, em caso mais extremo, o preconceito sobre a forma própria de alguém falar pode levar à evasão escolar. Conforme Bortoni-Ricardo (2004, p. 35), "[...] precisamos tomar conhecimento da magnitude e dos efeitos nefastos do preconceito linguístico para podermos nos municiar de informações científicas e combatê-lo".

Dessa forma, no ensino da língua portuguesa, a relação intrínseca entre a língua e a sociedade deve ser levada em consideração. As práticas pedagógicas desenvolvidas nas escolas devem levar os estudantes a refletirem sobre a realidade heterogênea da língua e sobre seus comportamentos em relação à diversidade linguística no Brasil. Nesse processo, é papel do professor deixar claro para os discentes que existe mais de uma possibilidade de usar a língua e que os usos feitos por nós serão recebidos de formas distintas pela sociedade, pois convergem na língua valores sociais e culturais. Dessa forma,

> uma das tarefas do ensino de língua na escola seria, então, *discutir os valores sociais atribuídos a cada variante linguística*, enfatizando a carga de discriminação que pesa sobre determinados usos da língua, de modo a conscientizar o aluno de que sua produção linguística, oral ou escrita, estará sempre sujeita a uma avaliação social, positiva ou negativa (Bagno; Stubbs; Gagné, 2002, p. 75, destaques dos autores).

Diante do exposto, consideramos essencial que a abordagem da temática da variação linguística seja uma prática constante nas escolas brasileiras e que tal abordagem vise conscientizar os alunos sobre as variedades linguísticas existentes no intuito de alcançar mudanças de postura

dos membros que compõem o ambiente escolar. Dessa maneira, a parte prática deste trabalho foi executada com a intenção de colaborar para essa mudança desejada, conforme acompanharemos na seção metodológica a seguir.

4. Metodologia da pesquisa

A metodologia adotada neste trabalho é de base qualitativa interpretativista, pautada em métodos da etnografia empregada na educação para a geração dos dados e utiliza como estratégias: pesquisa de percepção a partir da observação participante; anotações de campo; execução de oficinas para estudantes da 2ª e 3ª séries do Ensino Médio de uma escola pública situada em Ceilândia – DF. As oficinas aconteceram durante os meses de março e abril de 2025, no período de 10/03 a 07/04. Ao todo, foram nove encontros, sendo dois por semana, nas segundas e quartas-feiras, das 13h às 14h30min.

O *corpus* deste estudo é constituído por cartas produzidas pelos estudantes durante as oficinas. No total, foram produzidas quarenta cartas. As categorias estabelecidas para a análise dos dados gerados foram: a adequação da linguagem ao interlocutor e o atendimento ao objetivo da escrita. Os dados foram analisados a partir de critérios interpretativos.

Oito estudantes participaram deste estudo. Cinco deles regularmente matriculados na 2ª série e três, na 3ª série. Eles têm idades entre 17 e 18 anos. A forma como os alunos foram identificados na presente pesquisa foi a seguinte: i. nome de um personagem da obra Capitães da Areia[3], ii. gênero (masculino ou feminino), iii. idade. Observem o quadro que segue:

Quadro 1: Identificação dos estudantes colaboradores do estudo

Pedro Bala (m) 18 anos	Sem-Pernas (f) 17 anos
Professor (m) 17 anos	Gata (f) 17 anos
Pirulito (m) 17 anos	Volta Seca (f) 17 anos
Dora (f) 17 anos	Boa-Vida (f) 18 anos

Fonte: Elaboração das autoras

4.1. Do processo dialógico das oficinas – tomada de consciência

As oficinas foram realizadas ao longo de cinco semanas, durante o contraturno vespertino, envolvendo estudantes da 2ª e 3ª séries do turno matutino. A opção por esse horário visou preservar o cronograma regular das disciplinas e garantir que o engajamento dos discentes ocorresse sem prejuízo às suas atividades acadêmicas habituais. O propósito central dessas atividades foi promover a conscientização sobre a variação linguística e o preconceito em nosso país, oferecendo subsídios para que os participantes pudessem confrontar e mitigar práticas de discriminação decorrentes da diversidade linguística, especialmente no contexto escolar.

[3] Romance escrito por Jorge Amado na década de 30.

No primeiro encontro, os estudantes responderam a um questionário com perguntas abertas, cuja finalidade era nos permitir conhecê-los, saber quais as suas expectativas em relação às oficinas, bem como averiguar os conhecimentos que eles tinham sobre a temática da variação linguística e do preconceito linguístico.

Nesse momento, os participantes foram orientados sobre a dinâmica das atividades, sendo informados de que a produção de cartas seria o eixo condutor das discussões propostas. Após a aplicação do questionário diagnóstico, a primeira etapa consistiu no estudo detalhado desse gênero textual e de suas características composicionais. A decisão de priorizar essa fundamentação técnica visou proporcionar aos estudantes o domínio necessário sobre as particularidades do gênero, a fim de que eles se sentissem seguros e familiarizados com a estrutura da carta antes de utilizá-la como ferramenta de reflexão nas etapas subsequentes.

Os estudantes demonstraram interesse e curiosidade pelo conteúdo. Quando solicitados, participavam e davam suas opiniões. Mostramos e discutimos com eles trechos da primeira carta escrita em terras brasileiras - a carta de Pero Vaz de Caminha, datada de 1º de maio de 1500. Eles ficaram bastante interessados e surpresos. Além disso, apresentamos e comentamos sobre os materiais em que as cartas eram produzidas, desde os papiros às telas de computadores, e como foram enviadas ao longo dos anos. Abordamos ainda os tipos de cartas, a estrutura delas e, mais especificamente, sobre a carta pessoal, mostrando exemplos desse tipo de carta.

No segundo encontro, os estudantes realizaram a primeira produção textual, cujo objetivo foi exercitar a escrita do gênero carta. Essa atividade inicial funcionou como uma etapa de familiarização prática, preparando-os para as produções que ocorreriam nas fases posteriores da pesquisa. Para subsidiar a escrita, foi disponibilizado o material didático da aula anterior, de modo que os alunos pudessem recorrer às orientações sobre o gênero sempre que necessário.

Ademais, colocamo-nos à disposição deles. Para a escrita, os discentes deveriam considerar o seguinte comando: *Com base no que foi estudado sobre o gênero textual carta, escolha uma pessoa para quem você gostaria de escrever e envie-lhe uma carta. O assunto a ser tratado na carta fica a seu critério. Mãos à obra!!!*

Após a produção deles, demos início à segunda aula que tinha como objetivo: apresentar o histórico da língua portuguesa no Brasil e a contribuição de povos distintos na sua formação. Na aula, tratamos da história da língua portuguesa no Brasil. Começamos apresentando os oito países em que a língua portuguesa é a língua oficial e fizemos um breve panorama sobre as outras línguas faladas aqui, além das duas línguas de comunicação entre os brasileiros: a língua portuguesa, que é a língua majoritária, e a língua brasileira de sinais, língua usada pelos surdos do nosso país. Na sequência, abordamos como ocorreu a formação da língua portuguesa no Brasil, destacando a contribuição de diversas línguas na composição do nosso léxico. Posteriormente, versamos sobre os sotaques brasileiros e como eles surgiram, bem como apresentamos alguns dos dialetos do Brasil. Para encerrar a aula, exibimos um vídeo disponível no Youtube intitulado “Sotaques brasileiros: como as pessoas

acham que são", de autoria de Manoloyy, datado de 29/09/2015[4]. As discussões realizadas nesse encontro, serviram para mostrar aos alunos a influência de diferentes povos para a variação que a nossa língua materna apresenta.

Os alunos participaram da aula, ficaram atentos, divertiram-se, mostraram-se surpresos em alguns momentos, em outros, curiosos.

A aula do terceiro encontro teve o objetivo de discutir a variação linguística e os fatores que a condicionam. Nele, discorremos sobre o que é variação linguística, sobre a noção de certo e errado *versus* adequado e inadequado e sobre os fatores condicionantes da variação linguística (nesse tópico, tratamos da variação nos distintos níveis da língua), sempre exemplificando com expressões presentes no cotidiano. Observamos que os alunos ficavam fascinados, constatando que a variação linguística vai além dos sotaques. Eles foram bastante receptivos. Participaram da aula, trouxeram exemplos do dia a dia deles, reconstruíram os textos usados nos exemplos.

No quarto encontro, aprofundamos o estudo sobre a variação linguística, com foco nos fatores extralinguísticos que a condicionam e na sistematização de suas modalidades (diacrônica, diatópica, diastrática, diamésica e diafásica). Assim como na etapa anterior, a abordagem priorizou o uso concreto da língua, examinando gêneros textuais diversos — como tirinhas, bilhetes, letras de música, entrevistas e notícias — para ilustrar a dinâmica da variação.

Durante as discussões, observamos uma reação de surpresa por parte dos estudantes ao constatarem como variáveis sociais como gênero, mercado de trabalho e redes sociais exercem influência direta nas escolhas linguísticas dos falantes. Essa percepção revelou-se um ponto de inflexão, pois os participantes ainda não haviam refletido sobre a língua sob essa perspectiva da dimensão social da linguagem. O diálogo constante permitiu que a turma participasse ativamente, engajando-se de forma lúdica e reflexiva com as novas descobertas sobre o funcionamento do idioma por eles usado.

Após tratarmos desses assuntos, apresentamos algumas questões de vestibulares, extraídas do blog intitulado "Veredas da Língua", de autoria do professor Maurício Fernandes da Cunha[5]. Os estudantes as respondiam e após darem suas respostas, nós as discutíamos.

Em uma conversa posterior a essas aulas, alguns alunos comentaram conosco que, com os exemplos, ficava muito mais fácil compreender o assunto. Esse comentário nos faz perceber o quão importante é ensinar a língua pelo viés do uso, como prima a Sociolinguística Educacional, promovendo reflexões sobre ela a partir de exemplos do uso real.

No quinto encontro, os alunos produziram a segunda carta. O comando da atividade propunha uma situação hipotética de grande relevância pessoal: *Parabéns!!! Seu e-mail foi premiado com uma viagem para a Disney, nos Estados Unidos, com tudo pago! A viagem acontecerá no período de 25/05 a 31/05 de 2025 e será patrocinada pela empresa Cei Turismo. Pedimos que entre em contato com a*

[4] https://www.youtube.com/watch?v=gJHc5cDmweU. Acesso em: 26 jul. 2026.

[5] Disponível em: https://veredasdalingua.blogspot.com. Acesso em: 26 jul. 2026.

equipe responsável pelo setor de viagens pelo e-mail ceiturismo@gmail.com e informe o número do seu bilhete premiado, que é nº 12.409, até o dia 10/05 de 2025 para receber o seu prêmio. Mais uma vez, nossas felicitações! Aguardamos sua resposta!

Beltrano Mendes de Albuquerque, Departamento de Notificação aos Ganhadores, Rua Getúlio Vargas, nº 736, Edifício Pingo de Ouro, 5º andar, sala 06, Centro, Belo Horizonte – MG. CEP: 12345-678

Que notícia maravilhosa! Você precisa compartilhá-la! Escreva para o diretor da sua escola, solicitando a autorização dele para se ausentar sem nenhum prejuízo em seus estudos.

O objetivo central dessa atividade foi estimular os estudantes a estabelecerem uma conexão prática entre a teoria da adequação linguística e o contexto de interação, exigindo deles o exercício do monitoramento estilístico em suas produções.

Foi observado um nítido engajamento da turma na elaboração das cartas, e um visível entusiasmo ao aplicar os conhecimentos em um cenário de interlocução mais formal. Durante o processo, atuamos como mediadoras, permanecendo à disposição para sanar dúvidas, enquanto os estudantes utilizavam, como suporte, os materiais didáticos e as sistematizações desenvolvidas nos encontros anteriores.

A aula ministrada no sexto encontro teve o objetivo de promover reflexões sobre o preconceito linguístico, mais especificamente no ambiente escolar, com vistas a mudanças de posturas dos estudantes.

Iniciamos o encontro apresentando a letra de uma música aos alunos. Tratou-se da música "Volta pra escola", lançada em 2014, pelo grupo de pagode Molejo. Eles não a conheciam, então, colocamos o áudio para que pudessem acompanhar a letra, ouvindo a música. Após a audição, indagamo-los sobre o enredo da música. Eles disseram que era possível perceber a variação linguística e, também, o preconceito linguístico presente na história e citaram trechos que comprovavam suas percepções.

Após esse primeiro momento de discussão, demos início à aula sobre o preconceito linguístico, como ele acontece e suas consequências. Em seguida, mostramos alguns relatos de alunos que sofreram preconceito linguístico no ambiente escolar. Os estudantes ficaram muito atentos aos relatos. Percebemos que eles ficaram perplexos e comovidos. Isso nos deixou felizes porque notamos, em seus olhares, reflexões sobre a gravidade e sobre as complicações que o preconceito linguístico (discriminação) pode trazer à vida das pessoas que o sofrem. A participação deles durante a aula foi muito boa. Após a aula, como de costume, entregamos o texto com o conteúdo discutido. Enquanto o entregávamos, ouvimos um aluno dizer ao outro o seguinte: "Não tem outro jeito, tem que conscientizar!" Ouvir essa fala nos trouxe alegria, pois pudemos observar que as discussões realizadas foram capazes de promover reflexões necessárias para a mudança de postura almejada com este trabalho. Ao aproximarmo-nos deles, esse mesmo aluno indagou-nos: "Não tem solução?" Respondemos que sim, que a solução é conscientizar! Os dois estudantes concordaram conosco.

> A escola deve direcionar sua atuação para formar cidadãos capazes de se relacionar com qualquer grupo social sem nenhum tipo de preconceito e que saibam utilizar os diferentes

> usos da língua, adequando-os ao seu contexto de uso, pois em nossas salas de aula não temos apenas alunos, mas, sujeitos sociais (Araújo, 2014, p. 33).

No sétimo encontro, os estudantes produziram a terceira carta, cuja instrução era: *Imagine que você e sua família se mudaram de cidade. Nessa cidade, as pessoas falam diferente da forma como você e seus familiares falam. Na sua nova escola, as pessoas não respeitam o seu modo de falar, elas riem de você, criticam, corrigem, imitam, fazem brincadeiras de que você não gosta. Escolha uma pessoa de quem você goste e confie e escreva a ela para contar a experiência que você está vivendo na escola e como você se sente!*

Essa atividade visou promover a reflexão dos estudantes sobre a gravidade da discriminação em razão da variação linguística, intentando que o exercício da empatia proporcionado por ela pudesse provocar neles o desejo por uma sociedade mais justa e igualitária, pois

> numa sociedade que quer ser verdadeiramente democrática, é preciso conhecer, descrever, denunciar e combater os componentes do senso comum que funcionam como repressores da liberdade individual e coletiva, como instrumentos de discriminação social, de humilhação, de opressão psicológica e até mesmo de violência física (Bagno, 2007, p. 81).

No oitavo encontro, os alunos produziram a quarta carta, tendo como base o seguinte enunciado: *A partir do que você estudou e aprendeu sobre variação linguística nas oficinas, escreva uma carta para os seus colegas de classe que não participaram delas para falar sobre o que você aprendeu e sobre a importância que esses aprendizados tiveram e terão na sua vida!*

A intenção dessa atividade era que os estudantes recapitulassem os conteúdos que havíamos tratado sobre variação linguística e preconceito linguístico, nas oficinas, de modo a colaborar para a fixação deles e, ainda, que usassem os conhecimentos adquiridos para conscientizarem os colegas sobre a temática.

No nono encontro, os alunos produziram a última carta. O comando dado para a produção deles foi o seguinte: *Querido estudante, chegamos ao término das nossas oficinas. Gostaria de saber a sua opinião sobre elas. Por favor, escreva-me para contar o que você achou: se gostou ou não gostou; do que mais gostou ou do que não gostou; se foi importante para você participar das oficinas e por quê. Como você pretende usar os conhecimentos adquiridos no seu dia a dia na escola e fora dela?*

Após as produções, agradecemos a eles por terem participado das oficinas, comentamos a importância de eles usarem os aprendizados adquiridos, ao longo das oficinas, nas suas vidas e de os usarem como uma ferramenta de combate ao preconceito linguístico, a partir da conscientização das pessoas nos diversos ambientes em que eles se encontrarem.

5. Análise e discussão dos dados

Cada estudante escreveu cinco cartas para destinatários distintos. À medida que discutíamos, nas oficinas, sobre a variação linguística e o preconceito linguístico, os estudantes escreviam suas cartas. Diante disso, a produção das cartas foi um momento em que eles puderam refletir sobre a temática em

questão dentro e fora do espaço escolar. As propostas de escrita das cartas os levaram a essa reflexão, visando naturalizar a variação linguística e, consequentemente, desconstruir o preconceito que leva à discriminação em razão da diversidade linguística existente em nosso país. Os critérios estabelecidos para a análise dos dados gerados foram: a adequação da linguagem ao interlocutor e o atendimento ao objetivo da escrita.

A linguagem que usamos, quando escrevemos uma carta, é reveladora dos relacionamentos sociais existentes entre os correspondentes. Assim, quando o relacionamento é de proximidade e de intimidade a linguagem utilizada é mais informal. Quando a relação é de pouca intimidade, o registro usado é mais formal. Outro fator que influencia a linguagem utilizada nas correspondências são os papéis sociais dos correspondentes. Dessa maneira, quando esses papéis estão em equilíbrio (amigos, irmãos, namorados etc.) é possível perceber uma linguagem mais informal. Quando os papéis sociais estão em desequilíbrio (pais e filhos, médicos e pacientes, patrões e empregados, alunos e professores etc.) a linguagem tende a ser mais cuidadosa ou formal. Isso acontece porque

> quando escrevemos uma carta, sabemos bem quem é esse amigo, de que coisas ele gosta ou não, o que ele sabe sobre nós, se devemos escrever a carta numa folha de caderno, ou em um papel mais sofisticado, se devemos caprichar na letra, se podemos usar gírias (e até palavrões) ou não, se nos despedimos com um abraço, um tchau, um beijo, ou simplesmente assinamos nosso nome. Enfim, todos esses aspectos – longe de serem meros detalhes na produção textual – interferem, condicionam e orientam o processo de escrita de um texto (Vieira; Costa Val, 2005, p. 29).

Nas quarenta cartas escritas pelos estudantes, os destinatários foram diversificados. Nas cartas 1 e 3 eles puderam escolher seus destinatários. Nas cartas 2, 4 e 5 os destinatários foram previamente determinados. Na primeira carta escrita por eles, os destinatários escolhidos foram: amigas, amigos, *crush*, mãe, Presidente da República e o proprietário de um estabelecimento onde se consertam celulares. Na terceira carta escrita, eles optaram por escrever para: amigas, amigos, irmão e avó. Na segunda carta, o destinatário foi o diretor da escola onde eles estudam. Na quarta carta, eles escreveram para os colegas de classe que não participaram das oficinas e na quinta carta a destinatária foi a pesquisadora.

Vale destacar que a configuração linguística das cartas é determinada, primordialmente, pela relação entre a finalidade do texto e o perfil do interlocutor. Então, a escolha de um modelo específico - seja uma carta pessoal, oficial, comercial ou de reclamação - não é aleatória; ela é motivada pela intenção comunicativa do remetente. Nesse aspecto, o tipo de carta selecionado estabelece o protocolo de linguagem e o nível de formalidade exigidos, sendo o destinatário o elemento que valida essas escolhas. Dessa forma, ao redigir uma carta oficial ou de solicitação, o estudante é levado a reconhecer que a natureza institucional do gênero impõe um monitoramento estilístico mais rigoroso, evidenciando que a variação linguística é, acima de tudo, uma questão de adequação ao contexto e às convenções sociocomunicativas.

Assim, quanto ao primeiro critério estabelecido para a análise dos dados, o exame das cartas demonstrou que os estudantes adequaram sua linguagem aos interlocutores. Com os interlocutores amigas, amigos, *crush*, irmão e colegas de classe, os escreventes têm uma relação pessoal de intimidade e os papéis sociais exercidos por eles estão em equilíbrio. Dessa forma, a linguagem utilizada pelos remetentes para dialogarem com esses destinatários foi informal, apresentando gírias, palavras de afeto, expressões de intimidade, brincadeiras, além de ter sido mais livre, visto que a relação estabelecida entre os correspondentes permite mais liberdade. Já com as destinatárias mãe e avó, embora haja um relacionamento de intimidade entre elas e a remetente, o papel social está em desequilíbrio. Dessa maneira, a linguagem utilizada no diálogo foi mais respeitosa. Conforme explica Silva (2002, p. 99-100),

> Quando o relacionamento entre os participantes se funda no âmbito da interpessoalidade e da intimidade, as relações interativas por eles construídas encarnam um caráter de simetria e informalidade. Essa condição confere aos participantes um status relativamente semelhante, na medida em que a sua posição social parece não ser um fator que pode provocar relações de poder ou desequilíbrio na interlocução, a despeito das relações hierarquizadas entre mãe e filha; tia e sobrinho, por exemplo.

Nesse sentido, os diálogos estabelecidos com aqueles que estão na mesma posição social que os estudantes, foram mais espontâneos. Já com as destinatárias que ocupam uma posição superior na hierarquia familiar, a interação ocorreu em um tom mais cerimonioso. Observem os exemplos:

Quadro 2: Linguagem nas cartas em que há relações de intimidade: simétrica e assimétrica

Carta	Participante	Destinatário	Trechos
Carta 1	Pirulito (m) 17 anos	Amigo	"**Oi amigo**! [...] André você viu que vai longar um filme do Minecraft? **bora** assistir **mano** não parece tão bom mais dá para dar umas boas risadas. **Para sempre melhores amigos**! **De seu grande amigo** Pirulito"
Carta 1	Gata (f) 17 anos	*Crush*	"**Meu querido, *crush***! [...] Acho que encontrei minha metade aquilo que faltava em mim com os mesmos propósitos e sonhos que eu, **carinhoso**, **fofo**, **lindo**, **maravilhoso**,... o que sinto por você é indescritível, você conseguiu! **Arrepiou até minha alma**. **Com muito amor**, **Gatinha**"
Carta 1	Dora (f) 17 anos	Mãe	"**Minha querida mãe**, [...] Infelizmente, ontem a noite, esses sentimentos veio à tona, e logo entrei em desespero e medo de tudo voltar, **e principalmente, atrapalhar a ti novamente**. Por isso decidi que irei começar ir a academia para tentar fugir dessa angústia e cuidar do meu corpo e minha mente. Assim como, a **senhora** sempre me aconselhou, a cuidar de mim. **Obrigado por tudo mãe**. Dora"

Carta 3	Professor (m) 17 anos	Irmão	"**Caro irmão**, [...] Como eu havia comentado antes, eles me criticam na fala e apesar de toda essa tensão definitivamente não me importo e muito menos me sinto ofendido, porém, é uma situação **paía**, uma situação **chata** pois **não largam do meu pé**. Atenciosamente. **Seu querido irmão** Professor"
Carta 3	Volta Seca (f) 17 anos	Amiga	"**Querida amiga Alice**! [...] sempre que abro a boca vem um **idividuo que não tem o que fazer tirar sarro de mim** por isso evito falar. Queria tanto voltar para casa e falar avontade sem me preocupar com **ladainha**, **piadas de mau gosto e tudo mais**, espero revela de novo, **até a próxima**! thau thau. **Com amor** Volta Seca"
Carta 3	Dora (f) 17 anos	Avó	"**Querida vovó**, Sou eu, sua neta Dora. Como a **senhora** está? Espero que bem. [...] Porque para minha tristeza, não consegui me adaptar a essa escola. Já que descobri o preconceito das pessoas com o nosso modo de falar. Aparentemente, as pessoas acham feio ou até engraçado. **Com carinho e amor**, Dora."
Carta 4	Sem-Pernas (f) 17 anos	Alunos do 2º ano	"[...] teria sido bom vocês terem feito esse curso, principalmente **pelo fato de vocês não respeitarem a linguagem de muitas pessoas da sala**, [...]. É isso, **espero ter deixado claro que são desrespeitoso**, **obrigada**, **de nada**. **Com muito amor**, Sem-Pernas"
Carta 4	Pedro Bala (m) 18 anos	Alunos do 3º ano	"**Prezados colegas de classe**, [...] Enfim, estou aprendendo muito, vocês **vacilaram** em não participar, agregariam muito às suas vidas. **Agradeço a atenção de vocês**. Pedro Bala"
Carta 4	Professor (m) 17 anos	Alunos do 2º ano	"**Caros colegas**. [...] é uma ótima profissional. Com ela aprendi diversas coisas, e diversos jeitos de fazer uma carta, seja pessoal ou informal seja pessoal ou até publica, **por preguiça e determinação, quer dizer, falta dela, vocês perderam esta incrivel oficina de cartas**, espero que nas próximas oportunidades vocês pensem melhor. **Atenciosamente**. Professor"

Fonte: Elaboração das autoras

Com os destinatários Presidente da República e proprietário do estabelecimento onde se consertam celulares, os remetentes não mantêm nenhum tipo de relacionamento pessoal. Dessa maneira, a linguagem usada nas cartas remetidas a esses destinatários foi mais cerimoniosa e formal. Ademais, a posição social representada pelo cargo de Presidente da República pressupõe uma relação interativa de assimetria, exigindo do escrevente maior formalidade na escrita. Vejam os exemplos:

Quadro 3: Linguagem nas cartas em que não há nenhum tipo de relação pessoal

Carta	Participante	Destinatário	Trechos
Carta 1	Pedro Bala (m) 18 anos	Presidente da República	"**Prezada Vossa Excelência**, [...] **Dessa forma**, a requisição que faço é em relação a criação de um instituto, **do qual eu serei o representante** responsável por sua administração. **Nesse contexto**, **tratarei** de garantir a execução de medidas **que visem** a diminuição da queima de áreas florestais **por meio do monitoramento** das áreas verdes do território nacional e da adoção de protocolos **de contenção de incêndios**. [...] **Desde já agradeço a colaboração**. Pedro Bala"
Carta 1	Professor (m) 17 anos	Proprietário de estabelecimento	"**Olá prezado**, Como está?! Sinto uma imensa saudade do aparelho eletronico que eu deixei para o concerto, já faz algumas semanas e continuo sem nenhuma noticia **até o momento**. **Apesar disso, gostaria de ser informado da situação do meu aparelho**, ou então até mesmo ser atualizado sobre alguma informação, pois está me fazendo uma grande falta até o momento. [...] **Atenciosamente**, Professor"

Fonte: Elaboração das autoras

Já com os destinatários diretor da escola e pesquisadora os alunos mantêm uma relação de pouca ou nenhuma intimidade e há uma hierarquia entre eles que causa desequilíbrio nos papéis sociais exercidos. Sendo assim, para escreverem para esses destinatários, os estudantes usaram uma linguagem mais cuidadosa e respeitosa. Acompanhem os exemplos a seguir:

Quadro 4: Linguagem nas cartas em que há relação de pouca intimidade e assimetria

Carta	Participante	Destinatário	Trechos
Carta 2	Dora (f) 17 anos	Diretor da escola	"**Prezado Diretor Raul**[6], [...] Portanto, **lhe** peço que pense bem sobre, pois é o justo a se fazer. Sendo que posso recuperar tudo depois e acompanhar normalmente as aulas. E se o **senhor** autorizar, será justo e tranquilizador para mim fazer a viagem. **Agradeço** por ler esta carta, e espero que o **senhor** me compreenda. Por Dora"
Carta 2	Volta Seca (f) 17 anos	Diretor da escola	"**Encarecido, Diretor** [...] Portanto, peço **encarecidamente** que **considere meu pedido e que não haja nenhum problema ou imprevisto mediante a isso** aguardo sua resposta Até logo **Atenciosamente**: Volta Seca"

[6] Nome de um personagem da obra Capitães da Areia.

Carta 5	Boa-Vida (f) 18 anos	Pesquisadora	"**Querida professora**, A **senhora** é uma mulher muito sabia, possui muitos conhecimentos e explica muito bem, é uma expiração. Muito obrigada **por nos oferecer** o curso de Variação Linguística! [...] **Agradecida**, Boa-Vida"
Carta 5	Volta Seca (f) 17 anos	Pesquisadora	"Estou escrevendo essa carta para dizer que **estou deveras grata** pelas aulas até esse momento, cada aula e explicação **foram de grande ajuda e entendimento** sobre as variações linguisticas [...]. **Até logo**!! **Atenciosamente** Volta Seca"

Fonte: Elaboração das autoras

Encerrada a primeira parte da análise dos dados, passemos à análise do segundo critério elencado, qual seja: atendimento ao objetivo da escrita.

A produção da primeira carta tinha como propósito possibilitar aos estudantes o exercício da escrita do gênero em questão, de modo que pudessem familiarizar-se com ele e não sentissem dificuldades nas produções que realizariam ao longo das oficinas. Nesse sentido, a orientação para a escrita deu a eles a liberdade de escolher para quem e sobre o quê escrever, como se observa na instrução que segue: "*Com base no que foi estudado sobre o gênero textual carta, escolha uma pessoa para quem você gostaria de escrever e envie-lhe uma carta. O assunto a ser tratado na carta fica a seu critério. Mãos à obra!!!*"

Os destinatários, bem como os assuntos tratados por eles em suas cartas foram diversos. Entre os destinatários tivemos: amigas, proprietário de estabelecimento de conserto de celulares, *crush*, Presidente da República, amigos e mãe. Entre os assuntos tratados, observamos que eles escreveram para tecer elogios, fazer requerimento, desabafar, tratar de assuntos aleatórios, entre outros.

> Em geral, escrevemos cartas — de amor, de amizade, de agradecimento — movidos pela empatia que temos com aquele(a) para quem escrevemos. Queremos falar de nossa vida passada, presente e futura. A mão vai escrevendo, muitas vezes sem muito controle sobre o que escreve: os pensamentos e o desejo são mais fortes. Deixar que a mão deslize sobre o papel ou sobre o teclado é deixar fluir pensamentos, sentimentos (Maciel; Cafiero; Rangel, 2021, p. 12).

Dessa maneira, como os estudantes tiveram a liberdade de escolher seus destinatários e os assuntos a serem tratados em suas cartas e cumpriram com os seus propósitos, entendemos que o objetivo dessa primeira escrita foi atendido, conforme o quadro a seguir:

Quadro 5: Fragmentos da primeira carta

Participante	Trechos
Gata (f) 17 anos	"Meu querido, *crush*! Como você está? Início esta carta com imenso amor te conhecer foi a melhor coisa que me ocorreu, quando penso em você meu coração acelera, para e me pergunto, que sentimento bom, é tão bom amar alguém, tão bom amar você... [...] Com muito amor, Gata"
Boa-Vida (f) 18 anos	"Minha melhor amiga, [...] Muito obrigada pelos seus sábios conselhos, apoio, carinho e admiração, você é muito incrível! Espero poder te encontrar sábado para passarmos um tempo juntas. Podemos assistir dorama, tomar sorvete, correr, fazer várias coisas. Gosto muito da sua companhia[...] Para sempre melhores amigas, com muito carinho, Boa-Vida, para, Gata"
Pedro Bala (m) 18 anos	"Prezada Vossa Excelência, Venho por meio desta carta fazer um requerimento, pois, minha cidade vêm passando por algumas dificuldades, o ar se encontra muito poluído, a umidade esta extremamente baixa e a temperatura absurdamente quente. [...] Desde já agradeço a colaboração. Pedro Bala"
Dora (f) 17 anos	"Minha querida mãe, Como a senhora está? Tenho saudades das nossas conversas mais profundas. Inicio essa carta com um assunto que preciso desabafar. Portanto, preciso contar isso à alguém, e a senhora é a melhor pessoa para isso. [...] Obrigado por tudo mãe. Dora."
Sem-Pernas (f) 17 anos	"Caro Renan Como você está? Vim-lhe contar de uns acontecimentos que não lhe contei. Atualmente estou com 17 anos, ano que vem faço 18 anos quase terminando ensino o médio, estou pensando em fazer alguns cursos de informatica, e talvez, depois que eu concluir o ensino médio. [...] Com amor, Sua baixinha! Sem-Pernas"

Fonte: Elaboração das autoras

A segunda produção realizada por eles tinha como destinatário o diretor da escola. A proposta de escrita foi a seguinte: *Parabéns!!! Seu e-mail foi premiado com uma viagem para a Disney, nos Estados Unidos, com tudo pago! A viagem acontecerá no período de 25/05 a 31/05 de 2025 e será patrocinada pela empresa Cei Turismo. Pedimos que entre em contato com a equipe responsável pelo setor de viagens pelo e-mail ceiturismo@gmail.com e informe o número do seu bilhete premiado, que é nº 12.409, até o dia 10/05 de 2025 para receber o seu prêmio. Mais uma vez, nossas felicitações! Aguardamos sua resposta!*

Beltrano Mendes de Albuquerque, Departamento de Notificação aos Ganhadores, Rua Getúlio Vargas, nº 736, Edifício Pingo de Ouro, 5º andar, sala 06, Centro, Belo Horizonte – MG. CEP: 12345-678

Que notícia maravilhosa! Você precisa compartilhá-la! Escreva para o diretor da sua escola, solicitando a autorização dele para se ausentar sem nenhum prejuízo em seus estudos.

Como se percebe, o objetivo da escrita era dirigir-se ao diretor da escola para solicitar sua autorização para uma viagem que os discentes precisariam realizar. No entanto, em suas cartas, alguns alunos limitaram-se a informar que se ausentariam da escola no período da viagem, em vez de solicitarem a autorização do diretor para a sua realização. É certo que comunicar a ausência é parte do propósito da escrita, mas não constitui seu objetivo principal. Em virtude disso, fica evidente que tais estudantes não cumpriram, em sua completude, com o objetivo proposto. Observem os fragmentos que seguem:

Quadro 6: Fragmentos da segunda carta

Participante	Trechos
Gata (f) 17 anos	"Prezado diretor, [...] fui premiada com uma viagem para a Disney, nos Estados Unidos, com tudo pago! E peço-lhe humildemente que **me autorize a ir**. Para não prejudicar os meus estudos, estudarei *online*. Ademais, receberei um atestado de viagem. Isso ocorrerá no período de 25/05 a 31/05 de 2025. [...] Com carinho, da sua querida aluna, Gata"
Professor (m) 17 anos	"Prezado Diretor. Informo que no dia 24 de março recebi um comunicado diretamente da empresa CEI turismo, informando que fui sorteado para ir em uma viagem para a Disney no período de 25/05 a 31/05 de 2025, **estou informando para não ser afetado durante o período da minha ausência** e estou justificando o motivo [...] Atenciosamente, Professor"

Fonte: Elaboração das autoras

A terceira proposta de escrita proporcionou aos estudantes reflexões sobre o preconceito linguístico. A partir de uma situação hipotética de preconceito linguístico sofrido por eles, solicitamos-lhes que escolhessem uma pessoa da sua confiança para desabafar sobre os preconceitos sofridos, conforme a seguinte instrução: *Imagine que você e sua família se mudaram de cidade. Nessa cidade, as pessoas falam diferente da forma como você e seus familiares falam. Na sua nova escola, as pessoas não respeitam o seu modo de falar, elas riem de você, criticam, corrigem, imitam, fazem brincadeiras de que você não gosta. Escolha uma pessoa de quem você goste e confie e escreva a ela para contar a experiência que você está vivendo na escola e como você se sente!*

Conforme o enunciado, a proposta era que os alunos desabafassem com alguém sobre a experiência de preconceito linguístico vivida por eles, focando no ambiente escolar. A ênfase de tal preconceito nesse espaço deve-se ao fato de ser ele o nosso campo de pesquisa.

Entre os destinatários escolhidos pelos discentes para desabafarem estão amigos(as), irmão e avó. Ao analisarmos suas cartas, observamos que todos narraram e desabafaram sobre episódios de preconceito linguístico vivenciados por eles; todavia alguns não se concentraram no âmbito escolar, como solicitado na questão. Assim sendo, podemos afirmar que, novamente, parte dos alunos não cumpriu com o objetivo da escrita em sua totalidade. Verifiquem os trechos abaixo:

Quadro 7: Fragmentos da terceira carta

Participante	Trechos
Volta Seca (f) 17 anos	"Querida amiga Alice! [...] A nova casa é bonita, o bairro é calmo e devez em quando barulhento, mas dá para suporta. Já a escola mão posso dizer o mesmo, nunca poderia imaginar como as pessoas poderiam ser tão maldosas, se você me visse agora juraria que é outra pessoa, aqui eu fico mais quieta e quase nunca falo devido ao bulling que venho sofrendo por conta do meu sotaque, sempre que abro a boca vem um idividuo que não tem o que fazer tirar sarro de mim por isso evito falar. Com amor Volta Seca"
Boa-Vida (f) 18 anos	"Minha querida *best friend*, [...] Na escola, estou sofrendo preconceito por causa da minha forma de falar, as pessoas da minha turma ficam rindo de mim, me imitam, me criticam, me corrigem, fazem brincadeiras, das quais eu não gosto. Eu me sinto constrangida, envergonhada e triste com tudo isso. Isso não me faz bem. Assinado Boa-Vida para: Gata"
Gata (f) 17 anos	"Terceiramente foi quando meu grupinhos de amigos começaram a me zoar, pelo fato do meu sotaque ser distinto do deles. Confesso que tudo isso me deixou constrangida, arrependo-me de vir morar aqui. Logo mudarei, por isso deixo o aviso. Assinado: Gata Para: Raí"
Pedro Bala (m) 18 anos	"Além disso, as pessoas daqui desconhecem muitas das nossas palavras, outro dia fui comprar pão francês na padaria, fui atendido por uma moça que parecia vir do sul, quando pedi os pães tive que apontar qual deles era quando mostrei para ela, ela logo reclamou: 'se queria um cassetinho dizia logo, onde já se viu pão da França no Brasil?'". agradeço seu apoio, Seu camarada Pedro Bala."

Fonte: Elaboração das autoras

Para além da proposta, esclarecemos que o intuito maior dessa atividade era que os alunos, ainda que de forma hipotética, pudessem exercitar a empatia para, a partir desse exercício, colaborarem para uma sociedade mais ética e acolhedora, como intenta esta pesquisa.

A quarta carta escrita pelos estudantes colaboradores da pesquisa foi endereçada aos colegas de classe que não quiseram ou não puderam participar das oficinas. Nela, os discentes deveriam abordar os aprendizados adquiridos sobre a temática da variação linguística e ressaltar a importância desses conhecimentos em suas vidas, tanto no presente quanto no futuro. Acompanhem o enunciado da

atividade: *A partir do que você estudou e aprendeu sobre variação linguística nas oficinas, escreva uma carta para os seus colegas de classe que não participaram delas para falar sobre o que você aprendeu e sobre a importância que esses aprendizados tiveram e terão na sua vida!*

A intenção dessa atividade proposta era que os alunos fizessem um panorama a respeito do que foi discutido ao longo das oficinas acerca da temática em questão. Ou seja, que eles explicassem aos colegas o que é variação linguística, quais são os fatores que a condicionam, citassem exemplos e que abordassem o preconceito e a discriminação linguística sofridos por algumas pessoas em decorrência da variação linguística. Dessa forma, a atividade funcionaria como uma revisão do conteúdo estudado e, consequentemente, contribuiria para uma melhor fixação dos conceitos trabalhados.

A análise das oito produções dos alunos evidenciou que, com exceção do participante Professor, que focou somente na temática das cartas, todos os demais estudantes fizeram menção à variação linguística em seus textos. Alguns limitaram-se a afirmar que a variação linguística existe e comentaram sobre a importância de conhecê-la, sem, contudo, explicá-la. No entanto, embora não tenham explicado com clareza o que é variação linguística e porque ela ocorre, pode-se perceber nos textos desses alunos a ideia de que a língua apresenta variações em decorrência das situações de uso e do espaço geográfico, bem como pode-se perceber a ideia do que é o preconceito linguístico ao expressarem que algumas pessoas não são respeitadas ou recebem tratamentos distintos pelas formas como usam a língua.

Outros estudantes trataram do tema com maior propriedade, a partir de conceituações, da abordagem de fatores sociais que ocasionam a variação e de exemplificações. Contudo, nenhum deles mencionou sobre a importância que os aprendizados adquiridos tiveram ou terão nas suas vidas. Assim sendo, fica evidente que o objetivo da escrita não foi completamente alcançado. Observem os exemplos:

Quadro 8: Fragmentos da quarta carta

Participante	Trechos
Volta Seca (f) 17 anos	"[...] que é variação linguística, um assunto muito importante para o conhecimento geral sobre os sotaque de diferentes regiões do Brasil, onde se mosta as variedade linguísticas as semelhanças de fala e ao mesmo tempo a diferença gritante da mesma, como o modo de falar afeta diretamente a vida de alguém e como as pessoas recebem ou tratam outra apenas pelo modo como ela diz certas palavras."
Dora (f) 17 anos	"[...] Seria uma ótima oportunidade de aprendizado a vocês. Pois, é muito importante para a construção de uma sociedade melhor, mais comunicativa e respeitosa. E a oficina de variação linguistica me ajudou a entender melhor os diferentes tipos de linguagens, e como isso ocorre na sociedade. É um assunto que deveria ser abordado a todos, para o melhor entendimento das diversas formas como elementos de uma língua podem se apresentar em uma situação de uso, consequentemente, evitando o preconceito linguistico."
Sem-Pernas (f) 17 anos	"[...] honestamente eu entendi um pouco, sobre as variações linguisticas, preconceito linguístico, teria sido bom vocês terem feito esse curso, principalmente pelo fato de vocês não respeitarem a linguagem de muitas pessoas da sala, foi uma experiência agradável [...]"

Boa-Vida (f) 18 anos	"As variações linguísticas são diversas formas de como elementos de uma língua podem se apresentar em uma situação de uso. O uso da língua variam entre grupos sociais, entre regiões, devido a fatores que são próprios da língua. No curso de variação linguística, aprendi que que existe várias formas de falar em diferentes locais e devemos respeitar."
Pirulito (m) 17 anos	"Meus amigos vim aqui lhes informar sobre um tema pouco abordado nas escolas, a variação linguistica, ela fala sobre as diversas formas de comunicação que variam devido ao gênero, idade, local, classe social, entre outros, ela está presente em tudo que tenha comunicação, entretanto é urgente falar do preconceito que pode ocorrer o preconceito linguistico e causado por diversos fatores e quando feito pode causar diversos danos ao proximo como a tristeza, Fobia Social, em casos mais graves evasão escolar, o preconceito começa quando se julga o modo falar do outro pelas diferenças sonoras ou verbais [...]"
Pedro Bala (m) 18 anos	"Isso ocorre porque a língua sofre variações dependendo da idade, localidade, gênero, profissão ou realidade socioeconômica dos seus falantes. Além disso, a forma como o som de determinadas palavras é pronunciado afeta o modo como algumas palavras são escritas, bicicleta por exemplo pode ser escrita por alguns como "bicicreta", pois sua escolaridade influência também o modo como a língua será usada por essas pessoas."

Fonte: Elaboração das autoras

A quinta produção realizada pelos colaboradores tinha como destinatária a pesquisadora. A proposta da escrita era avaliar as oficinas realizadas, além de relatar se foi importante participar delas e indicar de que forma pretendem usar os conhecimentos adquiridos dentro e fora da escola. Observem o enunciado da questão: *Querido estudante, chegamos ao término das nossas oficinas. Gostaria de saber a sua opinião sobre elas. Por favor, escreva-me para contar o que você achou: se gostou ou não gostou; do que mais gostou ou do que não gostou; se foi importante para você participar das oficinas e por quê. Como você pretende usar os conhecimentos adquiridos no seu dia a dia na escola e fora dela?*

Os textos analisados mostram que os alunos gostaram das oficinas; contudo a maioria não mencionou a importância de participar delas, e somente dois deles versaram sobre como pretendem usar os conhecimentos adquiridos em suas vidas. Isso evidencia que, mais uma vez, o objetivo da escrita não foi atingido em sua completude por alguns alunos. É possível que a quantidade de questionamentos presentes no enunciado tenha dificultado a produção textual dos estudantes. Acompanhem os exemplos:

Quadro 9: Fragmentos da quinta carta

Participante	Trechos
Boa-Vida (f) 18 anos	"Gostei muito do curso e sei que sentirei falta dessa rotina de grande aprendizado. Nesse curso aprendi muitas coisas, tais qual, **o nome que se dar ao modo diferente que as pessoas falam chama-se dialeto**. [...] Aprendemos a fazer cartas, dessa forma pude treinar mais minha escrita. Com esse curso, **pude conhecer mais um pouquinho a língua portuguesa**."
Volta Seca (f) 17 anos	"Em geral, não tenho o que reclamar, pois **agora tenho conhecimento da existencia do bullyn que está presente na maneira que se fala** e mesmo com poucas horas de duração das aulas, foi possível transmitir muitos novos conhecimentos."
Dora (f) 17 anos	"Eu tenho de relatar que estou feliz com o curso, foi agradável, e **me ajudou a clarear a mente sobre variação linguística**. Bem como, colaborou para que eu treinasse minha escrita no gênero carta, pois é um tipo textual do PAS 2, no qual irei fazer esse ano. [...] Do mesmo modo que, **a reflexão sobre preconceito linguístico foi essencial para minha percepção do preconceito no ambiente escolar (e fora da escola)**."
Pirulito (m) 17 anos	"[...] nesta oficina eu aprendi o quão inportante e a comunicação e suas variações, **me ensinou a ter empatia e parar de julgar ao proximo me ensinou diversas culturas** e como foi o desenvolvimento da comunicação desde ponbos correios a conputadores super tecnológicos [...]"
Sem-Pernas (f) 17 anos	"Foi importante para mim por causa do diploma, pode ser bom para mim no futuro. Eu pretendo usar para redações, concursos, aulas normais em sala de aula, e **a respeita linguagens diferentes da minha** [...]"
Pedro Bala (m) 18 anos	"[...] Não só isso, o material utilizado foi de altíssima qualidade, permitiu o aprendizado de forma mais dinâmica e menos massiva. Além disso, **após aprender sobre variação linguística e preconceito linguístico serei capaz de agir de maneira mais ética, julgando menos a fala dos outros e aceitando a enorme variedade de dialetos que existem dentro da nossa língua**."

Fonte: Elaboração das autoras

Ficamos felizes em saber que os alunos gostaram das oficinas e de termos contribuído, de alguma forma, para o desenvolvimento pessoal e intelectual deles. É perceptível em suas cartas que as discussões e reflexões realizadas durante os nove encontros que tivemos proporcionaram a eles mais consciência sobre a variação linguística e, sobretudo, sobre a discriminação que existe em decorrência de tal variação. É possível observar nos trechos em destaque dos exemplos citados que as reflexões realizadas nas oficinas influenciaram as suas percepções sobre a heterogeneidade da nossa língua materna. Além de terem contribuído para a aquisição de novas percepções, as reflexões possibilitaram a avaliação dos estudantes sobre seus comportamentos diante da diversidade linguística brasileira, apontando para possíveis mudanças de comportamento.

Essas reflexões, ao impulsionarem uma revisão de comportamentos e atitudes, alinham-se à premissa de que o ensino de língua não deve ser uma prática isolada, mas um exercício de cidadania. Nessa perspectiva, a transformação observada nos estudantes revela o que Bagno define como a missão central do educador contemporâneo:

> À professora e ao professor de língua portuguesa cabe o trabalho da **reeducação sociolinguística** de seus alunos e de suas alunas. O que significa isso? Significa valer-se do espaço e do tempo escolares para formar cidadãs e cidadãos conscientes da complexidade da dinâmica social, conscientes das múltiplas escalas de valores que empregamos a todo momento em nossas relações com as outras pessoas por meio da linguagem (Bagno, 2007, p. 82, destaques do autor).

Portanto, ao promover a avaliação das condutas diante da diversidade linguística, a prática pedagógica, desenvolvida por meio da leitura e discussão de textos e da escrita de cartas com interlocutores e objetivos distintos, cumpre seu papel político e social, ao substituir o preconceito por uma consciência crítica capaz de reconhecer a legitimidade das variadas vozes que compõem a identidade brasileira.

Considerações finais

A sociedade brasileira é heterogênea e a língua, enquanto parte constitutiva da sociedade é, igualmente, heterogênea. O desconhecimento científico sobre essa realidade reforça ideologias linguísticas, entre as quais a ideia de que há variedades linguísticas mais bonitas, mais corretas e, portanto, superiores a outras. Tais percepções contribuem para a discriminação linguística e a exclusão social em diversos ambientes, até mesmo no escolar. Práticas de discriminação não condizem com um espaço de educação. Nesse sentido, é necessário informar sobre a realidade linguística brasileira e refletir sobre comportamentos de intolerância linguística, nesse ambiente. Assim sendo, este trabalho teve o objetivo de conscientizar alunos do Ensino Médio sobre a legitimidade da variação linguística no Brasil e as implicações sociais da discriminação baseada na fala.

Quanto ao primeiro critério estabelecido para a análise dos dados, constatamos que os estudantes adequaram a linguagem aos seus interlocutores. Quando interagiram com pessoas com as quais têm intimidade e que estão em uma posição social paralela à deles (amigos, amigas, *crush*, irmão) os estudantes optaram por uma linguagem mais informal. A linguagem que uma das participantes utilizou para dialogar com pessoas com as quais possui vínculo próximo, mas que a posição social é de assimetria (mãe e avó) foi mais cortês. Quando trataram com pessoas com as quais não se relacionam, os alunos usaram um registro formal, como nas cartas escritas ao Presidente da República e ao dono de um estabelecimento onde se consertam celulares. Quando escreveram para destinatários com os quais há uma relação de pouca ou nenhuma intimidade e desequilíbrio dos papéis sociais (diretor da escola e pesquisadora), a linguagem utilizada foi formal.

No que diz respeito ao segundo critério, verificamos que o objetivo da escrita foi alcançado apenas na primeira carta - na qual os estudantes escolheram para quem escrever e sobre o que tratar. Nas demais cartas, os objetivos não foram atendidos em sua completude por todos os estudantes.

No que tange ao objetivo pretendido neste trabalho, as observações realizadas ao longo das oficinas e a análise dos dados gerados pelos alunos permitiram-nos perceber que as discussões embasadas nas teorias da Sociolinguística Educacional bem como as reflexões realizadas contribuíram para a ampliação das concepções dos educandos sobre a variação linguística, a heterogeneidade linguística no Brasil e o preconceito linguístico. Além disso, notamos que as reflexões sobre o preconceito linguístico e suas consequências promoveram efeitos positivos nos estudantes, uma vez que verificamos, em suas produções, indícios de que novas posturas relacionadas à diversidade linguística serão adotadas.

Referências

ARAÚJO, Maria José Oliveira. *A variação linguística em sala de aula:* uma proposta de intervenção reflexiva sobre o preconceito linguístico. 2014. 70 f. Dissertação (Mestrado em Linguística e Ensino) – UFPB/CCHLA, João Pessoa, 2014.

BAGNO, Marcos; STUBBS, Michael; GAGNÉ, Gilles. *Língua materna*: letramento, variação e ensino. São Paulo: Parábola, 2002

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

BAGNO, Marcos. *Dicionário crítico de sociolinguística*. São Paulo: Parábola, 2017.

BARONAS, Joyce Elaine de Almeida. Nos caminhos da variação linguística. *In*: BRANDÃO-SILVA, Flávio; ROMUALDO, Edson Carlos; PEREIRA, Hélcius Batista. (org.) *Da Variação Linguística à "Pedagogia da Variação"*: descrição e ensino de português. São Carlos: Pedro & João Editores, 2022, pp. 123-141.

BORTONI-RICARDO, Stella Maris. *Educação em língua materna*: a sociolinguística na sala de aula. São Paulo: Parábola, 2004.

BORTONI-RICARDO, Stella Maris. *Nós cheguemu na escola, e agora?* Sociolinguística e educação. 2. ed. São Paulo: Parábola, 2005.

GNERRE, Maurizio. *Linguagem, escrita e poder*. 3. ed. São Paulo: Martins Fontes, 1991.

HAGUETTE, Teresa Maria Frota. *Metodologias qualitativas na sociologia*. Petrópolis: Vozes, 1987.

JUNG, Neiva Maria. Variação linguística no ensino: trocas linguísticas e simbólicas. *In*: BRANDÃO-SILVA, Flávio; ROMUALDO, Edson Carlos; PEREIRA, Hélcius Batista. (org.) *Da Variação Linguística à "Pedagogia da Variação"*: descrição e ensino de português. São Carlos: Pedro & João Editores, 2022, pp. 169-190.

MACIEL, Francisca; CAFIERO, Delaine; RANGEL, Egon de Oliveira. (org.). *Cartas para Magda.* 1. ed. São Paulo: Parábola, 2021.

MINAYO, Maria Cecília de Souza. Análise qualitativa: teoria, passos e fidedignidade. *Ciência & Saúde Coletiva.* v. 17, n. 3, pp. 621-626, 2012.

SCHERRE, Maria Marta Pereira. *Doa-se lindos filhotes de poodle*: variação linguística, mídia e preconceito. São Paulo: Parábola, 2005.

SCHERRE, Maria Marta Pereira; NARO, Anthony Julius. *Padrões sociolinguísticos*: análise de fenômenos variáveis. São Paulo: Parábola, 2008.

SILVA, Jane Quintiliano Guimarães. *Um estudo sobre o gênero carta pessoal*: das práticas comunicativas aos indícios de interatividade na escrita dos textos. 2002. 209 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Letras – Estudos Linguísticos -, Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, 2002.

SILVA, Mikaylson Rocha da; GOMES, Almir Anacleto de Araújo. O papel das atitudes linguísticas nos estudos variacionistas e de contato dialetal no pb. *Cuadernos de La Alfal*. v. 12 n. 1, pp. 53-70, maio 2020.

VIEIRA, Martha Lourenço; COSTA VAL, Maria da Graça. *Produção de textos escritos*: construção de espaços de interlocução / caderno do professor. Belo Horizonte: Ceale/FaE/UFMG, 2005.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72255

"Se você acha que capixaba não tem sotaque, fala arroz aí rapidinho": a ditongação na variedade capixaba[1]

"If you think capixabas don't have an accent, say 'arroz' real quick": diphthongization in the capixaba accent

Ana Carolina Carnieli Duarte[2]

Raquel Gomes Chaves[3]

RESUMO

Este trabalho investigou o processo de ditongação com glide /j/ em sílabas tônicas terminadas em /S/ em dados de fala da região Metropolitana da Grande Vitória (RMGV), a partir da perspectiva teórico-metodológica da Teoria da Variação e Mudança Linguística (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972], 1994) e com o suporte da fonética acústica. Como *corpus* de análise, utilizamos a Amostra Duarte (2025), composta por sete informantes estratificados por *sexo* (homens e mulheres), *faixa etária* (Faixa A: 22 a 35 anos e Faixa C: acima de 60 anos) e *nível de escolarização* (ensino médio completo e ensino superior completo). As variáveis controladas foram: *vogal-núcleo do segmento-alvo*, *posição do acento*, *estilo*, *sexo*, *faixa etária* e *escolarização* (variáveis de efeito fixo) e *item lexical* e *informante* (variáveis de efeito aleatório). Os resultados indicaram que o fenômeno apresentou taxa de aplicação de 35,5%, índice próximo ao registrado pela literatura para os dados de fala carioca. As variáveis linguísticas e extralinguísticas estatisticamente significativas foram *faixa etária*, a ditongação é mais frequente na fala dos informantes mais jovens; *posição do acento*, a ditongação é prevalente em monossílabos tônicos e oxítonas; *vogal núcleo do segmento-alvo*, a vogal média-alta anterior colabora com o fenômeno em oposição à vogal média-baixa posterior; e *estilo*, a leitura desfavorece a incidência de ditongação. Os dados sugerem, portanto, que a ditongação na RMGV é um processo influenciado pelo monitoramento estilístico, pela posição do acento na palavra, pela qualidade da vogal do item lexical e pela faixa etária dos falantes.

PALAVRAS-CHAVE: Ditongação. Variedade capixaba. Variação fonética-fonológica.

ABSTRACT

This study investigated the process of diphthongization with the glide /j/ in stressed syllables ending in /S/ in speech data from the Metropolitan Region of Vitória (RMGV), based on the theoretical-methodological framework of Variationist Theory (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972], 1994) and supported by acoustic phonetics. The *corpus* consists of the Duarte Sample (2025), comprising seven informants stratified by *sex* (men and women), *age group* (Group A: 22 to 35 years old; Group C: over 60 years old), and *education level* (high school and university graduates). The controlled variables included:

---

[1] Na etapa de redação deste manuscrito, utilizamos a ferramenta ChatGPT (OpenAI, 2026, versão GPT-4) com o objetivo de aprimorar a fluidez do texto. Após o uso da ferramenta, revisamos criticamente todo o conteúdo, assumindo integral responsabilidade pelo manuscrito publicado, incluindo eventuais imprecisões ou incorreções.

[2] Universidade Federal do Espírito Santo (UFES), Edital Piic 2024/2025 (Bolsa FAP), carnieli.carol@gmail.com, https://orcid.org/0009-0003-1835-9505.

[3] Universidade Federal do Espírito Santo (UFES), raquel.chaves@ufes.br, https://orcid.org/0000-0002-6310-0194.

Recebido em 20/01/2026 - Aceito em 10/06/2026

target segment nucleus vowel, stress position, style, sex, age group, and education (fixed effect variables), as well as *lexical item* and *informant* (random effect variables). Results indicated a 35.5% application rate, close to the rates reported in the literature for the Rio de Janeiro dialect. Statistically significant variables included: *age group*, with diphthongization being more frequent among younger speakers; *stress position*, prevalent in stressed monosyllables and oxytone words; *nucleus vowel*, where the high-mid front vowel favors the phenomenon compared to the low-mid back vowel; and *style*, as reading discourages diphthongization. The data suggest that diphthongization in the RMGV is a process influenced by stylistic monitoring, word stress position, the quality of the lexical item's vowel, and the speakers' age group.
KEYWORDS: Diphthongization. Capixaba variety. Phonetic-phonological variation.

1. Introdução

A discussão mais recorrente acerca da variedade capixaba gira em torno de quais seriam os traços linguísticos indexadores do falar espírito-santense. A dificuldade em determinar esses aspectos resulta na concepção geral de que o capixaba não teria sotaque (Tesch, 2022; Lacerda; Leppaus, 2023; Yacovenco, 2012). Segundo Yacovenco *et al.* (2012), o isolamento histórico do Espírito Santo e os fluxos migratórios posteriores contribuíram para uma variedade linguística pouco marcada, cujos traços passam despercebidos, inclusive pelos próprios falantes.

Nesse cenário, a ditongação variável em sílabas tônicas travadas por /S/ configura-se como um dos poucos traços salientes na identificação da variedade em questão. Tesch (2022), ao aplicar um questionário de percepção a 1400 respondentes, relativo às marcas identificadas como específicas do sotaque capixaba, apontou que alguns traços fonéticos/fonológicos foram mais recorrentes nas respostas, a saber: a realização do /S/ capixaba, identificado como diferente do /S/ carioca, e o fenômeno da ditongação (*arroz ~ arroiz*). Para os habitantes locais, a impressão de que a ditongação seria um traço identitário do capixaba também pode ser ilustrada por comentários, como os encontrados na rede social *X* (antigo *Twitter*): "Capixaba é uma raça engraçada. Acrescentam o i em todas as palavras: *arroiz, treis*; aí, quando a palavra tem *i*, ignoram e falam *pexe, fejão*." e "Se você acha que capixaba não tem sotaque, fala *arroz* aí rapidinho. Você fala *arroiz*.".

Partindo dessa discussão, neste trabalho, investigamos a variação da ditongação em sílabas tônicas travadas por /S/ (*três ~ treis, mas ~ mais, Jesus ~ Jesuis*) em dados de espírito-santenses que residem na Região Metropolitana da Grande Vitória (RMGV). Nos pautamos na perspectiva teórico-metodológica da Teoria da Variação e Mudança Linguística (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972], 1994), com o suporte da Fonética Experimental.

O processo variável referido, do ponto de vista fonológico, é caracterizado pela inserção de uma semivogal na mesma sílaba que o monotongo, criando um ditongo decrescente, isto é, uma sequência de dois segmentos vocálicos consecutivos (na qual o primeiro é uma vogal e o segundo, uma semivogal) ocupando uma mesma sílaba (Bisol, 1999). Do ponto de vista fonético, os ditongos decrescentes são abordados como eventos articulatórios gradientes que envolvem uma configuração vocálica estável, a qual se altera para uma segunda configuração mais fechada. Acusticamente, essa trajetória se manifesta nas transições formânticas, principalmente do primeiro e segundo formantes (Barbosa; Madureira, 2015).

As perguntas de pesquisa que guiam esta análise são as seguintes: (i) qual a frequência da ditongação variável em sílabas tônicas travadas por /S/ na Grande Vitória?; (ii) há evidências acústicas que corroboram a análise de oitiva na identificação das formas ditongadas *versus* monotongadas?; (iii) quais fatores linguísticos e extralinguísticos se correlacionam com a variante ditongada?

Adicionalmente, com base nos estudos revisados sobre o fenômeno na Região Sudeste (Silva, 2014; Aragão, 2014; Silveira, 2023; Callou, 2009), formulamos as seguintes hipóteses: (i) o índice de ditongação é alto na comunidade investigada e difundido entre as categorias sociais, com base na identificação desse traço pelos falantes (marca dialetal da comunidade como um todo); (ii) há diferenças significativas entre ditongos e monotongos, levando-se em conta que ditongos apresentariam transições de F1 e F2; (iii) os fatores favorecedores do fenômeno são: as vogais anteriores e a central, em comparação às posteriores; os monossílabos tônicos e oxítonas em relação às paroxítonas e proparoxítonas; o estilo de fala registra maior ocorrência do fenômeno do que o estilo leitura; os informantes mais jovens ditongam mais, assim como os menos escolarizados; mulheres tendem a ditongar mais do que homens.

No contexto capixaba não há registro de uma descrição abrangente e detalhada da atuação do fenômeno de ditongação, à exceção do trabalho de Silva (2014), que analisa a ditongação nas capitais brasileiras. Com vistas a preencher essa lacuna, este trabalho tem como objetivo geral investigar o processo de ditongação com glide /j/ em sílabas tônicas na variedade da Grande Vitória. Os objetivos específicos são: (i) compreender a influência dos fatores linguísticos e sociais nesse aspecto da fala; (ii) contribuir para a descrição fonético-fonológica da fala capixaba; e (iii) enquadrar essa discussão no quadro mais amplo da Teoria da Variação e Mudança Linguística (Weinreich; Labov; Herzog, 1968; Labov, 2008 [1972], 1994).

O presente trabalho organiza-se da seguinte forma: na seção 2, apresentamos a revisão de literatura acerca do fenômeno investigado, com foco em trabalhos sobre a Região Sudeste; na seção 3, descrevemos a metodologia empregada, com ênfase nos procedimentos de constituição da Amostra Duarte (2025), no tratamento dos dados, bem como na inspeção acústica dos segmentos analisados e na delimitação do envelope de variação; na seção 4, apresentamos e discutimos os resultados obtidos, contemplando tanto a análise quantitativa descritiva quanto a análise estatística inferencial, com base em modelos de regressão de efeitos mistos; e, por fim, na seção 5, tecemos as considerações finais, nas quais retomamos os principais achados do estudo e apontamos suas contribuições para a descrição sociolinguística da variedade capixaba.

2. Estudos sobre o fenômeno variável da ditongação na Região Sudeste

A ditongação por inserção de glide /j/ em sílabas tônicas terminadas em /S/ é um fenômeno recorrente em diversas variedades do PB, com ocorrências registradas em diferentes regiões do país. Silva (2014, p. 144), ao mapear o comportamento do fenômeno em 25 capitais brasileiras, por meio dos questionários lexicais que constituem o *corpus* do Projeto Atlas Linguístico do Brasil (ALiB),

conclui que a distribuição geográfica constitui fator importante para a aplicação do fenômeno, sendo o país dividido entre áreas fortemente favorecedoras da ditongação, como o Nordeste, e áreas de manutenção do monotongo, como o Sul.

Especificamente sobre a Região Sudeste, de particular interesse para a presente pesquisa, Silva (2014, p. 245) identifica um comportamento homogêneo nas frequências de ditongação entre as capitais: Rio de Janeiro apresentou índices de 19%, Vitória, frequência de 15%, Belo Horizonte, 12% e São Paulo registrou 13%. As taxas inferiores a 20%, no entanto, podem ser justificadas pela seleção de itens lexicais constantes no protocolo de coleta do Projeto ALIB, uma vez que Silva (2014) analisa as respostas ao Questionário Fonético-Fonológico (QFF), o qual é composto por 159 questões previamente elaboradas e organizadas e que incentiva respostas específicas dos participantes (com perguntas como: "a carne do porco não é magra porque tem…?", ao que o informante deveria responder "colesterol").[4] O questionário, porém, foi desenvolvido objetivando o estudo de diversos fenômenos, e não especificamente a ditongação diante de /S/ e, por isso, os contextos são, muitas vezes, pouco favorecedores. Além disso, a autora não analisou apenas a ditongação em sílabas tônicas, como foi feito no presente trabalho e em demais pesquisas sobre o fenômeno, considerando todos os contextos em que o /S/ ocupasse a posição de coda silábica. É importante destacar ainda que a análise da distribuição da ditongação em faixas etárias revelou indícios de mudança em curso, com favorecimento para a realização entre os mais jovens.

Outro estudo sobre a ditongação nas capitais brasileiras é o de Aragão (2014), que utiliza do mesmo *corpus* de Silva (2014). A autora também analisa itens lexicais isolados do QFF para testar se os processos da ditongação e da monotongação constituem variantes diatópicas, diastráticas ou puramente fonéticas (compreendendo a influência das variáveis linguísticas controladas, sendo elas: *tipo da vogal, contexto posterior, tonicidade, extensão da palavra* e *registro*). Os resultados globais apontaram para um alto percentual de ditongação no Brasil, com 72% de aplicação. No que tange à distribuição geográfica, incluindo as capitais da Região Sudeste, Aragão (2014, p. 12) nega a hipótese de que esses processos funcionem como marcadores regionais. A autora argumenta que o fenômeno ocorre em todo o território nacional, não sendo possível isolar o sudeste ou qualquer outra região como área exclusiva ou preferencial de ocorrência com base nos itens analisados. No entanto, apesar dos dados encontrados no estudo apontarem para esse resultado (já que a ditongação acontece em diversas regiões do país), se faz necessário compreender se o fenômeno ocorre da mesma maneira em todas as áreas do Brasil, identificando a influência de outros aspectos dos sotaques (como a realização do /S/), as trajetórias formânticas dos dados e a avaliação da população sobre esse aspecto de fala.

---

[4] As palavras analisadas por Silva (2014) a partir das respostas dos informantes ao QFF foram as seguintes: alastrou, aliás, arroz, as, asfalto, aspas, atrás, através, vasculhante, bastante, besta, bosque, busca, capataz, capaz, casca, casas, caspa, chaminés, churrasqueira, colegas, colesterol, coroas, costas, das, desvio, dez, dias, duas, elas, encosta, escola, esquerdo, fósforo, frutas, gasta, gosto, inglês, japonês, juntas, labaredas, lâmpadas, mas, mazela, mesma, muitas, nas, nos, os, palafitazinhas, pescoço, pesquisa, placas, português, prateleiras, produz, quiosques, rapaz, rasgando, rasgar, respostas, rodas, ruas, rústicas, talvez, três, vocês.

Os estudos que analisaram o comportamento variável da ditongação na fala carioca (Callou, 2009; Silveira, 2023) apontam essa variedade como de alta manifestação do fenômeno. No que se refere à frequência de aplicação da variante ditongada, Callou (2009, p. 47), ao analisar comparativamente dados do Projeto da Norma Urbana Linguística Culta (NURC) coletados em dois períodos (1970 e 1990), identifica taxas de ditongação entre 41% e 65% (na década de 1970) e de 19% a 62% (na década de 1990), a depender do contexto linguístico.

Silveira (2023, p. 30), por sua vez, ao desenvolver uma análise em tempo real comparando dados do Programa de Estudos sobre o Uso da Língua (PEUL) e contrapondo os resultados das décadas de 1980 e 2000, registra frequência de ditongação de 57% em monossílabos na década de 1980 e de 72% na de 2000. Ambos os estudos apontam para um condicionamento linguístico semelhante: a ditongação é verificada em maior escala quando a vogal núcleo é baixa ou média anterior, co-ocorre com a variante palatal em coda silábica e é fortemente condicionada pela tonicidade da sílaba, com os monossílabos apresentando frequência significativamente superior.

Quanto às variáveis extralinguísticas, embora o fenômeno seja analisado exclusivamente em dados de fala culta (o que indica que a ditongação não se trata de um marcador de baixa escolarização), observam-se padrões de variação relacionados à faixa etária, com distribuição curvilínea que sugere um processo de mudança linguística em curso, e ao sexo, com leve preferência pela ditongação entre falantes do sexo feminino.

3. Metodologia

Para a realização deste estudo, optamos pela constituição de uma amostra de dados de fala coletados segundo os moldes labovianos (Amostra Duarte, 2025). A referida amostra é composta por sete entrevistas sociolinguísticas realizadas com informantes residentes da Região Metropolitana da Grande Vitória[5], que, segundo dados do IBGE (2022), possui aproximadamente 1.880.000 habitantes distribuídos entre as sete cidades: Cariacica, Fundão, Guarapari, Serra, Viana, Vila Velha e a capital Vitória. Os sete informantes que fazem parte deste estudo residem em Cariacica, Serra, Vila Velha e Vitória. Na imagem 1, apresentamos um mapa da RMGV.

[5] A amostra encontra-se, ainda, em fase de constituição, sendo as sete entrevistas analisadas as disponíveis para estudo até o presente momento.

Imagem 1: Mapa da Região Metropolitana da Grande Vitória



Fonte: Instituto de Defesa Agropecuária e Florestal do Espírito Santo (IDAF-ES). Disponível em: https://idaf.es.gov.br/mapas-do-espirito-santo. Acesso em: 05 jan. 2026

Cabe destacar que não delimitamos como pré-requisito que os pais dos participantes da pesquisa fossem, também, nascidos na localidade, como se observa na constituição de outras amostras sociolinguísticas. No entanto, analisamos os dados de fala de cada um dos participantes, observando se atendiam aos traços fonéticos capixabas, nesse caso, considerando o /S/ alveolar e o /R/ aspirado como variantes características da comunidade em estudo[6] – o que foi aferido a partir da observação pessoal da entrevistadora no contato prévio com os participantes. Os informantes foram estratificados por *sexo* (homens e mulheres), *faixa etária* (*Faixa A*, de 22 a 35 anos, e *Faixa C*, acima de 60 anos) e *nível de escolarização* (ensino médio completo e ensino superior completo). A estratificação da amostra encontra-se disposta no quadro 1.

[6] Schmidt (2024) e Ruela (2025) realizaram pesquisas de iniciação científica analisando o /S/ e o /R/ em posição de coda final na variedade capixaba, respectivamente. Schmidt (2024, 2025) identificou a variante fricativa aoveolar como praticamente categórica para a realização do /S/ em coda e Ruela (2024, 2025) constatou a variante fricativa glotal para a realização do /R/ em posição final.

Quadro 1: Estratificação social e dados gerais dos informantes da Amostra Duarte (2025)

Informante	Sexo	Idade	Local de nascimento	Cidade que reside atualmente	Local de nascimento dos pais
Informante 1	Masculino	23 anos	Vargem Alta (ES)	Vitória (ES)	Vargem Alta (ES)
Informante 2	Masculino	27 anos	Vitória (ES)	Vila Velha (ES)	Mãe: Vitória (ES); Pai: Vila Velha (ES)
Informante 3	Masculino	71 anos	Vitória (ES)	Vitória (ES)	Baixo Guandu (ES)
Informante 4	Masculino	67 anos	Cariacica (ES)	Cariacica (ES)	Cariacica (ES)
Informante 5	Feminino	30 anos	Vitória (ES)	Vitória (ES)	Mãe: Vitória (ES); Pai: Rio de Janeiro (RJ)
Informante 6	Feminino	60 anos	São José do Calçado (ES)	Serra (ES)	São José do Calçado (ES)
Informante 7	Feminino	22 anos	Vitória (ES)	Vitória (ES)	Vitória (ES)

Fonte: Elaboração das autoras.

A coleta foi realizada em três momentos. Primeiramente, realizamos uma entrevista de experiência pessoal nos moldes variacionistas. Após a realização das entrevistas, procedemos à leitura de uma lista de itens lexicais, preparada utilizando itens previamente selecionados (sendo palavras paroxítonas e oxítonas com a sílaba tônica terminada em /S/, propiciando a realização da ditongação, além de itens distratores). Por fim, gravamos a leitura de cinco textos que incluíam os itens lexicais passíveis de ditongação[7]. Com vistas a garantir a qualidade da gravação, utilizamos o gravador digital ZOOM H4N e o microfone unidirecional AKG C520L. Os áudios foram gravados em taxa de amostragem de 44100Hz.

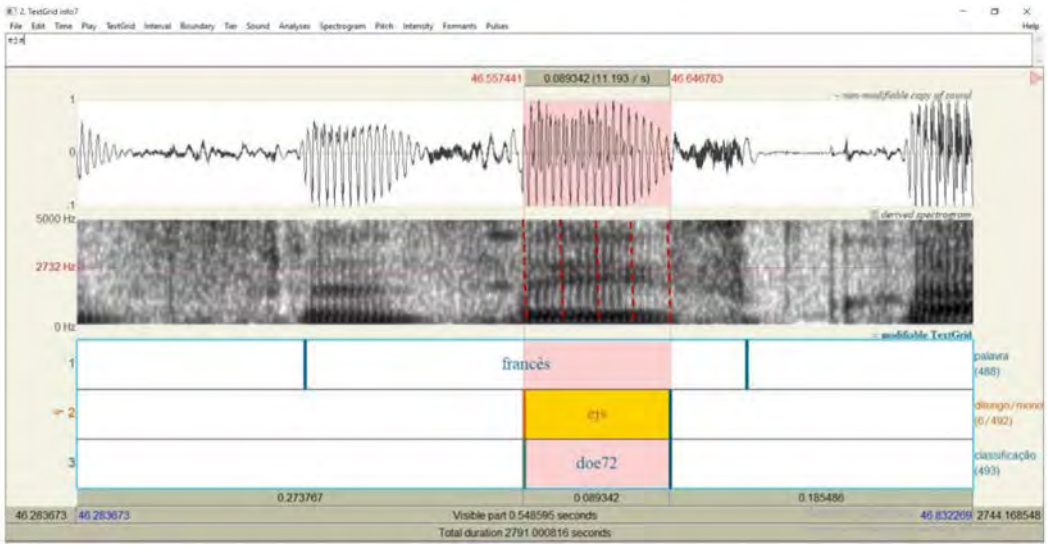
Posteriormente à coleta, realizamos as seguintes etapas: (i) transcrição, etiquetagem manual e classificação dos dados; (ii) levantamento das ocorrências que apresentam sílabas tônicas travadas por /S/; (iii) análise da trajetória formântica dos dados (valores de F1 e F2) no *software* Praat (versão 6.4.47); (iv) organização dos dados em planilhas de forma a possibilitar o estudo da influência das variáveis preditoras de efeito fixo – *posição do acento, estilo, vogal, sexo, nível de escolarização, faixa etária* – e das variáveis de efeito aleatório – *item lexical* e *informante*; (v) submissão dos dados à análise estatística na Plataforma R (RStudio); (vi) análise e interpretação dos resultados.

Para o exame das trajetórias formânticas das ocorrências, extraímos as informações dos primeiro e segundo formantes (F1 e F2)[8] de cinco pontos equidistantes no intervalo correspondente à etiquetagem do segmento-alvo, visando delimitar a trajetória dos valores. Esse procedimento pode ser verificado na imagem 2.

[7] Projeto aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal do Espírito Santo (CEPE/UFES), sendo 86924924.6.0000.5542 o número do processo.

[8] F1 e F2 são medidas em Hz que indicam, respectivamente, o grau de elevação da língua no momento de produção do dado e do segundo formante e o grau de anteriorização da língua durante a realização do segmento-alvo.

Imagem 2: Representação da coleta dos dados das trajetórias formânticas dos segmentos-alvo analisados em 5 pontos equidistantes da etiquetagem (linhas tracejadas em vermelho)



Fonte: Elaboração das autoras.

Com base na análise de oitiva atrelada à inspeção acústica dos dados, buscamos delimitar um critério que guiasse a categorização da variável-resposta em *forma ditongada versus forma não ditongada*.

No processo de definição das variáveis preditoras linguísticas e extralinguísticas, bem como os níveis que compõem cada uma delas, tomamos como base os resultados de estudos acerca do fenômeno da ditongação realizados em outras localidades do Brasil (*cf.* seção 2). O quadro 2 exibe as variáveis consideradas.

Quadro 2: Envelope de variação

Variável resposta	
ditongação	formas não ditongadas (*mas, três, cuscuz*) formas ditongadas (*mais, três, cuscuiz*)
Variáveis preditoras de efeito fixo	
vogal-núcleo do segmento-alvo	/a/ (*mas ~ mais*) /e/ (*mês ~ mêis*) /ɛ/ (*dez ~ deiz*) /o/ (*arroz ~ arroiz*) /ɔ/ (*voz ~ voiz*) /u/ (*luz ~ luiz*)
posição do acento na palavra	monossílabo tônico (*três ~trêis*) oxítona (*capuz ~ capuiz*) paroxítona (*mesmo ~ meismo*) proparoxítona (*ginástica ~ ginaistica*)

estilo	fala leitura
sexo	feminino masculino
faixa etária	faixa A (de 22 a 35 anos) faixa C (Mais de 60 anos)
escolarização	ensino médio ensino superior
Variáveis preditoras – efeito aleatório	
item lexical informante	

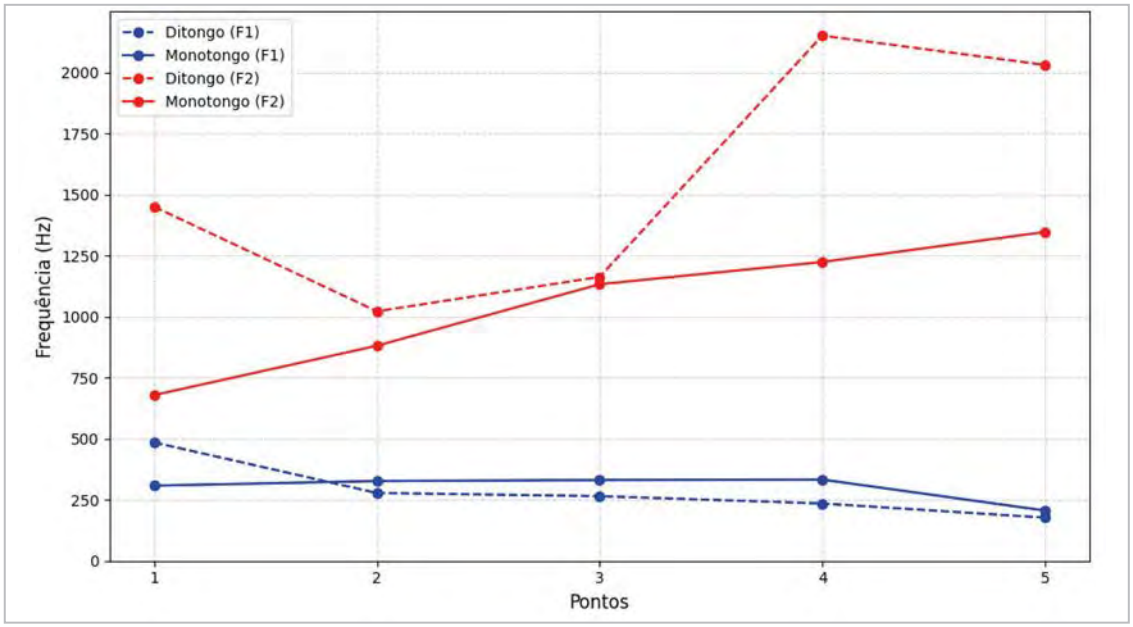
Fonte: Elaboração das autoras.

Na próxima seção, apresentamos os resultados, contemplando tanto a análise descritiva, com a exposição das distribuições gerais e das tendências observadas, quanto a análise inferencial, na qual se discutem os efeitos estatísticos das variáveis linguísticas e extralinguísticas sobre a ocorrência da ditongação.

4. Resultados e discussão

O primeiro passo da análise quantitativa da ditongação consistiu na definição de um critério acústico adequado para distinguir, de forma alinhada à percepção auditiva, vogais ditongadas de monotongos. A análise preliminar dos correlatos acústicos indicou que o segundo formante (F2), associado à anterioridade/posterioridade da língua, mostrou-se um parâmetro eficiente na identificação da presença/ausência do glide /j/ nos dados em exame. Em contraste, o primeiro formante (F1), relacionado à altura vocálica, permaneceu praticamente inalterado, independentemente da ocorrência de ditongação. O gráfico 1 ilustra essa diferença para a vogal /o/. Observamos que a trajetória de F1 se mantém estável tanto no monotongo [o] quanto no ditongo [oj]. Já a trajetória de F2 apresenta um aumento acentuado no final da produção do ditongo, refletindo o deslocamento da língua em direção à posição anterior, característico do glide /j/. Esse aumento ocorre no momento final da produção vocálica, imediatamente antes da transição para a sibilante.

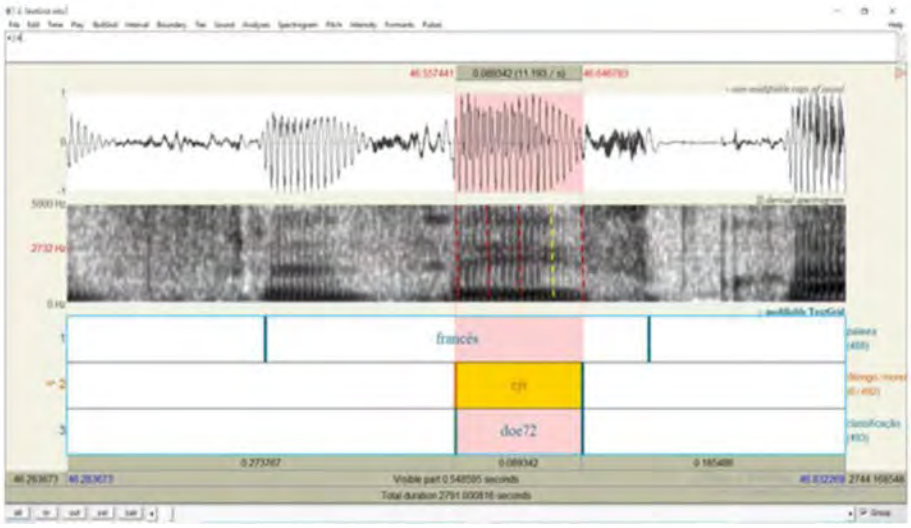
Gráfico 1: Comparação entre ditongo e monotongo da vogal [o]



Fonte: Elaboração das autoras.

A partir dessa constatação, o valor de F2 foi definido como o parâmetro para a classificação das ocorrências em monotongos ou ditongos. Para garantir a precisão desse parâmetro, foi necessário determinar um ponto específico na trajetória formântica que capturasse a articulação do glide amenizando a interferência dos efeitos de coarticulatórios do contexto seguinte. Como os segmentos-alvo da pesquisa são seguidos pela sibilante /S/ (realizada categoricamente como alveolar em nossa amostra), o ponto final da vogal (último tracejado) já apresenta influências da produção dessa consoante. Por essa razão, optamos por extrair o valor de F2 no penúltimo ponto de medição (linha tracejada amarela), conforme ilustrado na imagem 3.

Imagem 3: Representação da coleta dos dados das trajetórias formânticas dos segmentos-alvo analisados com enfoque no penúltimo ponto (linha tracejada amarela)



Fonte: Elaboração das autoras.

Na análise quantitativa dos dados, o fenômeno da ditongação em sílaba tônica ocorreu em 35,5% do *corpus* analisado, correspondendo a 428 de um total de 1.171 ocorrências. O índice geral de ditongação observado na amostra (35,5%) mostra-se superior ao valor de 15% reportado por Silva (2014) para Vitória – que analisa a ditongação em contextos tônicos e átonos, diferentemente do que realizamos neste estudo e nos aqui analisados sobre a fala carioca, que levam em consideração a realização da ditongação diante de /S/ em sílabas tônicas -, e próximo ao das taxas registradas para a variedade carioca, superior a 40% (Callou, 2009; Silveira, 2023).

O percentual geral é inferior à nossa expectativa inicial, de que o valor seria superior a 50%, hipótese conjecturada com base no fato de a ditongação ser frequentemente apontada como um traço característico da variedade capixaba. Acreditamos que, embora a ditongação em sílabas tônicas travadas por /S/ seja frequentemente apontada, tanto na literatura quanto no senso comum, como um marcador saliente do falar espírito-santense, os dados empíricos revelam um quadro mais complexo.

Em uma análise inicial, o valor que encontramos poderia ser interpretado como contraditório. No entanto, acreditamos que essa aparente discrepância (entre os valores esperados e os valores observados) seria reflexo da saliência sociolinguística do fenômeno. Em outras palavras, acreditamos que a ditongação mesmo não apresentando frequência elevada seria muito perceptível. Um exemplo clássico de fenômeno saliente no português é a não marcação da concordância verbal de terceira pessoa do plural (*eles come* vs. *eles comem*), que, conforme indica a literatura (Scherre; Naro, 1997; Monguilhott, 2009; Chaves, 2017; Naro, 1981; Vieira, 1997), nas comunidades urbanas, é verificada em cerca de 20 a 30% dos dados. Apesar da relativa baixa frequência, o fenômeno é altamente perceptível e estigmatizado (caracterizando-se, nos termos labovianos, como um estereótipo negativo).

Além disso, conforme Freitag (2018), há indícios de que os significados sociais da variação fonética (no caso a ditongação) seriam avaliados pelos ouvintes com maior facilidade do que variáveis de nível gramatical (no caso da concordância verbal). Essa percepção seria mediada pelo que Labov *et al*. (2006) denominaram *monitor sociolinguístico*, mecanismo cognitivo que atuaria no processamento de informações socialmente marcadas. Em estudo de percepção conduzido sobre a variação na pronúncia de – ing (*working ~ workin)*, Labov *et al.* demonstraram que os sujeitos são sensíveis a oscilações de frequência a partir de 10%.

Dessa forma, defendemos que a taxa de 35,5% é suficiente para que o fenômeno da ditongação tenha se cristalizado como um índice identitário da variedade capixaba ainda que não seja, em termos absolutos, um traço linguístico altamente frequente. Diferentemente de variáveis estigmatizadas que operam "abaixo da consciência" (indicadores), a ditongação na RMGV parece comportar-se como um marcador ou mesmo como um estereótipo, situando-se acima do nível de consciência social. Assim, acreditamos que a percepção de que a ditongação ocorre com maior frequência do que os dados quantitativos sugerem pode ser atribuída à sua elevada saliência. Estudos nesse sentido, como o de Tesch (2022), dão suporte a essa afirmação.

A seguir, na tabela 1, apresentamos os resultados referentes à distribuição dos dados para cada variável social e linguística considerada.

Tabela 1: Valores brutos e percentuais referentes a cada variável linguística e extralinguística controlada (análise descritiva)

variável		n (ditongação)	total	%
sexo	feminino	275	604	45,53
	masculino	153	567	26,98
faixa etária	A (mais jovens)	285	635	44,88
	C (mais velhos)	143	536	26,68
escolarização	EM	190	529	35,92
	ES	238	642	37,07
posição do acento	monossílabo	297	447	66,44
	oxítona	122	250	48,8
	paroxítona	8	467	1,71
	proparoxítona	1	7	14,29
vogal-núcleo	ɛ	3	28	17,3
	ɔ	29	216	8,8
	e	123	309	39,7
	o	6	38	15,8
	u	7	53	13,2
	a	245	393	62,3
estilo	fala	277	490	56,53
	leitura	151	681	22,17
		428	1.171	35,5%

Fonte: Elaboração das autoras.

A tabela 1 apresenta uma distribuição da ditongação em relação às variáveis sociais e linguísticas controladas em nosso *corpus*. No que tange aos fatores sociais, destacam-se as variáveis *sexo* e *faixa etária*. As informantes do sexo feminino apresentaram taxa de ditongação de 45,5%, consideravelmente superior aos 27% registrados para os sujeitos do sexo masculino. Em relação à *faixa etária*, a *faixa A* (mais jovens) ditongou em 44,9% dos casos, quase o dobro da *faixa C* (mais velhos), a qual apresentou uma taxa de 26,7%. Para a variável *escolarização*, por sua vez, registramos valores aproximados, 35,9% (Ensino Médio) e 37,1% (Ensino Superior).

No que toca aos fatores linguísticos, a variável *posição do acento* apresentou prevalência do fenômeno em monossílabos tônicos (66,4%), como em *mês ~ mêis*, e oxítonas (48,8%), como em *vocês ~ vocêis*, e aplicação reduzida em palavras paroxítonas (1,7%), como em *mesmo ~ meismo* e proparoxítonas (14,9%), como em *ginástica ~ gináistica*. A variável *vogal-núcleo do segmento-alvo* também se mostrou associada a diferentes taxas de ocorrência: o fenômeno foi mais recorrente nos dados com a vogal baixa (62,3%), como em *paz ~ paiz*, e com a vogal média-alta (39,8%), como em *vocês ~ vocêis*. Para as demais vogais, as proporções de ditongação observadas foram inferiores, situando-se entre 8,8% e 17,3%. Em reação à variável estilística, a *fala espontânea* apresentou mais ocorrência do fenômeno (56,5%) em detrimento do estilo de *leitura* (22,2%).

Para além das tendências gerais, a análise da variabilidade individual, exibida na tabela 2, revela a heterogeneidade do fenômeno entre os falantes.

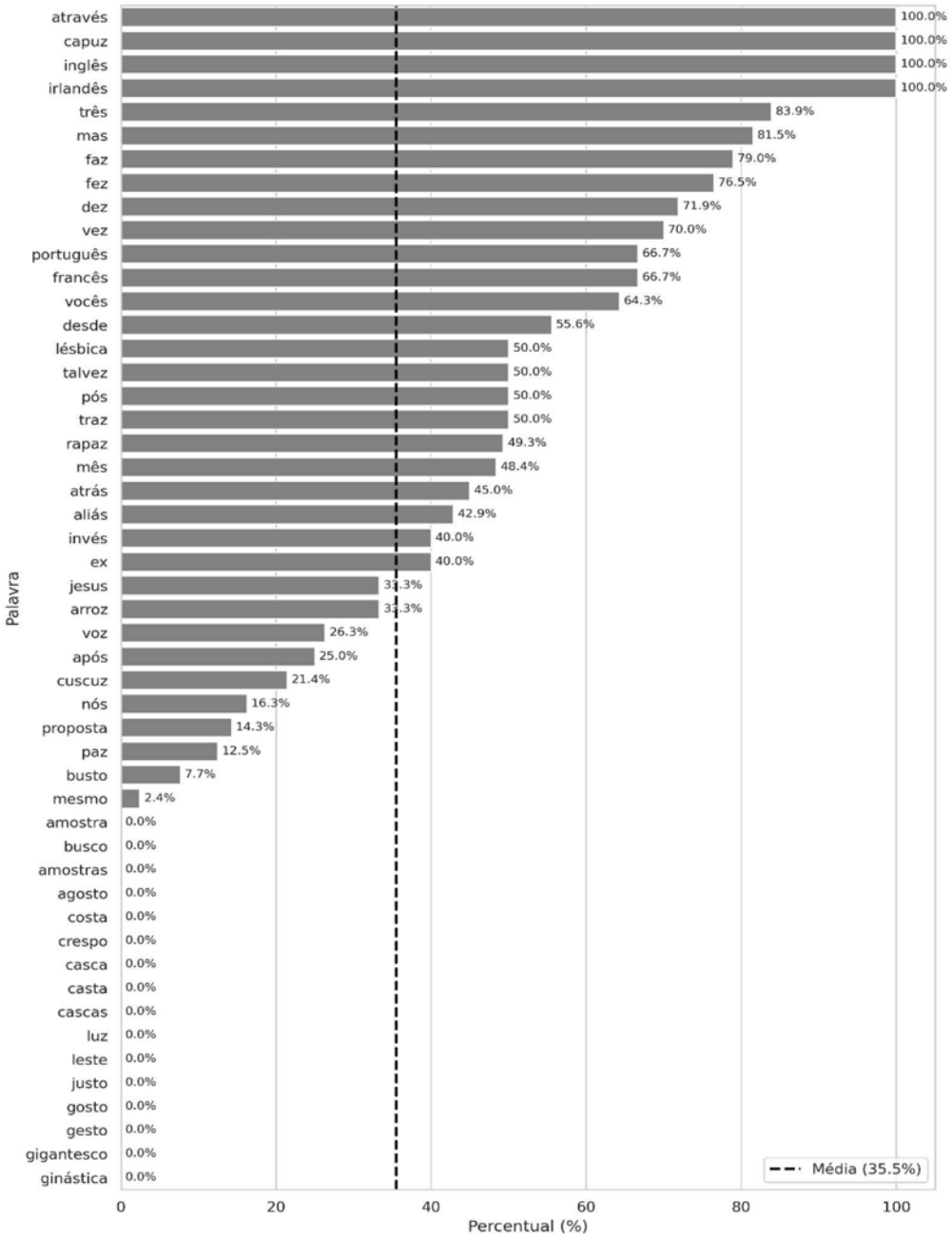
Tabela 2: Valores brutos e percentuais do fenômeno de ditongação por informante

Categoria	n (Ditongação)	Total	Percentual (%)
INFORMANTE 1 (23 anos, homem, Vitória)	54	129	41,86
INFORMANTE 2 (27 anos, homem, Vila Velha)	34	115	29,57
INFORMANTE 3 (71 anos, homem,Vitória)	33	154	21,43
INFORMANTE 4 (67 anos, homem, Cariacica)	32	169	18,93
INFORMANTE 5 (30 anos, mulher, Vitória)	94	145	64,83
INFORMANTE 6 (60 anos, mulher, Serra)	78	213	36,62
INFORMANTE 7 (22 anos, mulher,Vitória)	103	246	41,87

Fonte: Elaboração das autoras.

A *Informante 5*, jovem de 30 anos natural de Vitória, destaca-se como a participante que mais ditongou, com uma taxa de 64,8%. Cabe, no entanto, destacar que o pai da participante é carioca, o que pode ter impacto no índice verificado, já que no Rio de Janeiro a ditongação apresenta taxas próximas a essa, de acima de 40%. Seguindo padrão semelhante, o *Informante 1* (23 anos, Vargem Alta, residente em Vitória) e a *Informante 7* (22 anos, natural de Vitória) também apresentaram taxas elevadas, aproximadamente 42%. Em contrapartida, os informantes mais velhos, como o *Informante 3* (71 anos, Vitória) e o *Informante 4* (67 anos, Cariacica), exibiram um comportamento mais conservador, com as taxas inferiores de ditongação, 21,4% e 18,9%, respectivamente. Esse contraste entre falantes mais jovens e mais velhos reforça os padrões observados na tabela 1 e sugere que a ditongação pode estar passando por um processo de mudança liderado por mulheres jovens na variedade capixaba.

Gráfico 2: Proporção de ocorrência de ditongação por item lexical (50 das 73 palavras registradas em nosso *corpus*)



Fonte: Elaboração das autoras.

Tabela 3: Valores brutos e percentuais da ditongação por item

palavra	ditongação/ total de dados	% de ditongação
através	2/2	100
capuz	1/1	100
inglês	12/12	100
irlandês	1/1	100
três	26/31	83,87
mas	172/211	81,52
faz	15/19	78,95
fez	13/17	76,47
dez	23/32	71,88
vez	14/20	70
português	8/12	66,67
francês	2/3	66,67
vocês	18/28	64,29
desde	5/9	55,56
lésbica	1/2	50
talvez	3/6	50
pós	1/2	50
traz	2/4	50
rapaz	36/73	49,32
mês	15/31	48,39
atrás	9/20	45
aliás	9/21	42,86
invés	2/5	40
ex	4/10	40
Jesus	3/9	33,33
arroz	6/18	33,33
voz	5/19	26,32
após	6/24	25
cuscuz	3/14	21,43
nós	7/43	16,28
proposta	1/7	14,29
paz	1/8	12,5
busto	1/13	7,69
mesmo	1/41	2,44
amostra	0/5	0

Fonte: Elaboração das autoras.

Conforme podemos observar no gráfico 2 e tabela 3, a ditongação atua com maior ou menor frequência a depender do item lexical. Foi registrada aplicação categórica de 100% do fenômeno, por exemplo, na palavra *inglês*, a qual contabilizou 12 ocorrências, e de 0% em palavras como *texto*, computada 42 vezes. As palavras que apresentaram taxas acima da média encontrada de ditongação e que foram produzidas ao menos cinco vezes em nosso *corpus* foram: *inglês* (100%), *três* (83,9%),

mas (81,5%), *faz* (78,9%), *fez* (76,5%), *dez* (71,9%), *vez* (70%), *português* (66,7%), *vocês* (64,3%), *desde* (55,6%), *talvez* (50%), *rapaz* (49,3%), *mês* (48,4%), *atrás* (45%), *aliás* (42,9%), *invés* (40%). As palavras com maiores índices de ditongação são monossílabas tônicas, oxítonas e com duas sílabas (exceto *português*, com três sílabas). Uma única paroxítona figura entre as palavras com altos percentuais de ditongação, o item *desde* (*deisde*). Interpretamos que, nesse caso, os falantes poderiam estar analisando o item como *des de*. Nessa reanálise, o primeiro elemento estaria sendo interpretado como um monossílabo tônico e o segundo elemento, como uma preposição/clítico. Assim, esse não seria um exemplar paroxítono, mas sim, de um monossílabo tônico (*des* ~ *deis*). Por outro lado, as palavras em que não se registrou ocorrência do fenômeno em estudo foram majoritariamente paroxítonas. Dentre os itens desse grupo acentual, verificamos ditongação apenas em *proposta* (14,3%), *busto* (7,7%) e *mesmo* (4%).

Tal constatação indica a necessidade de uma investigação mais aprofundada, especialmente no que se refere à possível interação entre forças neogramáticas *versus* difusionistas (Labov, 1994; Phillips, 2006). Nesse sentido, acreditamos que o léxico, atrelado à saliência, possa desempenhar papel relevante na percepção do fenômeno como um traço reconhecível como típico do falar capixaba.

Para investigar os fatores linguísticos e sociais que se correlacionam ao fenômeno variável da ditongação, utilizamos um modelo de regressão logística de efeitos mistos. O modelo final incluiu as variáveis *faixa etária, estilo, posição do acento* e v*ogal núcleo do segmento-alvo* (variáveis de efeito fixo), além das variáveis para *informante* e *palavra* (variáveis de efeito aleatório). Os resultados deste modelo encontram-se dispostos na tabela 4.

Tabela 4: Resultados da análise de regressão de efeitos mistos

Variável (efeito fixo)	Estimativa	Erro-padrão	z-valor	p-valor
Intercepto: Faixa A - mais velhos, proparoxítona, vogal média-alta posterior, estilo:fala	-3,4163	0,73920	-4,622	< 0,001***
FAIXA ETÁRIA: Faixa A - mais jovens	1,8223	0,35788	5,092	< 0,001***
POSIÇÃO DO ACENTO: monossílabo	3,4814	0,66845	5,208	< 0,001***
POSIÇÃO DO ACENTO: oxítona	3,5651	0,61490	5,798	< 0,001***
POSIÇÃO DO ACENTO: paroxítona	-1,4178	0,46290	-3,063	0,002**
VOGAL: média-baixa anterior	0,9903	0,52389	1,890	0,059
VOGAL: média-baixa posterior	-0,9286	0,40216	-2,309	0,021*
VOGAL: baixa	0,6397	0,46722	1,369	0,171
VOGAL: média-alta anterior	0,9586	0,34515	2,777	0,005**
VOGAL: alta posterior	-0,6321	0,70428	-0,897	0,369
ESTILO: leitura	-1,9202	0,43523	-4,412	< 0,001***

Modelo (DITONCAÇÃO ~ FAIXA.ETARIA + POSIÇÃO DO ACENTO + VOGAL + ESTILO+(1|ITEM>LEXICAL) + (1|INFORMANTE), data = dados, family = binomial)

Fonte: Elaboração das autoras.

Os resultados do modelo, apresentados na tabela 4, indicam efeitos significativos das variáveis *faixa etária, posição do acento, vogal núcleo do segmento-alvo e estilo*. Observamos forte efeito da faixa etária, com a *Faixa A* (mais jovens) apresentando coeficiente positivo e estatisticamente significativo ($\beta = 1{,}8223$; $p < 0{,}001$), o que indica aumento na probabilidade de ocorrência da ditongação em comparação à faixa de referência (*Faixa C* – mais velhos).

A distribuição da ditongação nas faixas etárias investigadas sinaliza uma mudança em curso. Essa afirmação é feita com cautela, já que estamos diante de um estudo em tempo aparente. A maior probabilidade de aplicação da variante ditongada entre falantes mais jovens, aliada ao comportamento mais conservador dos falantes mais velhos, reproduz um padrão clássico de mudança linguística (Labov, 2008 [1972]; 1994). Esse resultado é particularmente relevante à luz da literatura sobre o sudeste, uma vez que Silva (2014) já havia identificado indícios de mudança nas capitais.

Entre os fatores linguísticos, o efeito mais expressivo foi o da *posição do acento*. As categorias *monossílabo* ($\beta = 3{,}4814$; $p < 0{,}001$) e *oxítona* ($\beta = 3{,}5651$; $p < 0{,}001$) elevaram significativamente a chance de ocorrência da ditongação em relação ao nível de referência (proparoxítonas). Esses resultados indicam que a posição do acento exerce papel central na realização da variante analisada.

No que se refere à *vogal núcleo do segmento-alvo*, apresentaram efeito significativos a vogal média-alta anterior ($\beta = 0{,}9586$; $p < 0{,}05$) e a vogal média-baixa posterior ($\beta = -0{,}9286$; $p < 0{,}05$). A vogal média-alta anterior (*três ~ treis, vocês ~ vocêis*) mostrou aumento da probabilidade de ditongação, ao passo que a vogal média baixa-posterior (*após ~ apóis, nós ~ nóis*) diminui a probabilidade de ocorrência do fenômeno ambas comparadas à vogal central. A tendência observada na análise descritiva de que a vogal baixa /a/ e as vogais médias anteriores (tanto a média-alta /e/ quanto a média-baixa /ɛ/) colaboram com a ocorrência do fenômeno, em consonância com os resultados reportados para a variedade carioca por Callou (2009) e Silveira (2023), foi confirmada para a vogal /e/. Tal favorecimento da vogal média-alta anterior pode ser explicado por fatores articulatórios e acústicos: vogais anteriores apresentam maior propensão à emergência do glide /j/ devido à maior amplitude do movimento articulatório da língua e à maior saliência das transições formânticas. Em contraste, vogais posteriores e mais fechadas impõem maiores restrições articulatórias a esse movimento, reduzindo a probabilidade de ditongação.

Por fim, o *estilo* também se mostrou relevante: o *estilo* leitura reduziu significativamente a probabilidade de ocorrência do fenômeno ($\beta = -1{,}9202$; $p < 0{,}001$). A redução da ditongação no estilo de leitura indica uma tentativa dos informantes de reprodução fiel da escrita. Esse comportamento estilístico aproxima a ditongação do estatuto de marcador sociolinguístico, definido por Labov (2008 [1972]) como uma variável sensível à atenção do falante.

5. Considerações finais

Este estudo teve como objetivo descrever o comportamento variável da ditongação com glide /j/ em sílabas tônicas travadas por /S/ na fala da RMGV, sob a perspectiva da Teoria da Variação e

Mudança Linguística. A análise de 1.171 ocorrências revelou um índice geral de 35,5% de ditongação, valor superior ao previamente registrado para Vitória. Esse resultado indica que a ditongação constitui um traço perceptível e socialmente reconhecido da fala capixaba, embora não se configure como um traço com altas taxas de aplicação.

No que se refere às perguntas de pesquisa formuladas na introdução, os resultados permitem responder que: (i) o índice de ditongação na Grande Vitória é intermediário no panorama do sudeste, sendo superior ao valor anteriormente registrado para Vitória (ii) as trajetórias formânticas, em especial os valores de F2, corroboram de forma consistente à análise de oitiva, validando a distinção entre monotongos e ditongos a partir de critérios acústicos objetivos; e (iii) os principais fatores correlacionados à variante ditongada foram a *posição do acento*, a *faixa etária*, o *estilo de fala* e a *vogal núcleo do segmento-alvo*.

Por fim, este trabalho contribui com a descrição de um aspecto fonético-fonológico do sotaque capixaba. Os resultados reforçam que a fala da Grande Vitória constitui dados que podem ser proveitosos para a discussão sobre saliência estrutural, especialmente em uma variedade em que poucos traços são apontados como marcas dialetais. Nesse sentido, a análise vai ao encontro a premissa de Labov (2006) de que traços identitários podem apresentar frequências de aplicação inferiores a 50%, por vezes em torno de 10%, e ainda assim serem socialmente salientes. Portanto, a ditongação na RMGV é um elemento relevante para o estudo do *monitor sociolinguístico*, sugerindo que pesquisas futuras focadas na percepção da própria população em relação a este traço da fala possam expandir as discussões aqui apresentadas, mapeando como esses índices de frequência se convertem em avaliações subjetivas e estigmas ou prestígios na comunidade de fala.

Referências

ARAGÃO, Maria do Socorro Silva de. Ditongação e Monotongação nas Capitais Brasileiras. *In*: CONGRESSO INTERNACIONAL DA ASSOCIAÇÃO DE LINGUÍSTICA E FILOLOGIA DA AMÉRICA LATINA, 17., 2014, João Pessoa. *Anais* [...]. João Pessoa: Ideia, 2014. pp. 2089-2101.

BARBOSA, Plínio; MADUREIRA, Sandra. *Manual de fonética acústica experimental*: aplicações a dados do português. São Paulo: Cortez, 2015.

BISOL, Leda. O ditongo na perspectiva da fonologia atual. D.E.L.T.A., Rio Grande do Sul, v. 5, n. 2, pp. 185-224, 1989.

CALLOU, Dinah. Variação e mudança no âmbito do consonantismo. *In:* PAIVA, Maria da Conceição de; D'ANGELIS, Wilmar da Rocha (org.). *Estudos de Sociolinguística:* variações linguísticas no português brasileiro. Rio de Janeiro: Editora UFRJ, 2009.

CHAVES, R. G. *A redução/desnasalização de ditongos nasais átonos finais e a marcação explicita de CVP6:* um estudo de correlação. 2017. 359 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2017.

DUARTE, Ana Carolina Carnieli. *A ditongação em sílabas tônicas travadas por /S/ no dialeto capixaba*. Vitória, 2025. Relatório Técnico-Científico.

FREITAG, Raquel Meister Ko. Saliência estrutural, distribucional e sociocognitiva. *Acta Scientiarum. Language and Culture*, Maringá, v. 40, n. 2, e41173, pp. 1-11, 2018. Disponível em: https://eduemojs.uem.br/ojs/index.php/ActaSciLangCult/article/view/41173. Acesso em: 13 jun. 2026.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Vitória (ES)*: Cidades e Estados. Rio de Janeiro: IBGE, 2022. Disponível em: https://www.ibge.gov.br/cidades-e-estados/es/vitoria.html. Acesso em: 20 abr. 2026.

LABOV, William. *Principles of linguistic change:* linguistic factors. Oxford: Blackwell, 1994.

LABOV, William *et al*. Listeners' sensitivity to the frequency of sociolinguistic variables. *In*: FRIESNER, M. L.; RAVINDRANATH, M. (eds.). *Penn Working Papers in Linguistics*: selected papers from NWAV 34. Philadelphia: University of Pennsylvania; Penn Linguistics Club, 2006. v. 12, n. 2, pp. 105-129.

LABOV, William. *Padrões Sociolinguísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].

LACERDA, Marcela Langa; LEPPAUS, Ana Clara S. Capixaba tem sotaque? Quem pode dar essa resposta? Questões para a sociolinguística da sociedade. *Confluência*, n. 65, pp. 73-94, 2023.

OPENAI. *ChatGPT* [software]. Versão GPT-4. [*S.l.*]: OpenAI, 2026. Disponível em: https://chat.openai.com. Acesso em: 31 jul. 2026.

PHILLIPS, Betty. Word frequency and the actuation of sound change. *Language*, v. 60, n. 2, pp. 320-342, 1984.

MONGUILHOTT, Isabel de Oliveira e Silva. *Estudo sincrônico e diacrônico da concordância verbal de terceira pessoa do plural no PB e no PE*. 2009. 229 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2009.

NARO, Anthony Julius. The social and structural dimensions of a syntactic change. *Language*, Baltimore, v. 57, n. 1, pp. 63-98, 1981.

RUELA, Emilene. */R/ em coda medial na fala capixaba*. Vitória, 2025. Relatório Técnico-Científico.

SCHERRE, Maria Marta Pereira; NARO, Anthony Julius. A concordância de número no português do Brasil: um caso típico de variação inerente. *In*: HORA, Dermeval da (org.). *Diversidade linguística no Brasil*. João Pessoa: Ideia, 1997. pp. 93-114.

SCHMIDT, Sophia. *A variação de /S/ em coda:* análise longitudinal da fala de Viviane Mosé e Elisa Lucinda. Vitória, 2024. Relatório Técnico-Científico.

SILVA, Amanda dos Reis. A ditongação em sílabas fechadas por /s/ no *corpus* do Projeto ALIB: nas trilhas das capitais brasileiras. 2014. Dissertação (Mestrado em Língua e Cultura) – Universidade Federal da Bahia, Salvador, 2014.

SILVEIRA, Anna Carolina Lopes da. A trajetória da ditongação de monossílabos na variedade carioca: um estudo em tempo real. 2023. Trabalho de Conclusão de curso. Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2023.

TESCH, Leila. O sotaque capixaba: um estudo de percepção. (*Con)Textos Linguísticos*, Vitória, v. 16, n. 34, pp. 225-242, 2022.

VIEIRA, S. A não-concordância em dialetos populares: uma regra variável. *Graphos*, v. 2, n. 1, pp. 115-133, 1997.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução de Marcos Bagno. São Paulo: Parábola, 2006 [1968].

YACOVENCO, Lilian Coutinho *et al*. Projeto PortVix: a fala de Vitória/ES em cena. *Alfa:* Revista de Linguística, São José do Rio Preto, v. 56, n. 2, pp. 771-792, 2012.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a71836

Gênero social e mudança linguística: reflexões teórico-metodológicas para a Sociolinguística

Gender and language change: theoretical and methodological reflections for Sociolinguistics

Camila Bordonal Clempi[1]

Angélica Rodrigues[2]

RESUMO

Neste artigo, reconhecemos que a correlação entre gênero social e variação linguística tem sido frequentemente orientada por interpretações essencialistas. Como alternativa, propomos tratar a variável gênero enquanto dimensão social e histórica, performativamente constituída e reiterada por práticas discursivas, por meio de uma releitura teórico-metodológica fundamentada nas teorias da performatividade (Butler, 2018a, 2018b) e da tecnologia de gênero (Lauretis, 1994). Para demonstrar a aplicabilidade dessa perspectiva, realizamos uma análise sociolinguística histórica da variação de marcas de segunda pessoa (2P/tu) e terceira pessoa (3P/você) na posição de sujeito no início do século XX (1915-1932) no português brasileiro. Utilizamos um córpus composto por 432 cartas publicadas na revista *A Cigarra*, voltada ao público feminino, contrapondo-o a um grupo controle, formado por outras 416 cartas distribuídas em duas diferentes amostras, definidas por um *continuum* de representações de gênero: a *Revista da Semana* ([+feminina ↔ -masculina]) e o *Correio Paulistano* (publicação [+masculina ↔ -feminina]). Os resultados evidenciam que o contexto social e a proposta discursiva das publicações influenciam o uso das variantes em estudo. Nossa análise revela como a imprensa feminina, enquanto tecnologia de gênero, constitui espaço de performatividades, oferecendo novos olhares para a compreensão do significado social na mudança linguística.

PALAVRAS-CHAVE: Sociolinguística Variacionista. Sociolinguística Histórica. Estudos de Gênero. Gênero feminino. Imprensa feminina.

ABSTRACT

In this article, we acknowledge that the correlation between gender and linguistic variation has often been guided by essentialist interpretations. As an alternative, we propose treating gender as a social and historical dimension, performatively constituted and continuously reiterated through discourse practices, based on a theoretical-methodological reinterpretation grounded on performativity theory (Butler, 2018a, 2018b) and gender as a technology (Lauretis, 1994). In order to demonstrate the applicability of this perspective, we conduct a historical sociolinguistic analysis of variation between second (2P/tu) and third-person (3P/você) forms in subject position in early twentieth-century Brazilian Portuguese (1915-1932). Our study draws on a *corpus* of 432 letters published in *A Cigarra*, a magazine directed toward female readership, contrasting it with a control group, composed of 416 letters distributed across two different samples, defined along a continuum of gendered representations: *Revista da Semana* ([+feminine ↔ −masculine]) and *Correio Paulistano*

[1] Doutora em Linguística e Língua Portuguesa pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP/FCLAr), camila.clempi@unesp.br, https://orcid.org/0000-0003-1549-400X.

[2] Professora Associada livre-docente na Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP/FCLAr), angelica.rodrigues@unesp.br, https://orcid.org/0000-0003-1470-4634.

Recebido em 15/12/2025 - Aceito em 10/06/2026

([+masculine ↔ −feminine]). Results suggest that both social context and discourse orientation shape the use of the variants in the material under investigation. Our analysis shows how women's press, understood as a gender technology, constitutes a space for performative practices, offering new perspectives for understanding social meaning in processes of language change.
KEYWORDS: Variationist Sociolinguistics. Historical Sociolinguistics. Gender Studies. Female gender. Women's press.

1. Introdução

Desde os trabalhos fundadores de Weinreich, Labov e Herzog (2006 [1968]) e de Labov (2016 [1972]), a Sociolinguística se consolidou como uma área de estudo da linguística que compreende a língua como heterogênea, condicionada por fatores linguísticos e sociais e passível de análise em contextos reais de comunicação. Nesse modelo teórico-metodológico, a heterogeneidade ou a variação linguística são analisadas, portanto, tendo em vista variáveis internas e externas ao sistema linguístico (Weinreich; Labov; Herzog, 2006 [1968]), tornando possível a investigação de padrões gerais de usos linguísticos na sociedade.

Sob essa perspectiva, variáveis como idade, escolaridade, classe social e gênero se estabeleceram na área como importantes categorias de estratificação social. No entanto, chamamos a atenção para o fato de que o gênero – também entendido, na literatura sociolinguística, sob outros rótulos, tais quais sexo, sexo/gênero – ainda é, muitas vezes, controlado e tratado como uma categoria estável. Esse modo de operacionalização pode resultar em interpretações essencialistas[3] que associam o comportamento linguístico feminino, por exemplo, a prestígio ou conservadorismo, sem considerar a complexidade e a circunstância nas quais as práticas generificadas ocorrem (Bucholtz, 2002; Eckert; McConnell-Ginet, 2010; Wodak; Benke, 2017).

Neste artigo, retomamos tal discussão propondo uma releitura teórico-metodológica do gênero na Sociolinguística, amparada nas teorias sociais de performatividade (Butler, 2018a, 2018b) e de tecnologia de gênero (Lauretis, 1994). Nosso propósito é demonstrar a operacionalidade dessas noções por meio de uma investigação empírica no âmbito da Sociolinguística Histórica (Romaine, 2009), tomando como objeto a variação de marcas de 2P/tu e 3P/você na posição sintática de sujeito, quer seja o pronome, quer seja o traço morfológico do verbo, em cartas publicadas na revista *A Cigarra*, direcionada a um público feminino, contrapostas a um grupo controle[4] (as cartas da *Revista da Semana* e do jornal *Correio Paulistano*), no início do século XX. Ao nos afastarmos de uma visão essencialista do gênero feminino, mostramos como as construções sócio-históricas especificamente em torno do papel social atribuído à mulher, em determinados contextos sociais, podem atuar nos processos de variação e mudança, ou seja, podem levar a um certo resultado quanto aos usos linguísticos. Ao articular fundamentos teóricos interdisciplinares à metodologia variacionista, propomos um avanço no

[3] Como propõe Tosald (2010, p. 169), essencialismo é a "necessidade de estabilizar determinados grupos sociais enquanto sujeitos políticos". Numa visão essencialista, as identidades que compõem os sujeitos são estanques.

[4] Um grupo controle, nesta pesquisa, é um córpus (ou um conjunto de córpus) de comparação, de referência, permitindo-nos confirmar ou não nossas hipóteses.

debate sobre gênero social e a ampliação das possibilidades de interpretação de resultados estatísticos na Sociolinguística brasileira.

O artigo está estruturado da seguinte maneira: na seção 2, destacamos como a variável gênero tem sido explorada nos estudos sociolinguísticos, evidenciando limitações e desafios teóricos. Na seção 3, apresentamos as contribuições das teorias da performatividade (Butler, 2018a, 2018b) e da tecnologia de gênero (Lauretis, 1994) para uma abordagem interdisciplinar de gênero em análises sociolinguísticas; em 3.1, especificamente, em diálogo com a discussão anterior, trazemos à cena a imprensa feminina enquanto tecnologia de gênero e espaço de performatividade. Já na seção 4, dedicamo-nos a mostrar a aplicabilidade desse modelo teórico-metodológico em uma pesquisa sociolinguística histórica. Por fim, na seção 5, sintetizamos as principais considerações deste trabalho.

2. Gênero na Sociolinguística Variacionista

Pioneiro a testar o condicionamento da categoria social gênero nos processos de variação linguística, Labov (1990, 2001) formula o Paradoxo do gênero com base em observações de fenômenos fonológicos. Segundo o autor, em comparação aos homens, as mulheres tendem a adotar com maior frequência variantes de prestígio em mudanças "vindas de cima" (acima do nível de consciência social), incorporadas pela classe socioeconômica dominante, ao mesmo tempo que favorecem o uso de variantes inovadoras em mudanças "vindas de baixo" (abaixo do nível de consciência social).

Nas "mudanças vindas de cima", Labov (2001) associa o comportamento linguístico das mulheres à sua menor inserção econômica, argumentando que elas recorreriam à língua como estratégia de gerar avaliações sociais positivas associadas a classes econômicas mais altas. Em sua interpretação, o uso de formas mais prestigiosas refletiria uma condição de vulnerabilidade socioeconômica e insegurança psicológica e social. Além disso, essa escolha estaria vinculada ao papel social da mulher como "mãe educadora", responsável pela transmissão de capital simbólico e pela mobilidade social dos filhos.

Já nas mudanças "vindas de baixo", as mulheres tendem a ser, numa análise em busca de padrões gerais, as principais impulsionadoras da mudança linguística (Labov, 2001). Essa liderança, evidenciada em seu estudo na Filadélfia, é explicada pela posição central que ocupavam em suas redes sociais (Milroy, 1980), que não se reduziam às localidades em que viviam. Tal contexto lhes conferiria, de acordo com Labov (2001), maior influência entre parentes e amigos, favorecendo a difusão de novas variantes na comunidade de fala.

As hipóteses labovianas, ao introduzirem o gênero como categoria analítica relevante e ao demonstrarem empiricamente sua correlação com padrões de variação e mudança linguísticas, passaram a ser incorporadas largamente nas pesquisas sociolinguísticas, trazendo contribuições valiosas para a explicação do encaixamento social da mudança. Todavia, a adoção dessas hipóteses como um pressuposto válido universalmente, ignorando ou minimizando diferenças sociais, históricas e contextuais entre os sujeitos, constitui um problema que buscamos endereçar neste texto. Inclusive, porque esse movimento pode levar ao viés da confirmação, que, como define Nickerson (1998),

consiste na tendência de buscar, selecionar ou interpretar resultados de modo a corroborar expectativas previamente estabelecidas.

No âmbito da Sociolinguística, isso significa que a análise dos dados pode ser conduzida para se alinhar às tendências previstas por Labov (1990, 2001), em vez de permitir que os próprios dados revelem padrões que eventualmente desafiem ou ampliem tais hipóteses. Esse risco é apontado por Oushiro (2020), que observa que a correlação entre gênero e uso de variantes avaliadas como prestigiosas, frequentemente documentada na literatura, em estudos como os de Fasold (1990) e Trudgill (1972), pode passar a ser tomada como expectativa de resultado em novas pesquisas, quando deveria permanecer como uma questão a ser observada e testada.

Os questionamentos de Freitag (2015, p. 38) também ilustram esse problema: "mulheres usam determinadas variáveis porque elas têm prestígio ou as variáveis têm prestígio porque as mulheres usam? O fato de constatar que as mulheres aderem a uma variante implica em generalizar que a variante tem prestígio?". Ou seja, quando a interpretação se ancora na hipótese laboviana clássica, corre-se o risco de naturalizar uma relação causal não comprovada: considerar que o emprego de uma variante por mulheres basta para qualificá-la como prestigiada ou, inversamente, assumir que uma variante é prestigiosa e, por isso, é mais utilizada por mulheres. Em ambos os casos, a correlação entre gênero e prestígio deixa de ser empiricamente verificada e passa a ser deduzida a partir da expectativa teórica, inclusive baseada em estereótipos sociais.

Vejamos que o próprio Labov (2016 [1972], p. 21) adverte:

> [...] não se pode entender o desenvolvimento de uma mudança linguística sem levar em conta a vida social da comunidade em que ela ocorre. Ou, dizendo de outro modo, as pressões sociais estão operando continuamente sobre a língua, não de algum ponto remoto do passado, mas como uma força social imanente agindo no presente vivo.

No Brasil, a consolidação de bancos de dados sociolinguísticos de "primeira onda" (Eckert, 2012)[5] colabora, em nosso entendimento, para a adoção desse tipo de abordagem. Ao ser registrado em fichas sociolinguísticas como um atributo demográfico estável, definido no momento da coleta, o gênero passa a integrar o perfil do sujeito como uma informação dada, raramente acompanhada de uma discussão sobre suas configurações sociais e históricas. Nas palavras de Freitag (2015, p. 46), por ser *default*, "controlar a variável sexo/gênero é uma praxe; tendo ou não uma hipótese do seu efeito sobre o fenômeno, não há custo operacional em incluí-la na análise, na medida que a categorização já vem pronta, sem requerer reflexões analíticas do pesquisador".

---

[5] As ondas sociolinguísticas, tal como prevê Eckert (2012), representam tendências distintas (e coexistentes) de análise do contexto social da variação linguística. Convém ressaltar que, apesar de os estudos de segunda e terceira ondas darem ênfase ao engajamento social nas práticas de interação, verificando o gênero como categoria social situada (Eckert, 2012), a Sociolinguística brasileira ainda se desenvolve majoritariamente a partir da abordagem de primeira onda (Freitag; Martin; Tavares, 2012). Para Eckert (2012), os trabalhos de primeira onda definiram a base metodológica do campo, relacionando variáveis linguísticas a categorias sociais, como classe, gênero e idade, e analisando o estilo de fala segundo traços de prestígio ou estigma.

Nessas circunstâncias, muitas vezes, as noções de "mulher" e de "homem" são pressupostas em construções de amostras estratificadas voltadas à identificação de padrões gerais, como advertem Eckert e McConnell-Ginet (2010). Talvez seja por isso que a ideia de "mulher universal", desprovida de raça, classe e sexualidade, por exemplo, tão rebatida por teóricas feministas (Crenshaw, 1991; Davis, 2016; hooks, 2020), faz-se presente nos trabalhos que se debruçam sobre fenômenos variáveis, em especial quando levantam conclusões generalizantes, como a de que as mulheres aderem mais a variantes de prestígio do que os homens por serem vistas como modelo linguístico a se seguir, tal como formulado por Labov (1990, 2001).

A concepção homogeneizadora da mulher como sujeito social, relegada a papéis tradicionais (de cuidado, domesticidade, docilidade etc.), reforça uma visão essencialista de gênero e opera no sentido de apagar as diferenças e hierarquias que estruturam as experiências dos indivíduos em contextos sociolinguísticos distintos (Eckert; McConnell-Ginet, 2010). É por essa razão que, segundo Barbosa (2025), quaisquer interpretações sobre o papel da mulher nos processos de variação e mudança linguística que não levem em conta a triangulação dos fatores gênero, raça e escolaridade estarão sempre tendendo a visões limitantes do comportamento social e linguístico das mulheres. Dessa maneira, concordamos com Levon e Mendes (2016) de que as análises da linguagem devam se beneficiar de uma perspectiva interseccional, na qual gênero seja considerado em articulação com outras dimensões da vida social, permitindo uma compreensão mais situada dos significados linguísticos.

Em estudos sociolinguísticos históricos (*cf*. Romaine, 2009), voltados, portanto, ao passado da língua, o desafio em controlar e interpretar o condicionamento dessa variável é ainda maior, uma vez que os textos escritos disponíveis dificilmente fornecem informações seguras acerca da autoria e, por conseguinte, dos perfis sociais dos informantes. Além disso, como apontam Auer *et al*. (2015), os textos produzidos em sincronias pretéritas são, em geral, atribuídos a homens brancos, letrados, pertencentes às classes sociais mais altas.

Como tentativa de superar esses desafios, propomos um alinhamento metodológico que visa à incorporação da teoria da performatividade, desenvolvida por Butler (2018a, 2018b) e incorporada ao debate sociolinguístico por Levon e Mendes (2016), e da tecnologia de gênero (Lauretis, 1994), apresentadas na seção seguinte, como instrumentos capazes de viabilizar uma análise mais profunda e socialmente sensível da mudança linguística.

3. Contribuições de teorias sociais para análise de gênero na Sociolinguística

A partir das discussões apresentadas, trazemos à luz, neste artigo, uma reorientação teórico-metodológica, fundamentada nos estudos de gênero. O objetivo é evitar interpretações essencialistas dos padrões linguísticos, adotando uma perspectiva que compreenda o gênero como processo socialmente situado. Para tanto, tomamos a imprensa voltada ao público feminino como córpus de investigação e a concebemos como tecnologia de gênero (Lauretis, 1994) e espaço de performatividade (Butler, 2018a, 2018b).

Para Butler (2018a), gênero não é uma essência interior nem uma identidade fixa do sujeito, mas uma realidade social fabricada por meio da repetição estilizada de atos, gestos e discursos. Gênero, em sua visão, é uma performance[6], uma reiteração de normas sociais impostas por práticas regulatórias de poder, ou seja, é o que se faz, não o que se é. Logo, a performance é moldada por certos significados e por certas expectativas sociais quanto a ser mulher/homem.

Valendo-se de diferentes reflexões (teatrais, filosóficas, antropológicas e fenomenológicas), Butler (2018b, p. 160) afirma:

> Como em outros dramas sociais rituais, a ação do gênero requer uma performance repetida. Essa repetição é a um só tempo reencenação e nova experiência de um conjunto de significados já estabelecidos socialmente; e também é a forma mundana e ritualizada de sua legitimação. Embora existam corpos individuais que encenam essas significações estilizando-se em formas do gênero, essa 'ação' é uma ação pública. Essas ações têm dimensões temporais e coletivas, e seu caráter público não deixa de ter consequências; na verdade, a performance é realizada com o objetivo estratégico de manter o gênero em sua estrutura binária – um objetivo que não pode ser atribuído a um sujeito, devendo, ao invés disso, ser compreendido como fundador e consolidador do sujeito.

A performatividade em Butler (2018b), dessa maneira, está relacionada à ideia de uma encenação coletiva, como no teatro, tendo em conta os papéis binários e heteronormativos de mulher/homem sócio-historicamente definidos. Nesse contexto, a noção de público (a audiência à qual o sujeito se dirige ao realizar suas performances de gênero) é central no pensamento butleriano. Nessa perspectiva, o(s) "outro(s)" integra(m) o imaginário social, uma vez que atitudes e comportamentos são identificados, mediante modelos normativos, como femininos ou masculinos. Em razão disso, a performance de gênero se constrói em consonância com expectativas sociais que determinam como devemos agir e nos enquadrar em um sistema binário.

Isso significa que, quando declaramos "somos mulheres", estamos assumindo um conjunto de condutas alinhadas a modelos culturalmente prescritos de feminilidade. De forma semelhante, Cyfer (2015), também ancorada na teoria butleriana, observa que o reconhecimento social de alguém como mulher ou homem é uma "condição alcançada pela manifestação pública de comportamentos associados à masculinidade e à feminilidade" (Cyfer, 2015, p. 46). Portanto, não existe uma essência natural que defina o sujeito. São os atos performativos que constroem a identidade de gênero e produzem a ilusão de um pertencimento a uma categoria.

Nessa lógica, a linguagem emerge como um dos elementos para a materialização do gênero, como uma instância produtora das identidades dos sujeitos. Na interface com a Sociolinguística, tal perspectiva revela que as escolhas linguísticas funcionam como atos performativos e recursos estilísticos pelos quais os falantes negociam, constroem e manifestam suas posições de gênero de maneira situada e interacional. As diferentes maneiras de falar e escrever tornam-se, então, estratégias

[6] Cumpre distinguir performatividade e performance na teoria butleriana: ao passo que a performance pode ser um ato individual, a performatividade é um ato coletivo.

discursivas que participam da encenação e da negociação contínua do gênero, não estando rigidamente associadas a "mulheres" ou "homens", uma vez que são guiadas pelos modos de agir e significar socialmente (Levon; Mendes, 2016; Podesva, 2007, 2011).

Tal entendimento, por um lado, pode ampliar o alcance da análise sociolinguística, pois desloca o foco das categorias fixas para os processos de construção social do gênero, permitindo examinar como valores, normas e representações de feminilidade e masculinidade são inscritos e reinscritos nas práticas discursivas, inclusive na escrita de diferentes épocas e meios de circulação. Por outro, a tradição quantitativa na Sociolinguística depende, em grande parte, da estabilização de categorias sociais. Nosso desafio é, portanto, unir abordagens quanti e qualitativas em um conjunto extenso de dados, de modo a identificar padrões linguísticos generalizáveis e interpretar os significados sociais presentes nos dados. O problema central reside, dessa maneira, em desenvolver metodologias que possibilitem identificar regularidades linguísticas em córpus de fala/escrita sem reduzir gênero a uma variável fixa.

Por isso, é útil trazer, em articulação à teoria da performatividade, a noção de tecnologia de gênero, proposta por Lauretis (1994). Conforme a autora, gênero é uma construção que se dá por intermédio de diferentes mecanismos, como as tecnologias sociais (revistas e jornais, rádio, televisão etc.), "e de discursos, epistemologias e práticas críticas institucionalizadas, bem como das práticas da vida cotidiana" (Lauretis, 1994, p. 209). Em outros dizeres, os meios de comunicação e os discursos sociais são dispositivos que constroem, normalizam e naturalizam o que é ser "mulher" e "homem" em um dado contexto.

Lauretis (1994, p. 220) traz um exemplo:

> a partir do momento em que assinalamos o F num formulário, ingressamos oficialmente no sistema sexo-gênero, nas relações sociais de gênero, e fomos 'engendradas' como mulheres, isto é, não são apenas os outros que nos consideram do sexo feminino, mas a partir daquele momento, nós passamos a nos representar como mulheres.

À medida que a noção de performatividade revela que o gênero se constitui na repetição de atos, o conceito de tecnologias de gênero demonstra que esses atos são sempre orientados por aparatos culturais mais amplos, como é o caso da imprensa feminina, explorada mais adiante (*cf.* subseção 3.1). Sob esse olhar, convém destacar que a escrita também pode ser compreendida como prática performativa, atravessada por convenções sociais que regulam modos de dizer e de representar o feminino, por exemplo. Em linha com Cixous (2022), escrever é inscrever-se na linguagem e, nesse movimento, expor as marcas do corpo, da experiência e da diferença. Pensar a escrita em contextos de produção feminina, portanto, é refletir sobre como certas formas de dizer se repetem, se consolidam e tensionam as normas de gênero que as produzem.

Nesse sentido, as teorias aqui discutidas, se atreladas em um estudo sociolinguístico, podem oferecer um caminho metodológico: a depender do córpus e dos objetivos de pesquisa, a exemplo de Clempi (2019, 2024) e Rodrigues, Clempi e Félix (2023), é possível analisar quanti e qualitativamente

os padrões sociolinguísticos que caracterizam performatividades generificadas, sem atribuí-los a identidades fixas dos falantes/escreventes, mas aos papéis e modelos que esses textos constroem e disseminam.

3.1. Imprensa feminina enquanto tecnologia de gênero e espaço de performatividade

Ao lançar mão de um córpus composto por cartas publicadas em revistas femininas nos anos novecentos, lidamos com as questões sobre a autoria e sobre a própria realidade sócio-histórica das mulheres em relação à escrita. Isso porque, ao longo da história, o ato de escrever foi atravessado por restrições simbólicas e institucionais que afastavam as mulheres dos espaços de produção textual e intelectual.

Entre os séculos XIX e XX, de acordo com Heller (2001) e Dumont e Santo (2007), a educação feminina, ao ser acessada por camadas mais altas da sociedade brasileira, voltava-se à formação moral e doméstica. Nesse cenário, a escrita e a leitura eram consideradas inadequadas ou mesmo perigosas quando associadas à autonomia e à reflexão crítica. Diante das limitações impostas, muitas autoras recorriam a pseudônimos masculinos para garantir circulação e legitimidade a seus textos.

Assim, assumimos que a imprensa feminina se configura como um dispositivo para a produção e difusão de modelos de feminilidade, operando como uma tecnologia de gênero. Consolidada por meio de revistas, mais do que um simples veículo de informação, a imprensa feminina institui um ideal feminino, com certas formas de pensar e agir, articulando e acionando discursos sobre comportamento, moralidade e domesticidade. Ao mesmo tempo, essas publicações também atuavam na regulação das práticas de leitura e escrita, delimitando os temas considerados apropriados às mulheres e os modos pelos quais elas poderiam se expressar.

Nessa orientação, em alinhamento com Bassanezi (1996, p. 15), compreendemos que:

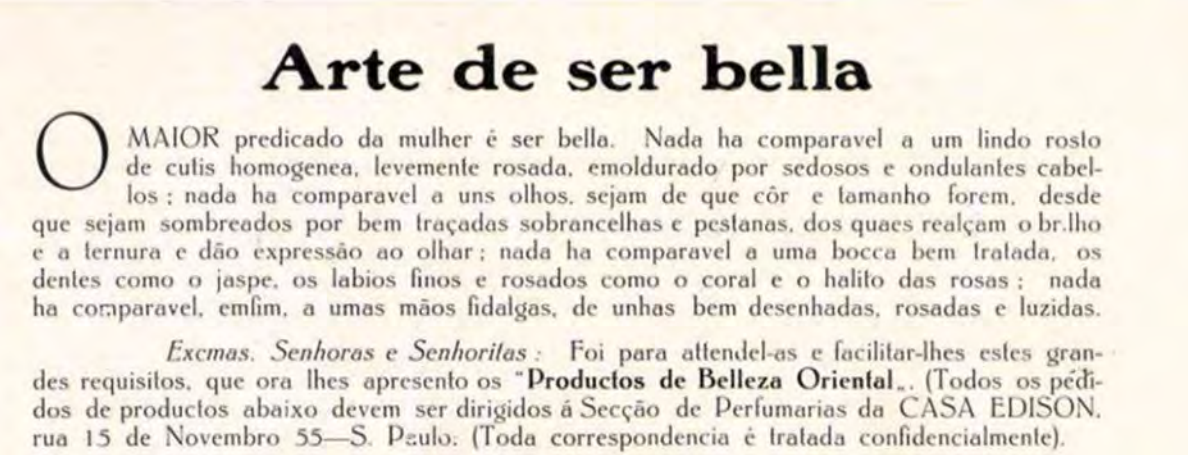
> As revistas buscam, na medida do possível, refletir um aparente consenso social, ou melhor, as ideias dominantes sobre o masculino, o feminino e as relações homem-mulher. Assim, elas também são espaços onde as relações sociais, com sua distribuição de poder, se estabelecem, são produzidas, reforçadas e reproduzidas.

Retomando o período de sua popularização no Brasil, nas primeiras décadas do século XX, evidenciamos que essa construção do ideal feminino pela imprensa especializada atuava como um mecanismo de regulação social. Longe de promover emancipação, essas publicações reforçavam os limites do espaço feminino, restringindo-o a uma esfera doméstica idealizada. O discurso da modernidade, quando presente, era instrumentalizado para atualizar e perpetuar estruturas patriarcais[7], definindo os termos aceitáveis para a atuação da mulher na sociedade. Dessa forma, a imprensa feminina alinhava os valores hegemônicos da época mediante repetição performática de normas de gênero em textos, imagens e anúncios, por exemplo.

[7] Entendemos patriarcado, da mesma maneira que Pedro (2005, p. 79), como "o poder dos homens na subordinação das mulheres".

É nesse contexto que se insere *A Cigarra*, uma revista que circulou entre os anos de 1914 e 1975, direcionada a um público consumidor feminino, por hipótese branco, jovem-adulto, culto, letrado, de camadas médias a altas e situado, de início (1914-1932), no estado e na cidade de São Paulo. Fundado por Gelásio Pimenta, o periódico quinzenal foi pensado, paradoxalmente, por homens. Não por acaso, de acordo com Moreira (2011, p. 94), havia "uma tendência particular da própria *A Cigarra* em destacar assuntos mais suaves que se 'adequavam' às mulheres". O anúncio adiante, em que o papel de mulher é atrelado à beleza, valida a questão:

Imagem 1: Anúncio publicado na revista *A Cigarra* (ano de 1915)



Arte de ser bella

O MAIOR predicado da mulher é ser bella. Nada ha comparavel a um lindo rosto de cutis homogenea, levemente rosada, emoldurado por sedosos e ondulantes cabellos; nada ha comparavel a uns olhos, sejam de que côr e tamanho forem, desde que sejam sombreados por bem traçadas sobrancelhas e pestanas, dos quaes realçam o br.lho e a ternura e dão expressão ao olhar; nada ha comparavel a uma bocca bem tratada, os dentes como o jaspe, os labios finos e rosados como o coral e o halito das rosas; nada ha comparavel, emfim, a umas mãos fidalgas, de unhas bem desenhadas, rosadas e luzidas.

Excmas. Senhoras e Senhoritas: Foi para attendel-as e facilitar-lhes estes grandes requisitos, que ora lhes apresento os "**Productos de Belleza Oriental**". (Todos os pédidos de productos abaixo devem ser dirigidos á Secção de Perfumarias da CASA EDISON, rua 15 de Novembro 55—S. Paulo. (Toda correspondencia é tratada confidencialmente).

Fonte: Arquivo Público do Estado de São Paulo ([2025]). Revista *A Cigarra*, edição 23 (ano de 1915).

Sob a perspectiva de Butler (2018a, 2018b), as páginas d'*A Cigarra* podem ser vistas como um espaço de encenação coletiva e repetida de feminilidade. A seção destinada a cartas, em particular, funcionava como um palco no qual se dramatizavam dilemas pessoais e se expressava intimidade, sempre mediada pelo olhar editorial. Além disso, a partir da noção de tecnologia de gênero, a revista se revela como um aparato que oferece o roteiro e o cenário para a performatividade de gênero. Ela não só retratava "a mulher": construía o "feminino" como uma categoria reconhecível, desejável e regulada por meio da repetição de histórias, conselhos e modelos de interação que orientavam comportamentos e modos de ser.

O excerto abaixo, extraído do início da seção de cartas *Collaboração das leitoras*, escrito pelo redator, ilustra qual era a perspectiva d'*A Cigarra*:

(1) É com immenso prazer que registramos o successo sempre crescente desta secção d'<*A Cigarra*>, destinada ás nossas gentilissimas leitoras, para que possam manifestar-se sobre **assumptos de nossa vida social**. Augmentou consideravelmente o numero de cartas enviadas á nossa redacção e escriptas por **delicadas mãos femininas**. Na impossibilidade de publical-as todas, daremos apenas algumas neste numero (*A Cigarra*, outubro de 1915, grifos nossos)[8].

[8] Trecho transcrito como no original. Dados disponíveis no Arquivo Público do Estado de São Paulo ([2025]) e na Biblioteca Nacional Digital ([2025]).

Na introdução da seção, a revista demarca bem o espaço das mulheres: a elas competia tratar de aspectos circunscritos à vida social com suas "delicadas mãos femininas", o que deixa implícito o papel de fragilidade relegado às leitoras. As cartas, nesse viés, promoviam um espaço para as mulheres poderem escrever, desde que fosse a respeito de determinados assuntos, ligados a um certo papel feminino e a um ideal de feminilidade, como exemplificamos em (2):

(2) Eis-me novamente aqui, 'Cigarrinha' gentil, deixando na gaze finissima das tuas asas seductoras os affectuos beijos que te prometti. [...]. Não te esqueças nunca querida, que eu sou tua amiga e que te adora immensamente; queres tambem ser minha amiguinha? [...]. Pois bem, então escuta-me agora, que eu te direi algumas cousas que sei de Mr. D. I. Tu o conheces sem duvida, não? [...] O jovem de quem falo é uma dessas creaturas para quem a vida não é senão um sorriso, um desabrochar perenne de flores. [...]. Já sabes quem é, 'Cigarrinha'? Eu vou terminar porque já fui bastante cacete, não é? Espero que não te esquecerás de mim e que acceitarás o convite que te faço de ires á missa das 9 no Coração de Maria. Lá encontrarás o meu perfilado e a tua sempre amiguinha dedicada e agradecida. Cruz Vermelha" (*A Cigarr*a, agosto de 1917).

Observamos, com o exemplo, que o segmento era composto por sugestões, comentários e observações envolvendo aspectos da sociedade da época. No trecho (2), aparece como remetente a revista *A Cigarra*, considerada uma amiga das leitoras. Nesse contexto amigável, era comum o uso de diminutivos ("Cigarrinha", "amiguinha") para fazer referência ao periódico. Verificamos também a presença de pseudônimos (Cruz Vermelha) e iniciais (D. I.), recursos comumente empregados nesses textos. A esse respeito, de acordo com Moreira (2011, p. 94), "ao contrário da literatura, em que mulheres eram obrigadas a assinar nomes masculinos para publicar seus escritos, na imprensa feminina foram os homens que tiveram que inventar pseudônimos com nomes de mulher".

A *Collaboração das leitoras* permaneceu ativa até 1930 e era um sucesso entre as consumidoras da revista, tendo em vista que ocupava muitas páginas (15-35 de 50-100 do total), o que a tornava próxima de um suplemento. De acordo com Besse (1999, p. 31), "essa seção proporcionava um foro aberto mas anônimo para o debate sobre os costumes modernos e o comportamento adequado às mulheres e servia como uma coluna de mexericos e assuntos pessoais". O pressuposto aqui era de que as leitoras escreviam tanto para a revista quanto para outras leitoras, ou mesmo para outros destinatários (amorosos ou não), o que poderia ampliar as possibilidades de interação propostas.

Fatores sócio-históricos, como acesso à educação e classe social, nos fazem deduzir que as publicações não eram escritas necessariamente por mulheres, mas representavam e reproduziam o que a sociedade interpretava como sendo (de) mulher(es), o que pode ser útil para uma análise sociolinguística que se volte às questões de gênero (feminino). Nesse sentido, podemos afirmar que a autoria desses textos se torna não central do ponto de vista metodológico.

Em outras palavras, o que defendemos é que a autoria de cada carta, nesta pesquisa, não precisa ser atestada nem controlada, como em uma estratificação sociolinguística tradicional, tendo em vista que o que está em jogo é uma escrita na qual há performatividade de um "feminino". Tal escrita, devemos destacar, não corresponde a um retrato fiel da realidade das diferentes mulheres da época.

Trata-se de uma representação normativa do real, um modelo discursivo de feminilidade que a revista, em seu conjunto, enquanto tecnologia de gênero, produzia e disseminava.

Mesmo que represente uma determinada mulher (branca, de classes socioeconômicas mais altas), o trabalho com o texto escrito, em particular o relacionado a esse tipo de material, para a formação de um córpus de pesquisa, pode favorecer uma discussão sobre a construção social do gênero feminino em dados linguísticos.

4. Gênero feminino, imprensa feminina e mudança linguística: uma análise da variação entre tu/você

Visamos a demonstrar, nesta seção, a aplicabilidade da abordagem teórico-metodológica sobre a qual discorremos até aqui. Propomos explorar, por meio de evidências de textos pretéritos, como processos de variação e mudança linguísticas podem ser sensíveis à construção social de papéis associados ao gênero feminino. Para isso, investigamos as marcas (quer seja o pronome, quer seja o traço morfológico do verbo) de segunda pessoa do singular (2P/tu)[9] e de terceira pessoa do singular (3P/você) na posição sintática de sujeito, no início do século XX (anos de 1915 a 1932).

O recorte temporal se justifica por ser esse um período em que a variante 'você' parece competir mais assiduamente com 'tu', variante mais antiga na língua (Lopes; Duarte, 2007; Pereira, 2012; Rumeu, 2008; Souza, 2012). Do ponto de vista sociopragmático, 'tu' tendia a indexar relações marcadas por proximidade, familiaridade e envolvimento entre os participantes da interação. Já 'você', resultante da gramaticalização de 'Vossa Mercê', preservava traços de indiretividade e distanciamento herdados da forma-fonte, sendo, nos termos de Koch (2008), uma estratégia menos invasiva para se referir a um interlocutor. Esses valores, socialmente construídos a partir de percursos históricos distintos, embora tenham se dissipado atualmente, encontravam-se em circulação no período em análise (Faraco, 2017; Lopes *et al*., 2018; Rumeu, 2008)[10].

O córpus principal de nossa pesquisa é constituído por 432 cartas publicadas na revista *A Cigarra*, direcionada ao público feminino. Tal como discutimos anteriormente, esse periódico pode ser entendido como uma tecnologia de gênero, que produzia um modelo de feminilidade, performando um tipo de mulher mediante um discurso de intimidade e familiaridade, o que torna o material produtivo à análise do fenômeno variável em questão. Com isso, buscamos analisar o gênero feminino de maneira a contemplar o conjunto-revista, sem nos ater, portanto, ao sujeito escrevente de forma individualizada.

Os resultados obtidos foram confrontados com aqueles provenientes de um grupo controle, composto, do mesmo modo, por 432 cartas distribuídas em duas diferentes amostras. Diferentemente da revista *A Cigarra*, cuja segmentação voltada ao público feminino é mais explícita, como demostramos

[9] Nos córpus, não foram atestados usos de 'tu' com morfologia verbal de terceira pessoa.

[10] A literatura indica que 'você', derivado de 'Vossa Mercê', perde progressivamente seu valor cerimonioso e passa a concorrer com 'tu' nos mesmos domínios sociofuncionais, sobretudo a partir da década de 1930 (Faraco, 2017; Lopes *et al*., 2018; Rumeu, 2008).

até aqui, os periódicos selecionados não apresentam uma definição de público baseada em gênero. Todavia, na prática, podem ter um direcionamento maior dos conteúdos, ora para mulheres, ora para homens, como nos sugerem capas, textos, anúncios e vendas dos periódicos, além de estudos sobre a história de revistas e jornais do século XX de maneira geral (Casadei, 2017; Czrnorski; Meyrer, 2016; Mira, 2001; Scalzo, 2016).

Por esse motivo, o grupo controle foi organizado a partir de um *continuum* de práticas sociais e de representações de gênero, levando em conta o direcionamento predominante de seus conteúdos. Nesse *continuum*, a *Revista da Semana* é situada como publicação [+feminina ↔ −masculina], enquanto o *Correio Paulistano* é caracterizado como publicação [+masculina ↔ −feminina][11].

Transcritos na forma original, os trechos a seguir ilustram o fenômeno linguístico e os córpus em análise[12]:

(i) *A Cigarra* (publicação feminina):

(3) [...]. ah! 'Cigarrinha', não **imaginas** o meu constante soffrer com essa ingratidão! Se **tu** visses a magoa do meu coração, talvez **virias** enxugar as minhas lagrimas. Julgo-me a mais infeliz entre as moças. A unica consolação que me resta é fazer de ti a minha confidente, [...] (*A Cigarra*, novembro de 1917).

(4) Sou de facto muito animada. Diz **você** que se sentiu lisonjeada com a leitura de minha cartinha, pois, confessando-me ao Alcebiades, usei das mesmas expressões que **você** empregou numa carta dirigida ao poeta Guilherme de Almeida [...] (*A Cigarra*, março de 1922).

(ii) Grupo controle de publicação [+feminina ↔ -masculina] *Revista da Semana*:

(5) [...] Ah! como **tu** foste cruel, não me dissimulando o terror que se apossou do teu espirito, deixando-me entrevêr as maiores desgraças e accusando-me de negligencia, como se tivesse sido eu, só eu, a culpada da perda das tuas cartas! [...]. Como **tu** foste egoista! [...] Já tenho as tuas cartas. [...] (*Revista da Semana*, julho de 1915).

(6) [...] Vocês gostariam de ver o viveiro que fizemos no pomar: canarios, pintassilgos, sonhaços, arapongas... são tantos os passarinhos! **Você** bem sabe como é o Mario: de quando em quando muda de mania. Agora já não quer saber da collecção de sellos [...] (*Revista da Semana*, julho de 1931).

(iii) Grupo controle de publicação [+masculina ↔ -feminina] *Correio Paulistano*:

(7) [...] Visses **tu** esse fragmento de pelle exposto á voracidade dos meus olhos allucinados... Perdi o negocio do feijão, porquanto, preso ao engôdo desse escandaloso e divino decote, não desci, como projectára na praça Antonio Prado. Segui, preso áquelle trecho sublime de paizegem humana, mudo, fascinado, extatico, cataleptico... [...] (*Correio Paulistano*, abril de 1920).

[11] Ambos os materiais estão disponíveis em formato digitalizado no site da Biblioteca Nacional Digital ([2025]).

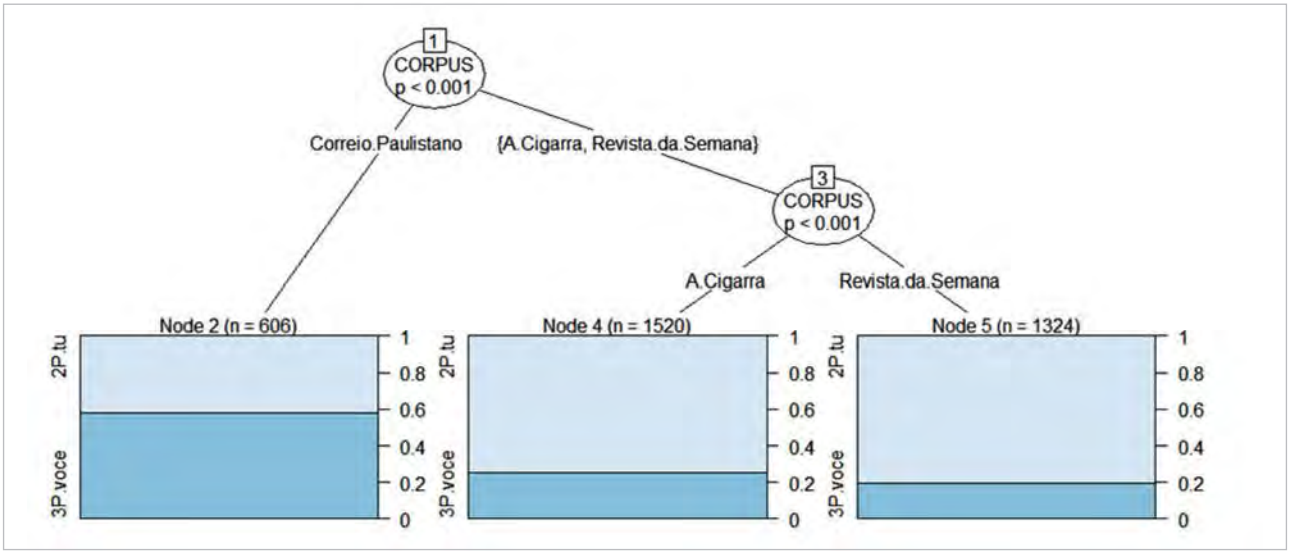
[12] A apresentação dos exemplos obedece ao seguinte critério tipográfico: em construções de sujeito pleno, a variável em estudo (tu/você) é grafada em negrito e seu verbo é sublinhado. Nos contextos de sujeito nulo, o verbo flexionado, que carrega o traço morfológico de pessoa, recebe negrito e é sublinhado.

(8) Si **você**, depois, quizer detalhes sobre a doutrina e processos dessa grey, poderei, com documentação farta, evidenciar a absoluta verdade do que allego. Deante de taes attitudes, comprehendo que, por obra de patroticas devoções civicas, o Brasil entra numa phase de disciplina rigorosa e reconstrucção nacional, [...]. Após 7 annos de pesquisas, é tempo de se entrar numa phase de reconstrucção sadia e fecunda [...] (*Correio Paulistano*, maio de 1929).

No intento de verificar padrões indicativos de mudança linguística na direção de 3P/você pela perspectiva sociolinguística, os dados foram submetidos a uma análise quantitativa, com o tratamento estatístico, após coleta e codificação, feito com o auxílio da linguagem de programação R (R Core Team, 2023).

Apresentamos, abaixo, os resultados gerais[13], gerados a partir de uma análise inferencial, em relação à distribuição de uso da variável resposta (marcas de 2P/tu e 3P/você) quanto aos córpus:

Figura 1: Árvore de inferência condicional para a variável resposta quanto aos córpus



Fonte: Elaborada pelas autoras.

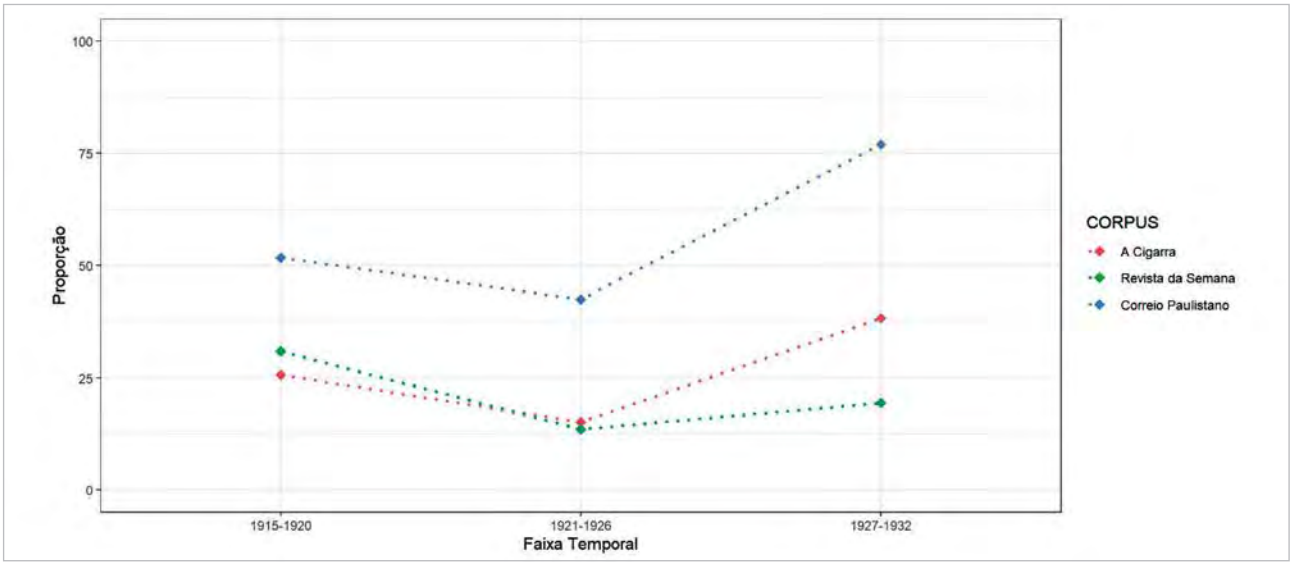
Com a figura 1, podemos visualizar os padrões de uso das variantes da variável resposta nos diferentes córpus. O primeiro nó da árvore apresenta duas ramificações para os córpus: a primeira, posicionada à esquerda, é referente aos dados do *Correio Paulistano* (grupo controle de publicação [+masculina ↔ -feminina]), e a segunda, posicionada ao centro e à direita da figura, diz respeito aos dados d'*A Cigarra* (publicação feminina) e da *Revista da Semana* (grupo controle de publicação [+feminina ↔ -masculina]). As divisões em "galhos" também indicam valores de $p < 0,05$, obtidos

[13] Dado o objetivo deste artigo, nosso foco, nesta seção, é apresentar apenas resultados gerais. Contudo, convém salientar que, na pesquisa, fizemos testes de regressão logística considerando (i) as variáveis respostas: marcas de 2P/tu e marcas de 3P/você; (ii) as variáveis previsoras: tipo de sujeito, contexto verbal, paralelismo linguístico, tipo de comunicação, temática, papel social do referente e faixa temporal; e (iii) a variável aleatória: identificação de autoria. Além de discutir os resultados das análises multivariadas em termos estatísticos, também olhamos para os dados de um ponto de vista qualitativo (Clempi, 2024), tendo em conta a pertinência do contexto para o entendimento do fenômeno.

pelo conjunto de variáveis (nesse caso, a variável resposta + a variável previsora córpus) inseridas no modelo. De modo geral, do ponto de vista estatístico, os resultados indicam que o jornal tem uma distribuição das variantes significativamente diferente em comparação a ambas as revistas.

Enquanto nas cartas d'*A Cigarra* e da *Revista da Semana* o uso de 3P/você corresponde, respectivamente, a 26% (N = 388/1520) e 20% (N = 262/1324), no *Correio Paulistano* a mesma variante atinge 58% de ocorrências (N = 352/606). Ao verificar a faixa temporal[14], confirmamos que o *Correio Paulistano* tende a avançar mais rápido em direção a 3P/você:

Gráfico 1: Proporção de uso de 3P/você quanto à faixa temporal em *A Cigarra* e grupo controle



Fonte: Elaborado pelas autoras.

Esses achados colocam-se, à primeira vista, em desacordo com estudos sociolinguísticos anteriores (Rumeu, 2008; Pereira, 2012; Souza, 2012), os quais, apoiados no Paradoxo do gênero formulado por Labov (1990, 2001), associam as mulheres à liderança na mudança em favor de 'você' com base na análise de cartas pessoais. Tal diferença, no entanto, torna-se compreensível quando consideramos as condições específicas de produção e circulação dos córpus examinados nesses trabalhos. Em Rumeu (2008), Pereira (2012) e Souza (2012), as amostras são compostas por cartas pessoais em âmbito privado, produzidas por sujeitos-autores identificáveis, pertencentes a redes familiares e afetivas da elite letrada, e dirigidas a destinatários específicos no interior da esfera íntima.

Nessas condições, sugere-se que as mulheres adotariam uma variante considerada mais "neutra" ou de "melhor tom" (você) como estratégia de atenuação em uma sociedade patriarcal, interpretação explicitada por Lopes, Marcotulio e Oliveira (2018), que ligam o uso de 'você' à herança de traços de indiretividade e menor ameaça à face do interlocutor:

[14] A decisão metodológica de controlar três faixas de tempo (1915-1920; 1921-1926; e 1927-1932) se deve às limitações, naturais em um estudo sociolinguístico histórico, no que se refere à coleta de dados por ano em cada córpus, considerando o material digitalizado disponível para acesso.

> [...] a forma você teria herdado, primeiramente, o caráter atenuante da estratégia- fonte Vossa Mercê, por isso seria menos "ameaçador ao interlocutor". Empregar tu poderia não ser tão "adequado" às figuras femininas no contexto social de fins do século XIX e início do XX, uma vez que tratar o interlocutor de forma mais direta com um pronome de 2SG poderia ser um ato invasivo à face do destinatário (Brown e Levinson, 1987) (Lopes; Marcotulio; Oliveira, 2018, p. 39).

Já nesta pesquisa, voltamo-nos para cartas mediadas pela imprensa, de circulação pública, muitas vezes assinadas de forma anônima ou por pseudônimos[15], nas quais as posições de gênero se constroem performativamente diante de um público leitor ampliado. Desse modo, a correlação entre gênero feminino e mudança em favor de 'você' mostra-se condicionada pelo tipo de prática discursiva que os córpus materializam, o que permite entender os padrões observados aqui e aqueles encontrados em estudos pautados em correspondências privadas.

Portanto, nossos resultados não invalidam as conclusões dos estudos citados, levando em conta, sobretudo, que partem de outros contextos (córpus compostos por cartas pessoais trocadas entre familiares ilustres) e de correlações com outras variáveis (faixa etária, subdivisão de cartas, remetentes/destinatários e outras). Nessa lógica, as especificidades verificadas nos comportamentos linguísticos chamam a atenção para a relevância de considerar o gênero de forma situada e relacional, sensível às práticas discursivas e às condições históricas e sociais de produção dos textos.

É preciso ter em vista que *A Cigarra* não deve ser assumida como um simples repositório de "textos de mulheres", já que construía, através da escrita, *uma representação* de um ideal de feminilidade centrado em mulheres brancas, jovens-adultas, das classes alta e intermediária, letradas e situadas no estado de São Paulo. O periódico não apenas delimitava quem poderia escrever, mas também quais modos de dizer se tornavam reconhecíveis como legitimamente femininos. Assim, o espaço dedicado às cartas engendrava um contexto discursivo no qual a variante 2P/tu se mostrava particularmente adequada, na medida em que, no período em análise, tal estratégia era mobilizada em relações de maior proximidade, familiaridade e envolvimento interpessoal (Faraco, 2017; Lopes *et al.*, 2018; Rumeu, 2008). Nesse sentido, o uso de 2P/tu nas cartas da revista era constitutivo da própria performance do gênero feminino.

A *Revista da Semana*, ao compartilhar esse mesmo projeto de endereçamento a um público feminino similar, ajuda-nos a sustentar tal argumento, já que as distribuições das variantes se aproximam daquelas encontradas n'*A Cigarra*. Tendo em conta a concepção de gênero social como uma prática reiterada de atos (Butler, 2018a, 2918b) e a premissa de que a imprensa feminina funciona como tecnologia de gênero (Lauretis, 1994), as revistas favoreciam e difundiam padrões linguísticos associados a valores sociopragmáticos específicos. Em outros termos, o uso recorrente de 2P/tu parece se ligar a expectativas socialmente atribuídas ao feminino, que se estabilizavam no âmbito das práticas mediadas pela imprensa, responsável por regular modos legítimos de sentir, dizer e se posicionar linguisticamente como uma mulher.

[15] É importante relembrar que, por meio das cartas, as leitoras conversavam entre si, embora os textos fossem públicos.

Vejamos alguns exemplos:

(9) Querida amiguinha. Conheço-te e sei que **és** boa. **Dizes** que tudo passa. Sim, passou aquelle sonho que foi talvez o primeiro... Mas, Porque? Por ser subjugado por outro. **Amas** com ardor. Porque **és** tão sensivel? Não **deves** ser assim. **Sabes** que numa época como esta, em que a miseria e decadencia predominam, mostrar teu amor, é fazer que elle se proste por terra e seja espisinhado! **Sê** altiva, **sepulta** em teu coraçãosinho esse affecto que é puro em ti, **procura** até encontrar (cousa difficil!) aquele que saiba avaliar o poder desse sacrosanto amor. **Escolhe**-o bem, **estuda**-o com attenção, para depositar-lhe n'alma a tua felicidade. **Podes** ficar sciente de que os homens só merecem credito depois de um minucioso estudo. Mesmo com as multiplas experiencias que tenho tido, acho que nunca lhe devemos ter inteira confiança. Já passei por muitos enganos e hoje, o meu <Eu>, só se recorda de ingratidões. Da leitora e colleguinha - Turmalina Verde (*A Cigarra*, abril de 1921).

(10) Minha linda Boneca [...] Viver é uma de suas terriveis expressões. Eis a verdade em que **deves**, sabiamente, iniciar teus filhos. Nada, de mais fecunda forma, lhes semeará na alma sementes de solidariedade humana. [...]. E tudo isto **conseguirás** com a simples ajuda de uma pequena bibliotheca. Os livros podem ser, minha linda Boneca, um veneno mortal. É escolher os livros amparo e protecção e ensinamento; é colloca-los nas estantes do pabilhão para que as crianças os encontrem. Antes do mais: não lh'os **apontes**, não lh'os **offereças**, **deixa** que tenham a illusão de escolher, de descobrir, principalmente de pensar todos os livros mais ou menos bons; **evita**-lhes a curiosidade do mau livro [...] Beija-te tua Elena (*Revista da Semana*, novembro de 1924).

No excerto (9), extraído d'*A Cigarra*, a autora, que se identifica pelo pseudônimo Turmalina Verde, dirige-se a outra leitora como "querida amiguinha", ilustrando a prática editorial de publicar correspondência entre leitoras, comum nos anos de 1915-1932. Essa forma de endereçamento estabelece proximidade entre remetente e destinatária e, não por acaso, inscreve ambas em uma rede de sociabilidade feminina que se constrói e se legitima na própria circulação dessas cartas.

Pela troca, depreendemos que a destinatária vivenciou uma desilusão amorosa, o que orienta o propósito comunicativo da carta: aconselhar, com base na experiência própria, a busca por um novo amor, ainda que seja difícil. Embora o conselho assuma contornos normativos, a estratégia escolhida (2P/tu) contribui para mitigar a assimetria entre quem aconselha e quem é aconselhada. Inclusive, o uso de imperativos em 2P/tu ("sê", "sepulta", "escolhe-o") não rompe com a cena de intimidade construída, mas se insere nela. Em vez de instaurar uma posição de autoridade, o enunciado se ancora na intimidade e na cumplicidade.

A escrita recorre, ainda, a recursos como o diminutivo ("coraçãozinho") para atenuar o teor prescritivo do conselho, que inclui ainda a recomendação de estudar bem o futuro amado antes de depositar-lhe a confiança. Ademais, as escolhas linguísticas no trecho ("querida amiguinha", "conheço-te", "és boa", "coraçãozinho", "afeto puro", "estuda-o com atenção", "depositar-lhe na alma tua felicidade") produzem um registro marcadamente íntimo.

Em (10), passagem de uma das cartas publicadas na *Revista da Semana*, observamos uma configuração semelhante no que diz respeito à construção de intimidade, mesmo que em outro tipo de relação. A autora dirige-se à destinatária como "minha linda Boneca", mobilizando um vocativo afetivo

que delimita uma cena de intimidade. A carta se organiza em torno de um discurso de aconselhamento voltado à educação dos filhos, especificamente no que se refere à formação moral e intelectual por meio da leitura. Como em (9), há a presença de formas verbais imperativas com marcas de 2P/tu ("não lh'os apontes", "não lh'os offereças", "deixa que tenham a illusão de escolher"), que estruturam o texto como orientação prática.

O que fica evidente, nesse caso, é a articulação entre maternidade, cuidado e responsabilidade formativa, elementos centrais na construção de um ideal de feminilidade no período. A escrevente, nomeada como Elena, assume uma posição de quem detém saber e experiência, mas essa posição é mediada por estratégias de aproximação, que evitam a constituição de um distanciamento no enunciado.

O contexto de familiaridade, ilustrado nos fragmentos (9) e (10), se coaduna à preferência pelo uso de 2P/tu, estratégia pronominal associada, no período em análise, a trocas de maior proximidade entre interlocutores (Lopes; Marcotulio, 2011). Tal aspecto auxilia a construção de uma escrita orientada pelo cuidado, pelo aconselhamento e pelo envolvimento afetivo, em consonância com o ideal feminino promovido por essas publicações.

À luz desses dados, a baixa frequência de 3P/você não deve, portanto, ser interpretada como indicador de que "as mulheres eram conservadoras", do ponto de vista linguístico. O resultado, na verdade, pode corresponder a um efeito das práticas reiteradas por meio das quais se produzem posições do que é reconhecido como feminino. Na pesquisa, essa produção se dá no interior de tecnologias de gênero (no caso, *A Cigarra* e *Revista da Semana*) que regulavam modos legítimos de dizer, sentir e se posicionar como mulher, restringindo o uso de uma variante com o traço de [+distanciamento] (3P/você).

Ao mesmo tempo, o próprio contexto de recepção dessas cartas precisa ser considerado: tratava-se de textos publicados em revistas, cujo funcionamento discursivo demandava a construção de uma cena enunciativa marcada pela intimidade. Nessa situação, o uso recorrente de 2P/tu pode ser interpretado também como uma estratégia de legitimação da carta pessoal, simulando uma confidencialidade que não se sustentava espontaneamente em um espaço de circulação midiática. Ou seja, os resultados a que chegamos não decorrem de uma característica intrínseca das mulheres, e sim da relativa incompatibilidade sociopragmática entre a variante 3P/você e o discurso de intimidade e proximidade reiterado por essas publicações.

5. Considerações finais

Neste artigo, partimos do reconhecimento de que a interpretação da variável social gênero na Sociolinguística Variacionista tem sido orientada por perspectivas essencialistas, baseadas em pressupostos dicotômicos, que tendem a tratar padrões como homogêneos, frequentemente reproduzindo as tendências do Paradoxo do gênero (Labov, 1990, 2001) como uma verdade universal. Ao articular os pressupostos sociolinguísticos às teorias sociais de gênero (Butler, 2018a,

2018b; Lauretis, 1994), consideramos a imprensa voltada ao público feminino como um espaço de performatividade e tecnologia de gênero, por seu papel na construção e disseminação de representações socialmente legitimadas de feminilidade. Essa perspectiva, conforme procuramos sustentar, permite apreender gênero não como categoria demográfica, e sim como dimensão social e histórica.

A aplicação do modelo teórico-metodológico ao qual recorremos na análise da variação entre as marcas de 2P/tu e 3P/você em cartas publicadas na revista *A Cigarra* e em um grupo controle (*Revista da Semana* e *Correio Paulistano*) revelou a viabilidade e a pertinência da proposta. Os resultados indicaram que a performatividade de uma feminilidade específica, guiada pela intimidade, incentivada por essa tecnologia de gênero, atua como um fator restritivo à adoção da variante 3P/você, ligada a traços de [+distanciamento]. A retenção da variante 2P/tu, vinculada a uma [+intimidade], nos dados analisados mostrou-se, portanto, menos como um indicador do sexo biológico dos escreventes e mais como um elemento constitutivo do próprio modelo de feminilidade que tecnologias de gênero (*A Cigarra* e a *Revista da Semana*) produziam e difundiam. Desse modo, o padrão linguístico identificado tem relação com uma performatividade do feminino, e não com uma suposta essência ou propriedade inerente a um sujeito universal "mulher".

Nesse sentido, as reflexões propostas avançam na problematização do uso do gênero social como categoria analítica na Sociolinguística Variacionista, ao fornecerem um instrumental teórico-metodológico para tratá-lo em sua complexidade. Ainda que nosso estudo apresente especificidades, por nos basearmos em dados escritos e em amostras de sincronias pretéritas, marcadas por condições próprias de produção e circulação, entendemos que as questões aqui desenvolvidas não se restringem a esse recorte empírico. Ao contrário, buscamos delinear uma possibilidade teórico-metodológica aplicável a diferentes investigações sociolinguísticas, desde que orientadas por uma perspectiva não essencialista. Tal encaminhamento implica questionar o uso recorrente do gênero como categoria *default* nas pesquisas sociolinguísticas, uma vez que as mulheres não podem ser tomadas como um bloco uniforme, mas devem ser analisadas em sua pluralidade, considerando-se as múltiplas maneiras de constituição de suas posições sociais e discursivas.

Nossa posição se junta à advertência de Wodak e Benke (2017, p. 150, tradução nossa) de que "se as teorias sociológicas do contexto e das conceitualizações de gênero, etc. continuarem a ser ignoradas, as explicações contraditórias não desaparecerão"[16]. Portanto, tendo em conta a promoção de um diálogo interdisciplinar, um olhar social de gênero em todas as suas complexidades históricas e sociais apresenta-se como um caminho, o qual tentamos percorrer neste artigo, para repensar os processos de variação e mudança linguística. Nessa direção, defendemos que cabe ao analista explicitar os limites de seu córpus e orientar suas interpretações de maneira a não naturalizar diferenças, mas a problematizá-las, de modo que os padrões observados não sejam vistos enquanto expressão de essências atribuídas a categorias como "homens" e "mulheres".

---

[16] No original: "If sociological theories of context and of conceptualizations of gender, etc. continue to be ignored, contradictory explanations will not disappear" (Wodak; Benke, 2017, p. 150).

Referências

ARQUIVO PÚBLICO DO ESTADO DE SÃO PAULO. *Acervo Digitalizado (A Cigarra)*. São Paulo: Governo do Estado, [2025]. Disponível em: http://www.arquivoestado.sp.gov.br/site/acervo/repositorio_digital/jornais_revistas. Acesso em: 24 out. 2025.

AUER, A.; PEERSMAN, C.; PICKL, S.; RUTTEN, G.; VOSTERS, R. Historical sociolinguistics: the field and its future. *Journal of Historical Sociolinguistics*, [*s. l.*], v. 1, n. 1, pp.1-12, 2015. Disponível em: https://www.degruyterbrill.com/document/doi/10.1515/jhsl-2015-0001/html?srsltid=AfmBOopbBpogjFQodRDhtEildNu4c1I2XPFBnyI389j-SSeRCubl4G9y. Acesso em: 20 out. 2025. DOI: https://doi.org/10.1515/jhsl-2015-0001.

BARBOSA, M. dos S. *A cor e o silêncio*: apagamento da mulher negra na sociolinguística. 2025. 25 f. Trabalho de Conclusão de Curso (Graduação em Letras) – Universidade Estadual Paulista, Araraquara, 2025.

BASSANEZI, C. B. *Virando as páginas, revendo as mulheres*: revistas femininas e relações homem-mulher, 1945-1964. Rio de Janeiro: Civilização Brasileira, 1996.

BESSE, S. K. *Modernizando a desigualdade*: reestruturação da ideologia de gênero no Brasil (1914-1940). São Paulo: Edusp, 1999.

BIBLIOTECA NACIONAL DIGITAL. *Hemeroteca digital brasileira*. Rio de Janeiro: Fundação Biblioteca Nacional, [2025]. Disponível em: https://bndigital.bn.gov.br/hemeroteca-digital/. Acesso em: 24 out. 2025.

BUCHOLTZ, M. From 'sex differences' to gender variation in sociolinguistics. *Working Papers in Linguistics*, [*s. l.*], v. 8, pp. 33-45, 2002.

BUTLER, J. Os atos performativos e a constituição do gênero: um ensaio sobre fenomenologia e teoria feminista. Tradução: Jamille Pinheiro Dias. *Caderno de leituras*, [*s. l.*], n. 78, pp. 1-16, 2018a. Disponível em: https://chaodafeira.com/wp-content/uploads/2018/06/caderno_de_leituras_n.78-final.pdf. Acesso em: 20 out. 2025.

BUTLER, J. *Problemas de gênero*. Feminismo e subversão da identidade. Tradução de Renato Aguiar. 1. ed. Rio de Janeiro: Civilização Brasileira, 2018b.

CASADEI, E. B. Questões de consumo e a feminização da Revista da Semana. *Revista Z Cultural*, Rio de Janeiro, v. 7, pp. 1-15, 2017. Disponível em: https://revistaz.letras.ufrj.br/questoes-de-consumo-e-a-feminizacao-da-revista-da-semana/. Acesso em: 13 abr. 2026.

CIXOUS, H. *O riso da Medusa*. Rio de Janeiro: Bazar do Tempo, 2022.

CLEMPI, C. B. *Variação entre futuro sintético e perifrástico em textos escritos*: um estudo diacrônico sobre o papel do gênero feminino na mudança linguística. 2019. Dissertação (Mestrado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista, Araraquara, 2019. 139 f. Disponível em: https://sucupira-legado.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popup=true&id_trabalho=7749882. Acesso em: 11 nov. 2025.

CLEMPI, C. B. *Contribuições de teorias sociais de gênero para o estudo da mudança linguística*: uma discussão sobre a implementação de marcas de 3P associadas a 'você'. 2024. Tese (Doutorado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista, Araraquara, 212 f. 2024. Disponível em: https://repositorio.unesp.br/entities/publication/55d4bf85-6f92-4f84-bf77-5e34487fccc9. Acesso em: 11 nov. 2025.

CRENSHAW, K. W. Mapping the Margins: intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, Palo Alto, v. 43, n. 6, pp. 1241-1299, 1991. Disponível em: https://blogs.law.columbia.edu/critique1313/files/2020/02/1229039.pdf. Acesso em: 20 out. 2025.

CYFER, I. Afinal, o que é uma mulher? Simone de Beauvoir e "a questão do sujeito" na teoria crítica feminista. *Lua Nova*: Revista de Cultura e Política, São Paulo, n. 94, pp. 41-77, abr. 2015. Disponível em: https://www.scielo.br/j/ln/a/YgZNbXJJFXCMmCKzKbnnP6t/abstract/?lang=pt. Acesso em: 20 out. 2025. DOI: https://doi.org/10.1590/0102-64452015009400003.

CZRNORSKI, S. R.; MEYRER, M. R. Chics, elegantes e distintas: a moda na seção Jornal das famílias da Revista da Semana (1915-1918). *Revista Brasileira de História & Ciências Sociais*, Rio Grande, v. 8, n. 15, pp. 247-271, jul. 2016. Disponível em: https://periodicos.furg.br/rbhcs/article/view/10626. Acesso em: 13 abr. 2026. DOI: https://doi.org/10.14295/rbhcs.v8i15.295.

DAVIS, A. *Mulheres, raça e classe*. São Paulo: Boitempo Editorial, 2016.

DUMONT, L. M. M.; SANTO, P. E. Leitura feminina: motivação, contexto e conhecimento. *Ciências & Cognição*, Rio de Janeiro, v. 10, pp. 28-37, 2007. Disponível em: https://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1806-58212007000100004. Acesso em: 11 nov. 2025.

ECKERT, P. Three waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *Annual Review of Anthropology*, [*s. l.*], v. 41, p. 87-100, 2012. Disponível em: https://web.stanford.edu/~eckert/PDF/ThreeWaves.pdf. Acesso em: 20 out. 2025.

ECKERT, P.; MCCONNELL-GINET, S. Comunidades de práticas: lugar onde co-habitam linguagem, gênero e poder. *In*: OSTERMANN, A. C.; FONTANA, B. (org.). *Linguagem, sexo, sexualidade*. Clássicos traduzidos. São Paulo: Parábola Editorial, 2010. p. 93-107.

FARACO, C. A. O tratamento você em português: uma abordagem histórica. *LaborHistórico*, Rio de Janeiro, v. 3 n. 2, pp. 114-132, 2017. Disponível em: https://revistas.ufrj.br/index.php/lh/article/view/17150. Acesso em: 24 out. 2025. DOI: https://doi.org/10.24206/lh.v3i2.17150.

FASOLD, R. *The Sociolinguistics of Language*. Oxford: Blackwell, 1990.

FREITAG, R. M. K. (Re)discutindo sexo/gênero na sociolinguística. *In*: FREITAG, R. M. K.; SEVERO, C. G. (org.). *Mulheres, Linguagem e Poder*. Estudos de Gênero na Sociolinguística Brasileira. São Paulo: Blucher, 2015. pp. 17-74.

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. *Alfa*, São José do Rio Preto, v. 56, n. 3, pp. 917-944, 2012. Disponível em: https://www.scielo.br/j/alfa/a/J6ZcH9z3RPYz5ZGxnQkZJkr/?lang=pt. Acesso em: 20 out. 2025. DOI: https://doi.org/10.1590/S1981-57942012000300009.

HELLER, B. Vossas filhas sabem ler? *In*: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 24., 2001, Campo Grande. *Anais* [...]. Campo Grande: Intercom, 2001. pp. 1-17. Disponível em: http://www.intercom.org.br/papers/nacionais/2001/papers/np4heller.pdf. Acesso em: 11 nov. 2025.

HOOKS, B. *E eu não sou uma mulher?*: mulheres negras e feminismo. 4. ed. Rio de Janeiro: Rosa dos Tempos, 2020.

KOCH, P. Tradiciones discursivas y cambio linguístico: el ejemplo del tratamento vuestra merced en español. *In*: KABATEK, J. (ed.). *Sintaxis histórica del español y cambio linguístico*: nuevas perspectivas desde las tradiciones discursivas. Madrid: Vervuert, 2008. pp. 53-88.

LABOV, W. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*, v. 2, n. 2, pp. 205-254, 1990.

LABOV, W. *Principles of Linguistic Change*, vol. 2: Social Factors. Oxford: Blackwell, 2001.

LABOV, W. *Padrões Sociolinguísticos*. Tradução de Marcos Bagno, Marta Scherre e Caroline Cardoso. 1. ed. São Paulo: Parábola, 2016 [1972].

LAURETIS, T. de. A tecnologia do gênero. Tradução de Susana Bornéo Funck. *In*: HOLLANDA, H. B. de (org.). *Tendências e impasses*: o feminismo como crítica da cultura. Rio de Janeiro: Rocco, 1994. pp. 206-242.

LEVON, E.; MENDES, R. B. Introduction. Locating sexuality in language. *In*: LEVON, E.; MENDES, R. B. (ed.). *Language, sexuality and power*: studies in intersectional sociolinguistics. Oxford: Nova Iorque, 2016. pp. 1-18.

LOPES, C. R. S.; DUARTE, M. E. L. Notícias sobre o tratamento em cartas escritas no Brasil dos séculos XVIII e XIX. *In*: RAMOS, J.M.; ALKMIM, M.A. (org.) *Para a história do português brasileiro*: Estudos sobre Mudança Lingüística e História Social. Belo Horizonte: Faculdade de Letras da UFMG, 2007. v. 5, pp. 329-357.

LOPES, C. R. S.; MARCOTULIO, L, L. O tratamento a Rui Barbosa. *In*: CALLOU, D.; BARBOSA, A. (org.). *A norma brasileira em construção*: cartas a Rui Barbosa (1866 a 1899), 1. ed. Rio de Janeiro: Fundação Casa de Rui Barbosa, 2011. v. 2, p. 265-292.

LOPES, C. R. s.; MARCOTULIO, L. L.; OLIVEIRA, T. L. de. A atuação dos papéis sociais na mudança no sistema de tratamento no português brasileiro: análise de cartas pessoais (1870-1979). *Estudos de Linguística Galega*, [*s. l.*], pp. 29-44, 2018. Disponível em: https://revistas.usc.gal/index.php/elg/article/view/3543. Acesso em: 24 out. 2025. DOI https://doi.org/10.15304/elg.ve1.3543.

LOPES, C. R. S.; MARCOTULIO, L. L.; RUMEU, M. C. B.; ANDRADE, A.; COELHO, I. L.; MARTINS, M. A.; LACERDA, M. O.; GOMES, V. S.; MONTE, V. M. CARNEIRO, SOUZA, C. M. N.; BALSALOBRE, S.; SOUZA, J. P.; OLIVEIRA, T. L.; MOURA, K. K.; CRUZ, I.; CARDOSO, N. D. A reorganização do sistema pronominal de 2ª pessoa na história do português brasileiro: a posição de sujeito. *In*: CASTILHO, A. T. (coord.). *História do português brasileiro*: mudança sintática das classes de palavra: perspectiva funcionalista. São Paulo: Contexto, 2018. v. 4, pp. 26- 141.

MILROY, L. *Language and Social Networks*. Oxford: Blackwell, 1980.

MIRA, M. C. *O leitor e a banca de revistas*: a segmentação da cultura do século XX. São Paulo: Olho D'Água/ Fapesp, 2001.

MOREIRA, L. G. *Em dia com a moda*. Da emancipação feminina ao culto à juventude nos anos 60: a representação da feminilidade nas páginas de *A Cigarra*. 2011. 159 f. Dissertação (Mestrado em Comunicação Social), Pontifícia Universidade Católica, Rio de Janeiro, 2011. Disponível em: https://www.maxwell.vrac.puc-rio.br/colecao.php?strSecao=resultado&nrSeq=34584&idi1=&rc=1. Acesso em: 20 out. 2025.

NICKERSON, R. S. Confirmation bias: a ubiquitous phenomenon in many guises. *Sage Journals*, [*s. l.*], v. 2, n. 2, pp. 175-220, 1998. Disponível em: https://journals.sagepub.com/doi/10.1037/1089-2680.2.2.175. Acesso em: 20 out. 2025. DOI: https://doi.org/10.1037/1089-2680.2.2.175.

OUSHIRO, L. As variáveis sexo/gênero e indivíduo em situação de contato dialetal. *In*: CARVALHO, D.; BRITO, D. (org.). *Gênero e língua(gem)*: formas e usos. Salvador: Editora da Universidade Federal da Bahia, 2020. pp. 43-66.

PEDRO, J. M. Traduzindo o debate: o uso da categoria gênero na pesquisa histórica. *História*, São Paulo, São Paulo, v. 24, pp. 77-98, 2005. Disponível em: https://www.scielo.br/j/his/a/fhHv5BQ6tvXs9X4P3fR4rtr/?format=pdf&lang=pt. Acesso em: 24 out. 2025. DOI: https://doi.org/10.1590/S0101-90742005000100004.

PEREIRA, R. de O. *O tratamento em cartas amorosas e familiares da Família Penna*: um estudo diacrônico. 2012. 157 f. Dissertação (Mestrado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012. Disponível em: https://laborhistorico.letras.ufrj.br/Mestrado/PereiraRO.pdf. Acesso em: 24 out. 2025.

PODESVA, R. J. Phonation type as a stylistic variable: the use of falsetto in constructing a persona. *Journal of Sociolinguistic*, Hoboken, v. 11, n. 4, pp. 478-504, 2007.

PODESVA, R. J. Salience and the social meaning of declarative contours: three case studies of gay professionals. *Journal of English Linguistics*, [*s. l.*], v. 39, n. 3, pp. 233-264, jul. 2011. Disponível em: https://journals.sagepub.com/doi/10.1177/0075424211405161. Acesso em: 13 abr. 2026. DOI: https://doi.org/10.1177/0075424211405161.

R CORE TEAM. *R*: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2023. Disponível em: https://www.R-project.org/. Acesso em: 20 jul. 2023.

ROMAINE, S. *Socio-Historical Linguistics*: its status and methodology. Cambridge: Cambridge University Press, 2009.

RODRIGUES, A.; CLEMPI, C. B.; FELIX, R. de A. A. Approaching Gender and Sexuality Issues from a Sociolinguist Corpus-Based Analysis: A Methodologic Challenge. *In*: MASSINI-CAGLIARI, G.; BERLINCK, R. A.; RODRIGUES, A. (ed.) *Understanding Linguistic Prejudice*. *Cham: Springer*, 2023. pp. 139-155. Disponível em: https://link.springer.com/chapter/10.1007/978-3-031-25806-0_9. Acesso em; 11 nov. 2025. DOI: https://doi.org/10.1007/978-3-031-25806-0_9.

RUMEU, M. C. B. *A implementação do 'você' no português brasileiro oitocentista e novecentista*: um estudo de painel. 2008. 275 f. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008. Disponível em: http://www.letras.ufmg.br/padrao_cms/documentos/profs/marciarumeu/Rumeu%20[2008].pdf. Acesso em: 24 out. 2025.

SCALZO, M. *Jornalismo de revista*. 4. ed. São Paulo: Contexto, 2016.

SOUZA, J. P. F. *Mapeando a entrada de você no quadro pronominal*: análise de cartas familiares dos séculos XIX-XX. 2012. 147 f. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012. Disponível em: https://laborhistorico.letras.ufrj.br/Mestrado/SouzaJPF.pdf. Acesso em: 24 out. 2025.

TOSALD, L. Do problema do essencialismo a outra maneira de se fazer política: retomando o potencial transformador das políticas de diferença. *Mediações*, Londrina, v. 15, n. 2, pp. 166-183, 2010. Disponível em: https://ojs.uel.br/revistas/uel/index.php/mediacoes/article/view/8228. Acesso em: 20 out. 2025. DOI: https://doi.org/10.5433/2176-6665.2010v15n2p166.

TRUDGILL, P. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, Cambridge, n. 1, pp. 179-195, out. 1972. Disponível em: https://www.cambridge.org/core/journals/language-in-society/article/sex-covert-prestige-and-linguistic-change-in-the-urban-british-english-of-norwich/6A687B9A4422D36E3C85E5A8407EC00D. Acesso em: 13 abr. 2026. DOI: https://doi.org/10.1017/S0047404500000488.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução: Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].

WODAK, R.; BENKE, G. Gender as a sociolinguistic variable: New perspectives on variation studies. *In*: COULMAS, F. (org.). *The handbook of sociolinguistics*. Oxford: Blackwell, 2017. pp. 127-150.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a71988

Amigos gaysssss: efeitos da duração de /-s/ final em percepções de gênero/sexualidade em vozes masculinas

Gay friendssss: effects of word-final /-s/ lengthening on perceptions of gender/sexuality in male voices.

Marcus Garcia de Sene[1]

Ronald Beline Mendes[2]

RESUMO

Este artigo explora efeitos da duração da coda /-s/ em final de palavra em percepções de gênero/sexualidade acerca de vozes masculinas em português brasileiro. Tais efeitos foram testados experimentalmente, com a técnica matched-guise, a partir de estímulos acústicos caracterizados por coda /-s/ breve (s-) ou alongada (s+). De acordo com as respostas de 122 ouvintes, em escalas de diferenciais semânticos, verifica-se que os quatro falantes cujas vozes foram utilizadas soam "menos masculinos/mais gay" quando ouvidos no disfarce s+. Tal efeito não varia de acordo com o sexo do ouvinte, sua escolaridade, faixa etária ou orientação sexual - o que sugere generalidade da correlação entre a variável linguística e percepções de gênero/sexualidade. Por outro lado, ouvintes que declaram ter poucos amigos gays em suas redes de relacionamento (relativamente a ouvintes que dizem ter muitos amigos gays) tendem a perceber os falantes do experimento menos vezes como "menos masculino/mais gay" diante de s+, o que sugere que o funcionamento de /-s/ como índice de gênero/sexualidade em vozes masculinas é sujeito ao acesso que os ouvintes têm a essa variação (cuja produção, associada a práticas sociais, ainda não foi estudada). Este estudo insere o português brasileiro no rol de línguas em que a indicialidade de /-s/ vem sendo investigada e aponta para encaminhamentos na continuação das pesquisas sobre esse tema.

PALAVRAS-CHAVE: Duração da coda /-s/. Gênero. Sexualidade. Percepção sociolinguística.

ABSTRACT

This paper explores effects of word-final coda /-s/ lengthening in perceptions of gender/sexuality in male voices, in Brazilian Portuguese. Such effects have been tested experimentally, according to the matched-guise technique, with acoustic stimuli defined by short (s-) or longer /-s/ (s+). Responses by 122 listeners, in semantic differential scales, show that four speakers were perceived as "less masculine/gayer-sounding" in their s+ guises. Such effect does not vary according to listener's sex, education, age group or sexual orientation – which indicates an overall correlation between the linguistic variable and perceived gender/sexuality. However, listeners who have few gay friends in their social networks (relative to those who have many gay friends) tend to perceive the four speakers fewer times as "less masculine/gayer-sounding" in their s+ guises. This suggests that /-s/ as an index of gender/sexuality in male voices is subject to the access listeners have to /-s/ variation in their social practice (though the production of /-s/ lengthening has not yet been investigated). This study inserts Brazilian Portuguese among the languages in which the indexicality of /-s/ has been scrutinized, and it points out to directions in future research.

KEYWORDS: /-s/ lengthening. Gender. Sexuality. Sociolinguistic perception.

[1] Universidade de Pernambuco (UPE), marcus.sene@upe.br, https://orcid.org/0000-0002-2715-5294.

[2] Universidade de São Paulo (USP), rbeline@usp.br, https://orcid.org/0000-0002-1510-7180.

Recebido em 30/12/2025 - Aceito em 16/06/2026

1. Introdução

Este artigo discute os resultados principais de um experimento de percepção sociolinguística, centrado na seguinte questão: a duração da coda /-s/ (em final de palavra) exerce efeito na percepção de gênero e sexualidade em vozes masculinas? Tal pergunta, bem como o experimento que a partir dela se organiza, justifica-se por três motivações principais:

(i) Mendes (2016, 2018) verificou efeitos da variável concordância nominal (CN) na percepção de efeminidade nas vozes de 4 falantes do sexo masculino: todos foram percebidos como mais efeminados quando ouvidos na versão com concordância padrão (CNp, como em "meus amigo-S"), relativamente a quando foram ouvidos na sua versão não padrão (CNØ, "meus amigo-Ø"). Chama a atenção, contudo, a variação na percepção de efeminidade na voz de um dos falantes (cujo pseudônimo é Robson). Em um dos trechos de fala utilizado no experimento, Robson foi percebido como relativamente muito menos efeminado na versão CNØ, tal como no resultado geral. Entretanto, num segundo trecho, a diferença das respostas para Robson - a depender da versão CNp ou CNØ - foi muito menor; a diferença entre os dois trechos hipoteticamente reside no fato de que, no segundo, Robson, na versão CNp, pronuncia a coda /-s/ de maneira marcadamente alongada – o que sugere que, além da variável gramatical (CN), a duração da coda /-s/ poderia ter um efeito na percepção de quão efeminada soa uma voz masculina.

(ii) O efeito da duração de /-s/, bem como o de seu centro de gravidade, em percepções de gênero e sexualidade, já foi desenvolvido em espanhol (Mack, 2016), em inglês (Levon, 2006, 2007), em dinamarquês (Maegaard; Pharao, 2016), nas variantes White Afrikaans e White South African English (WSAfE) (Bekker; Levon, 2017). Já em português brasileiro, estudos de tais efeitos são praticamente inexistentes (ver, contudo, Sene, 2022). Interessa, portanto, investigar se a amplitude dos efeitos da duração de /-s/, em termos da quantidade e diversidade de línguas em que eles se verificam, inclui também o português brasileiro.

(iii) Nesta língua, não é raro o desenvolvimento de discursos metalinguísticos nos quais se associa "soar gay"[3] à pronúncia do segmento /-s/, tanto em termos de duração (imagem 1a) quanto em termos de posteriorização (imagem 1b). Esses comentários apontam para um possível estereótipo, que convém explorar – inclusive experimentalmente, ou seja, com técnicas que permitam verificar a ocorrência de tais associações quando ouvintes não têm consciência acerca da variável linguística em foco no estudo.

[3] Para uma discussão sobre a diferença, para estudos sociolinguísticos, entre e "soar" e "ser" gay, recomenda-se Mendes (2018) e Sene (2022).

Imagem 1: Discursos metalinguísticos em que se revela uma associação entre a pronúncia de /s/ e percepções de gênero/sexualidade

a)



b)



Fonte: Sene (2022)

Com vistas à delimitação deste estudo, focaliza-se especificamente a duração da coda /-s/ em posição final de palavra (como em *meninosssss* - imagem 1a). Analisar essa variável justifica-se na medida em que a duração integra o repertório de pistas linguísticas mobilizadas pelos falantes na construção de significados sociais, conforme ilustra a imagem 1. Ainda que os falantes não disponham de um vocabulário técnico ou socialmente estabilizado para descrever formalmente as formas linguísticas (*cf*. Mendes, 2016, sobre Labov, 1996, p. 266-267), a duração emerge como um dos correlatos linguísticos associados à percepção do "soar gay". Além disso, a duração manifesta-se acusticamente como alongamento de um segmento sonoro, configurando um parâmetro relativamente acessível à audição e, portanto, passível de avaliação social por parte dos falantes.

A duração de /s/ apresenta elevada complexidade fonética e social, uma vez que a ela se associam conteúdos que extrapolam aquilo que é estritamente linguístico. Em torno dessa variável, articulam-se diferentes significados sociais, relacionados, por exemplo, às dimensões de gênero, orientação sexual e até mesmo pertencimento dialetal. Tal fenômeno decorre do fato de que as variáveis linguísticas operam como índices de significados sociais localmente construídos, ativados e negociados em práticas discursivas específicas (Eckert, 2008; Mendes, 2018; Sene, 2022). Do ponto de vista perceptual, essas associações podem ser compreendidas à luz de vieses cognitivos, entendidos, então, como relações implícitas entre variantes linguísticas e 'estereótipos' sociais (Sene, 2022). Tais relações são construídas a partir da experiência linguística de falantes e ouvintes e se sustentam, por exemplo, na sensibilidade ao detalhe linguístico e aos conteúdos indiciais mobilizados pela fala.

O experimento aqui apresentado insere-se, portanto, em uma sociolinguística interessada, de um lado, na indicialidade (Silverstein, 2003; Eckert, 2008, 2012) – isto é, na investigação de como significados sociais são ancorados em elementos linguísticos e de que modo esses significados podem ser acessados empiricamente – e, de outro, nas articulações entre linguagem, gênero e sexualidade. Estudos realizados em diferentes línguas (Levon, 2006, 2007; Mack, 2016; Bekker; Levon, 2017; Maegaard; Pharao, 2016) têm demonstrado, de forma consistente, que detalhes acústicos e parâmetros espectrais do /s/, como o centro de gravidade, são mobilizados pelos ouvintes na construção de avaliações sociais relativas a masculinidades, efeminidade e orientação sexual, ainda que tais avaliações possam ocorrer abaixo do nível da consciência metalinguística.

Ao transpor esse debate para o português brasileiro, o presente estudo busca não apenas ajudar a preencher uma lacuna empírica, mas também testar o efeito de uma variável linguística na percepção sociolinguística de gênero e sexualidade. Assim, ao focalizar experimentalmente a duração da coda /-s/, este trabalho contribui para o avanço das pesquisas sobre percepção sociolinguística no Brasil, oferecendo evidências sobre como significados sociais ligados a gênero e sexualidade podem ser indiciados por formas linguísticas específicas e acessados por meio de julgamentos perceptuais, em consonância com uma agenda contemporânea da sociolinguística que articula variação, percepção e o significado social da variação.

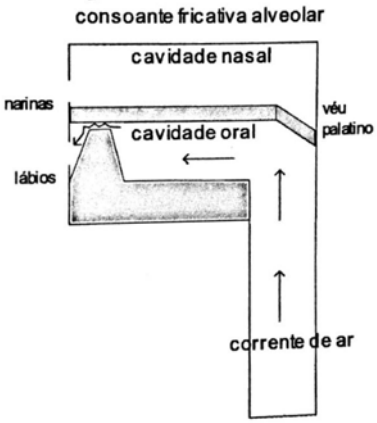
2. /-s/ como índice de gênero e sexualidade

Com vistas a embasar as decisões metodológicas relativas ao preparo do experimento de percepção sociolinguística apresentado neste artigo, revemos brevemente, nesta seção, alguns trabalhos que investigaram a relação entre a variável /-s/ e percepções de gênero e sexualidade. Crist (1997), Smyth, Jacobs e Rogers (2003) e Stuart-Smith, Timmins e Wrench (2003) focalizam dados de produção e, aqui, servem para mostrar que propriedades fonético-acústicas de /s/, em especial a duração e a sibilância, variam de modo sistemático em correlação a estilos sociais associados a gênero, sexualidade e classe – propriedades essas que não podem ser explicadas exclusivamente por diferenças anatômicas entre falantes.

Por outro lado, os estudos de percepção de Campbell-Kibler (2011) e Levon (2007, 2014) são centrais para este trabalho porque mostram como ouvintes associam variantes de /s/ a julgamentos sociais relacionados a masculinidades e a orientação sexual, frequentemente mediados por estereótipos socialmente compartilhados (como classe e competência), bem como por outros elementos linguísticos. Aqui, o maior interesse por esses estudos reside no fato de que eles demonstram, por meio de experimentos controlados, que pequenas manipulações acústicas são suficientes para produzir alterações na percepção social, oferecendo, assim, respaldo ao experimento que desenvolvemos e cujos resultados principais aqui relatamos e discutimos.

Dentre os sinais acústicos que o trato vocal pode produzir, as fricativas são articuladas por um estreitamento do trato oral, em diferentes pontos de articulação (desde mais anteriores até mais posteriores), que resulta numa resistência à passagem da corrente de ar e, assim, uma fricção e um ruído a ela correlato. As fricativas podem ainda ser articuladas como sibilantes, caracterizadas pelo choque "da corrente de ar egressiva contra um obstáculo localizado após o ponto de constrição, isto é, após o ponto em que o articulador ativo e o articulador passivo se aproximam formando o estreitamento" (Biasibetti, 2018, p. 20).

Figura 1: Diagrama esquemático da consoante fricativa alveolar



Fonte: Silva *et al.*, 2019

Como mostra o diagrama, a pressão exercida em determinado ponto é o que origina o efeito de turbulência e, a partir de uma pequena abertura, a passagem de ar produz sons fricativos como [f], [v], [s], [z], etc. Entre os parâmetros acústicos mais relevantes para a caracterização do /s/ estão os picos espectrais, momentos espectrais e duração (relativa e absoluta) (Haupt, 2007; Oliveira, 2011; Bassi; Seara, 2017). Esses parâmetros são responsáveis pela distinção entre, de um lado as fricativas alveolares /s/ e /z/ e, de outro, as palato-alveolares /ʃ/ e /ʒ/. Para além dessa diferenciação, a variabilidade acústica existente nesses parâmetros pode ser explicada por meio de condicionadores como a coarticulação, bem como fatores não linguísticos, como gênero e estilo.

Do ponto de vista anatômico, homens tendem, em média, a apresentar maior estatura corporal do que mulheres (Heffernan, 2004; Biasibetti, 2018) e, de modo correlato, tratos vocais mais longos e câmaras de ressonância mais volumosas (Kent; Read, 2015). Além disso, as pregas vocais dos homens costumam ser mais pesadas, o que tem relação direta com sua menor vibração e, consequentemente, com uma voz mais grave (f0 menor). Essas diferenças anatômicas costumam resultar em valores mais baixos de frequência fundamental para falantes do sexo masculino e mais altos para falantes do sexo feminino. No entanto, tais diferenças localizam-se predominantemente na porção mais posterior do trato vocal (Biasibetti, 2018) e não são suficientes para explicar, por si só, a variabilidade observada em determinados detalhes fonético-acústicos. Evidências nesse sentido são apresentadas em Crist (1997), que analisou a duração de /s/ na fala de seis homens – dois que se autodeclararam heterossexuais e quatro que se autodeclaram homossexuais. Seus resultados indicam que a duração da fricção associada a /s/ foi sistematicamente mais longa na fala dos participantes que se identificaram como gays, divergindo de análises que restringem a explicação da produção desses sons a aspectos estritamente anatômicos entre homens e mulheres. Achados convergentes são reportados por Smyth, Jacobs e Rogers (2003), que observaram não apenas maior duração de /s/ e /z/, mas também picos de frequência mais elevados em vozes masculinas percebidas como vozes que soam gay.

Stuart-Smith, Timmins e Wrench (2003) examinam a relação entre a produção de /s/ e a diferenciação de sexo/gênero, considerando sua interação com fatores como idade e classe social. Eles observam que, de modo geral, mulheres apresentam valores médios mais elevados e picos espectrais mais altos na realização de /s/, com base na frequência do ponto médio da curva espectral. No entanto, tais padrões não se distribuem de forma homogênea no interior do grupo feminino. Em particular, os valores de média e de frequência mínima de /s/ diferem entre mulheres que se alinham mais fortemente a estilos associados à classe trabalhadora de *Glasgow* e aquelas vinculadas à classe média. Um achado central do estudo é que as medidas acústicas das mulheres da classe trabalhadora se aproximam, em vários aspectos, daquelas observadas entre homens da mesma classe social. Esses resultados indicam, portanto, que as diferenças na produção de /s/ não podem ser explicadas apenas por fatores anatômicos ou biológicos relacionados ao sexo, mas devem ser compreendidas à luz de práticas estilísticas socialmente situadas. Nesse sentido, /s/ emerge como um recurso sociolinguístico sensível à performance de gênero (Sene, 2022), articulada em intersecção com a classe social, evidenciando que a variação fonética participa ativamente da construção de estilos de fala que indiciam posições sociais e identitárias específicas.

Nos estudos de percepção, Campbell-Kibler (2011) apresenta evidências experimentais sobre estereótipos de gênero e sexualidade associados à produção do /s/, examinando a intersecção entre percepções de masculinidade, orientação sexual e competência em um experimento sociolinguístico a partir de uma modificação da técnica de *matched-guise,* originalmente proposta por Lambert *et. al*. (1960). A autora manipula acusticamente o /s/, contrastando realizações posteriorizadas e anteriorizadas, além de controlar outras variáveis sociolinguísticas, como (ING) – cujas variantes são a pronúncia velar ou alveolar, em palavras como *talking* 'falando'. Disfarces (*guises*) com realização mais posterior de /s/ levam sistematicamente a percepções de menos masculinidade e maior probabilidade de o falante ser percebido como 'gay', além de desencadear julgamentos de competência mais baixa. Nesses disfarces, mesmo quando o falante é avaliado como inteligente ou educado — especialmente em combinação com a pronúncia velar de (ING) – a percepção de masculinidade atribuída permanece baixa, o que rompe a associação usual entre competência e masculinidade – *cf*. Mendes (2016, 2018).

Em contraste, nos guises com /s/ anteriorizado, quando os ouvintes percebem o falante como uma pessoa que não soa gay, avaliações negativas de competência não implicam uma correspondente redução na masculinidade percebida. Esses resultados evidenciam que /s/ posterior e /s/ anterior operam como recursos estilísticos distintos, capazes de reorganizar de maneiras diferentes as relações indiciais entre gênero, sexualidade e competência, reforçando a noção de que tais avaliações derivam de performances sociolinguísticas socialmente indiciadas e não de atributos biológicos do falante.

Levon (2014), por sua vez, investigou a relação entre certos estereótipos de gênero e sexualidade e a atribuição de características como "masculino" ou "gay" a excertos da fala de um homem britânico, além de analisar quanto um significado social potencial ("classe baixa") pode bloquear a emergência, no experimento, do significado "(soa) gay". Valendo-se da técnica *matched-guise* modificada, criou

disfarces que combinam as variantes de três variáveis linguísticas: *pitch médio* (estereotipicamente associado a gênero, de acordo com um conjunto de trabalhos desenvolvidos em inglês); sibilância de /s/, que outros trabalhos sobre percepção sociolinguística já haviam demonstrado correlacionar-se à identificação de vozes masculinas como vozes que soam gay; e realizações lábio-dentais [f] de TH (em palavras como *think* 'pensar'), cuja pronúncia padrão é uma fricativa interdental /θ/ (a pronúncia lábio-dental é considerada típica das classes sociais mais baixas no Reino Unido). Para estudar a relação dos estereótipos com as percepções sociolinguísticas, o pesquisador solicitou aos participantes que respondessem ao questionário "Male Role Attitudes Survey" (MRAS), desenvolvido por Pleck, Sonenstein e Ku (1993, 1994), a respeito do quanto concordavam com algumas afirmações acerca de masculinidade e dos papéis sociais dos homens. Levon (2014) verificou correlações, então, entre o grau de identificação dos participantes com tais afirmações e suas percepções de "masculinidade" e "homossexualidade" acerca dos falantes cujas vozes foram utilizadas no experimento.

Embora a variante /s/ tenha sido fortemente associada à percepção de "homossexualidade", quando combinada à pronúncia [f] para TH observou-se relativa diminuição nas avaliações de homossexualidade, bem como um aumento no número de vezes em que as vozes foram percebidas como mais masculinas. Dessa forma, Levon (2014) chama a atenção para o fato de que a percepção de gênero/sexualidade associada a /s/ pode depender de percepções de classe – ou seja, a associação dos significados "masculino" ou "soa gay" a variantes de /s/ não é absoluta, mas dependente de outros estereótipos. Além disso, a avaliação de que os falantes soam "mais gay" e "menos masculinos" nos disfarces com *pitch* médio mais alto e maior sibilância de /s/ foi mais frequente no caso dos respondentes que mais se afiliaram a estereótipos sobre masculinidade no MRAS; diferentemente, não fizeram tal associação os ouvintes que demonstraram menor concordância com as conceituações heteronormativas de gênero do MRAS.

Resultados como os encontrados por Levon (2007, 2014) e Campbell-Kibler (2011) nos servem para sempre ter em conta que não se devem interpretar associações entre formas linguísticas e significados sociais (tais como os baseados em gênero e sexualidade), que emergem em experimentos de percepção sociolinguística, como evidência cabal de um elo "natural" (ou essencial) entre certa variante (como um /s/ mais sibilante ou um /s/ mais longo) e certa identidade social. Mendes (2018, p. 41) ressalta que pode passar a "fazer parte de uma visão social a relação entre certos modos de falar e características/práticas normalmente associadas a certos grupos". Por outro lado, nas práticas sociais propriamente, indivíduos não são passivos, como se fossem aprisionados a uma identidade (com base em gênero e sexualidade, por exemplo), enquanto uma característica "pronta", que está "dada" (ver, por exemplo, Eckert; McConnell-Ginet, 1992, Cameron, 1997; Hall, 1995, 2005; Hall; O'Donovan, 1997; Bucholtz; Hall, 2004). Ao contrário disso, eles são atores sociais que escolhem, com alguma liberdade, como soar e como se definir (ainda que tal "escolha" possa ser cerceada por fatores sociais e, em particular, pela forma como os outros os percebem).

À luz desse conjunto de trabalhos, torna-se claro que a variável /s/ ocupa um lugar central nas investigações sociolinguísticas sobre gênero e sexualidade. Os estudos aqui brevemente retomados mostram que propriedades como duração e sibilância de /s/ podem ser tanto mobilizadas por falantes na construção de estilos sociais quanto interpretadas por ouvintes como pistas relevantes para avaliações de masculinidade, efeminidade e orientação sexual, ainda que tais avaliações operem abaixo do nível da consciência linguística. É nesse quadro teórico que se insere o experimento de percepção proposto neste artigo, que focaliza a duração da coda /-s/ em português brasileiro, com o objetivo de testar se (e em que medida) esse detalhe acústico influencia julgamentos acerca de gênero e sexualidade em vozes masculinas.

3. O experimento

Tal como vimos na seção introdutória, a principal pergunta que motivou o experimento que desenvolvemos é sobre se a duração variável da coda /-s/ tem efeito em percepções de gênero e sexualidade acerca de vozes masculinas. Na presente seção, descrevem-se os procedimentos, desde a preparação dos estímulos acústicos até a aplicação do experimento, passando:

(i) pela seleção de trechos da fala de quatro paulistanos do sexo masculino;

(ii) pela sua manipulação digital (no sentido de alongar as ocorrências de /-s/ final e criar pares de estímulos para cada falante - com /-s/ na duração original (s-) ou com /-s/ alongado (s+));

(iii) pela preparação do chamado "formulário de percepção" (por meio do qual ouvintes voluntários descrevem a pessoa que imaginaram a partir da audição de um estímulo).

No que toca ao primeiro desses três pontos, os falantes selecionados – cujos pseudônimos são, em ordem alfabética, Carlos, Jaime, Lucas e Robson – foram entrevistados pelo Projeto SP2010 (Mendes; Oushiro, 2012). Das suas respectivas entrevistas sociolinguísticas, foram extraídos trechos de fala (transcritos a seguir, tal e qual nas transcrições do Projeto SP2010) no qual houvesse dois tokens de /-s/ final que não fossem seguidos de vogal – um contexto fonético favorecedor de sândi, como em 'meus amigos', que dificultaria a manipulação digital no sentido do alongamento do /-s/. Interessa mencionar que esses são os mesmos falantes cujas vozes foram utilizadas por Mendes (2016, 2018), em seu experimento com foco nos efeitos da (CN) em percepções de gênero/sexualidade (embora os trechos selecionados tenham sido diferentes)[4].

[4] O segmento /s/ da forma verbal trabalhamos – no estímulo do Lucas - não foi submetido à manipulação acústica por duas razões principais. Em primeiro lugar, o item lexical não era seguido por uma pausa prosódica, de modo que qualquer intervenção na duração do /s/ resultaria em um efeito perceptivo artificial, gerando estranheza do ponto de vista auditivo. Em segundo lugar, mesmo na hipótese de haver uma pausa subsequente, a manipulação do /s/ em trabalhamos, quando seguida do pronome nós, produziria um estímulo altamente saliente, com grande potencial de atrair de forma imediata a atenção dos ouvintes, comprometendo a naturalidade do material experimental.

Carlos
Falam que aqui é Alto do Boa Vista que é um lugar mai**s** .... nobre ... aqui para cima então é um bairro mai**s** ... popular

Jaime
Estados Unidos eu não sei por que não conheço, Nova York é uma cidade que me apetece um pouco, eu acho .... Londre**s** .... ma**s** ..... eu amo São Paulo

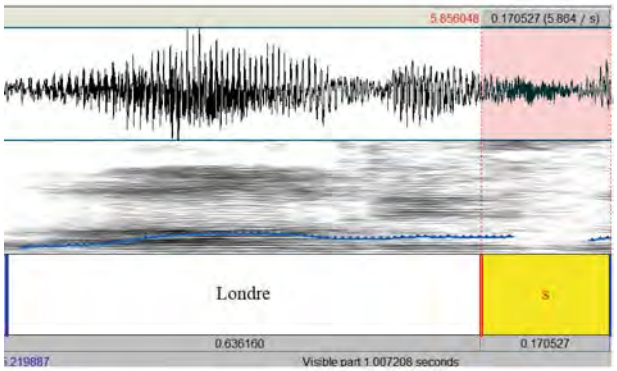
Lucas
Eu passei lá e estava meio assim: "nó**s** ... trabalhamos para ganhar um dinheiro e gastar com coisa que não são importante**s**

Robson
Eu era muito aloprado perto dos mai**s** velho, eu sempre fui uma criança que tacava o terror mai**s** ... que ao mesmo tempo era muito bonzinho, inocente.

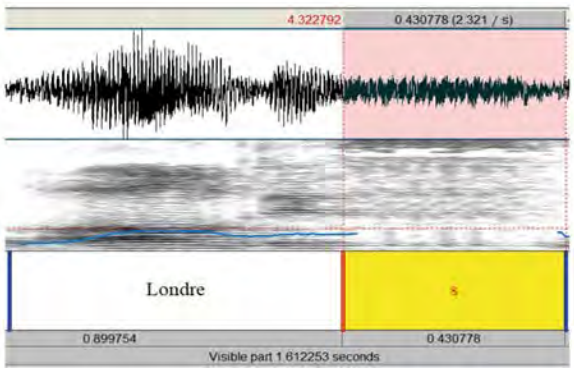
Após a seleção desses trechos, eles foram manipulados no Praat (Boersma; Weenink, 2014) no sentido de aumentar a duração dos /-s/ em destaque e de, assim, criar duas versões de cada trecho: aquela em que os tokens de /s/ permaneceram na duração original, como no exemplo da imagem 2a, e aquela em que a duração dos tokens da fricativa foi aumentada (imagem 2b).

Imagem 2: Exemplo de /-s/ final na palavra Londres (no trecho da fala de Jaime), na duração original (0.17 ms, em 3a) e na versão alongada (0.43 ms, em 3b)

a)



b)



Fonte: Elaboração dos autores

Daqui em diante, vamos nos referir a essas duas versões de cada trecho como "disfarces" (*guises*): o disfarce "s-" (duração original, mais curta – entre 0.07 e 0.17 milissegundos) e o disfarce "s+" (duração alongada, por meio de manipulação digital, entre 0.27 e 0.43 ms – aproximadamente 59% mais longa que a original). Embora a nomeação dos disfarces seja feita com base na variável linguística, o termo disfarce, a rigor, se refere ao falante. Em outras palavras, o ouvinte voluntário ouve cada falante em apenas um dos disfarces – sem ter consciência de que o que os difere é uma variável linguística específica (e apenas ela).

Este é o cerne da chamada técnica *matched-guise* (Lambert *et al*., 1960; Campbell-Kibler, 2009) – ou técnica dos estímulos pareados. O único elemento que os distingue é a variável linguística

em foco; como, no nosso caso, a variável "duração de /-s/ final" é constituída de duas variantes ("s-" ou "s+"), temos um par de estímulos para cada falante. Em cada par, nada além da variável /-s/ os diferencia: nem o volume da voz do falante; nem sua velocidade de fala; nem a frequência fundamental (f0) - por exemplo. Dessa forma, se as respostas dos ouvintes a cada estímulo de um par variarem significativamente, podemos entender que tal variação se deve à diferenciação linguística entre eles, ou seja, à variável em foco.

Os estímulos foram organizados em duas condições experimentais (ver quadro 1). Na Condição 1 (C1), Carlos, por exemplo, se apresenta no disfarce s+, ao passo que, na Condição 2 (C2), ele aparece no disfarce s-. Nesse desenho experimental, os ouvintes voluntários são estimulados pelos trechos de fala dos quatro falantes, mas em apenas um de seus disfarces. Para outros desenhos possíveis, ver Drager (2018).

Quadro 1: Organização dos estímulos

Falante	**Disfarce (guise)**	
	C1	C2
Carlos	s+	s-
Jaime	s-	s+
Lucas	s-	s+
Robson	s+	s-

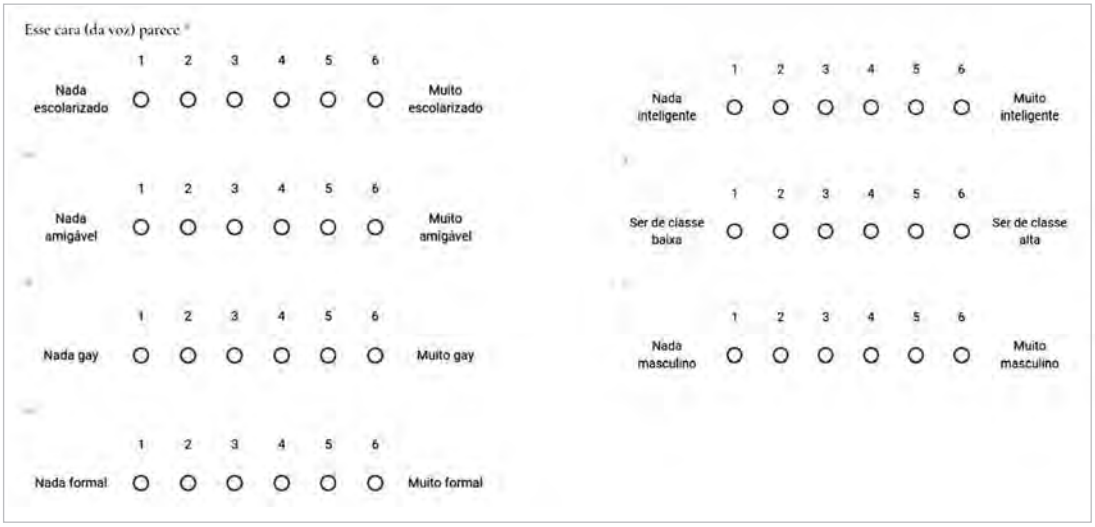
Fonte: Elaboração dos autores

Esse tipo de desenho experimental, em que o ouvinte interage apenas com uma de duas condições (isto é, ouve cada falante em apenas um de dois disfarces), é conhecido como *design between-subject* ou desenho intersujeito – que tem sido utilizado em boa parte dos trabalhos de percepção sociolinguística (Smyth, Rogers, 2008; Campbell-Kibler, 2006; Podesva, 2015; Oushiro, 2015; Santos, 2020; Mendes, 2016, 2018; Sene, 2022), em virtude da combinação de duas razões principais: (i) organizando-se os estímulos de modo que os ouvintes interagissem com todos os oito (isto é, com os falantes em ambos os disfarces – desenho denominado intrasujeito, *within-subject*), muito provavelmente eles se dariam conta de que ouviram uma mesma voz duas vezes, tendo em vista que é reduzido o número de vozes diferentes; desse modo, seria imperiosa a utilização de estímulos distratores, no sentido de que os ouvintes não identificassem a variável linguística em foco; (ii) consequentemente, cada ouvinte demoraria consideravelmente mais tempo para finalizar sua participação no experimento, tendo em vista que teria que ouvir cada falante duas vezes, além dos distratores e, para cada estímulo, preencher o "formulário de percepção", por meio do qual descrevem a pessoa que imaginaram ao ouvir cada um dos estímulos; diferentes pesquisas sobre percepção desenvolvidas com a mesma técnica sugerem que quanto mais longa a duração de um experimento, maiores são as chances de que os ouvintes não cheguem ao final dele (o que, evidentemente, leva a respostas incompletas por parte do conjunto de respondentes voluntários).

Adicionalmente, fez parte de nossas decisões em torno do desenho experimental não variar a ordem de apresentação dos estímulos. Sem dúvida, tal ordem pode influenciar as respostas dos ouvintes: pode ser que, primeiro ouvir um estímulo s- e, depois, um estímulo s+ (em vez do contrário) conduza a respostas diferentes por parte dos ouvintes. Trata-se do que se chama de efeito de preparação (ou *priming*) – processo pelo qual o contato prévio com 'algo' (uma forma linguística/significado) pode influenciar, de algum modo, o processamento linguístico subsequente (McDonough; Trofimovich, 2008). Por outro lado, incluir a variável 'ordem de apresentação dos estímulos' no desenho do experimento requereria um número maior de ouvintes voluntários, no sentido de obter um conjunto quantitativamente satisfatório de respostas para cada ordem possível, que de fato pudessem ser analisadas posteriormente.

O instrumento elaborado para a coleta das percepções dos ouvintes acerca dos falantes – o chamado "formulário de percepção" (o terceiro ponto apresentado ao início desta seção) – é uma replicação parcial daquele proposto por Mendes (2016, 2018); das 7 escalas aqui utilizadas, 5 aparecem nos trabalhos daquele pesquisador que, a partir dos resultados que obteve sobre a análise dos efeitos de (CN) em percepções de gênero/sexualidade em vozes masculinas, discute o interesse de, em um mesmo experimento, utilizar tanto uma escala para percepção de masculinidade, quanto uma para percepção de efeminidade, no sentido de verificar se as respostas dos ouvintes em uma se correlacionam negativamente com as respostas na outra (Mendes, 2018, p. 67). No presente estudo, com foco nos efeitos da duração de /-s/ em percepções de gênero/sexualidade, seguimos a sugestão de Mendes (2018) e incluímos uma escala para "soa nada a muito masculino" e outra para "soa nada a muito gay".

Imagem 3: Escalas de diferenciais semânticos



Fonte: Elaboração dos autores

Além das escalas de diferenciais semânticos, o instrumento de pesquisa incluiu uma ficha de perfil social, na qual os respondentes nos forneceram as seguintes informações: seu sexo, sua escolaridade, sua idade, sua orientação sexual e se tem amigos gays etc. O experimento foi aplicado na plataforma *Google Form*s. Com vistas a evitar que um mesmo ouvinte interagisse com ambas as condições experimentais e a garantir aleatoriedade na atribuição de uma das duas condições a cada ouvinte que se voluntariasse a iniciar o experimento na plataforma, utilizou-se a ferramenta de Caixa Suspensa para que eles pudessem escolher, sem consciência disso, em qual das condições (C1 ou C2 - ver quadro 1) participariam do experimento.

O experimento foi divulgado em diversas redes sociais (*Instagram*, *Facebook* e *Whatsapp*) e acabaram participando dele um total de 122 voluntários: 62 na Condição 1 e 60 na C2. Como cada ouvinte forneceu respostas (nas escalas) para cada um dos quatro estímulos (numa certa condição), geraram-se 488 formulários preenchidos, no total. Participaram do estudo 56 homens e 66 mulheres; quanto à orientação sexual, 65 se identificaram como heterossexuais e 57 como homossexuais. Em relação à idade, os participantes foram agrupados em duas faixas etárias (20 a 34 anos e 36 a 54 anos). Quanto à escolaridade, 41 ouvintes possuíam ensino médio e 81 ensino superior. Essas variáveis – sexo, idade e escolaridade – foram incluídas em razão de sua tradicional presença em estudos sociolinguísticos, bem como diante da hipótese de que podem influenciar padrões de percepção e julgamento linguístico.

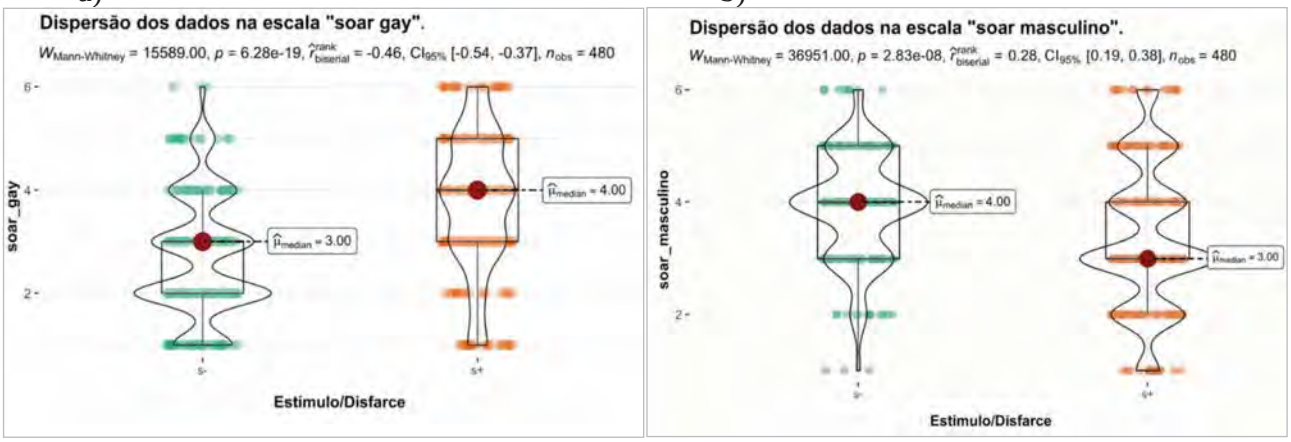
Adicionalmente, o questionário contemplou uma pergunta relativa à rede de sociabilidade dos ouvintes, especificamente quanto à quantidade de amigos gays, distinguindo-se entre aqueles que relataram possuir poucos amigos gays (48 participantes) e aqueles que indicaram possuir muitos amigos gays (74 participantes). A inclusão dessa variável baseia-se na hipótese de que maior exposição cotidiana a falantes gays poderia ampliar a familiaridade dos ouvintes com elementos linguísticos que, no repertório perceptual dos ouvintes, podem ser mobilizados na construção de julgamentos de masculinidade e sexualidade.

4. Análise dos dados

As respostas fornecidas pelos ouvintes foram analisadas no programa R (R Core Team, 2025). Para criar os gráficos que se apresentam a seguir, junto de testes estatísticos de significância, utilizou-se o *ggstatsplot* (Patil, 2021). Inicialmente, verifica-se a dispersão geral das respostas dos ouvintes nas escalas de percepção de “soar gay” e “soar masculino”.

Gráfico 1: Dispersão geral dos dados de percepção – respostas dos ouvintes nas escalas "soar gay" e "soar masculino"

a) b)



Fonte: Elaboração dos autores

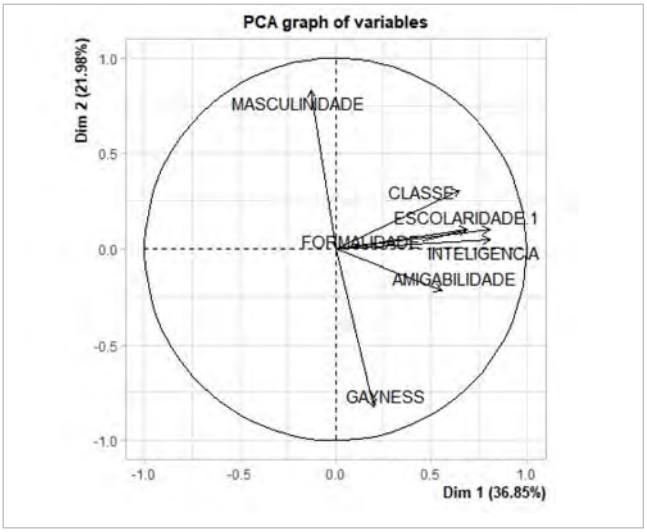
Na gráfico 1, tanto ao lado esquerdo (gráfico 1a – respostas para "soar gay") quanto ao direito (gráfico 1b – respostas para "soar masculino"), observam-se diferenças significativas entre ouvir os falantes (em conjunto) no disfarce s- ou s+. No geral, os falantes foram ouvidos como homens que soam "mais gay" no disfarce s+ (dispersão de respostas à direita, no par de respostas em gráfico 1a). Já no que toca às respostas na escala para "soar masculino", os falantes, no geral, foram ouvidos como homens que soam "menos masculinos" no mesmo disfarce s+. Essa diferença é confirmada pelo resultado do teste estatístico apresentado abaixo do título de cada painel[5]. O valor *p* inferior a 0,05 indica diferença estatisticamente significativa entre as condições, mostrando que os julgamentos atribuídos aos dois tipos de estímulo não são intercambiáveis.

Essa inspeção inicial sugere que as respostas dos ouvintes que participaram do experimento, nas duas escalas ora em foco, potencialmente se correlacionam entre si, negativamente: s+ leva à percepção de que os falantes, globalmente, soam "*mais* gay" e "*menos* masculinos". No interesse de testar esta e outras possíveis correlações entre as respostas dadas pelos ouvintes nas escalas, realizou-se também uma Análise de Componentes Principais (ACP) do pacote FactorMineR (Husson; Le; Pages, 2017), que indicou dois componentes que explicam a maior parte da variação das respostas.

A gráfico 2 ilustra esses dois componentes e indica a correlação negativa entre as respostas nas escalas "soar masculino" e "soar gay".

[5] Utilizou-se o teste de Mann-Whitney, versão não paramétrica do *teste t de Student* para amostras independentes, apropriado para comparações em desenhos experimentais *between-subject* como o deste estudo, pois compara as distribuições a partir das medianas em vez das médias e não exige normalidade estrita.

Gráfico 2: Gráfico das correlações entre respostas nas diferentes escalas disponibilizadas aos ouvintes no formulário de percepção



Fonte: Elaboração dos autores

Tal correlação negativa entre as respostas das duas escalas também é indicada na tabela 1 (índice 0.85 para "soar gay" e -0.84 para "soar masculino" – na última coluna, correspondente ao segundo componente principal, CP2). Dessa forma, no geral, quando os ouvintes perceberam as vozes dos falantes como vozes que soam "menos masculinas", também as perceberam como vozes que soam "mais gay". Esses resultados se equiparam aos de outros trabalhos (ver, por exemplo, Mendes, 2018; Sene, 2022, 2024) e podem ser interpretados como um estereótipo sobre vozes masculinas – no nosso caso, em associação à variável duração de /-s/.

Tabela 1: Correlações entre as respostas nas sete escalas de diferenciais semânticos

(Análise de Componentes Principais com rotação Promax)

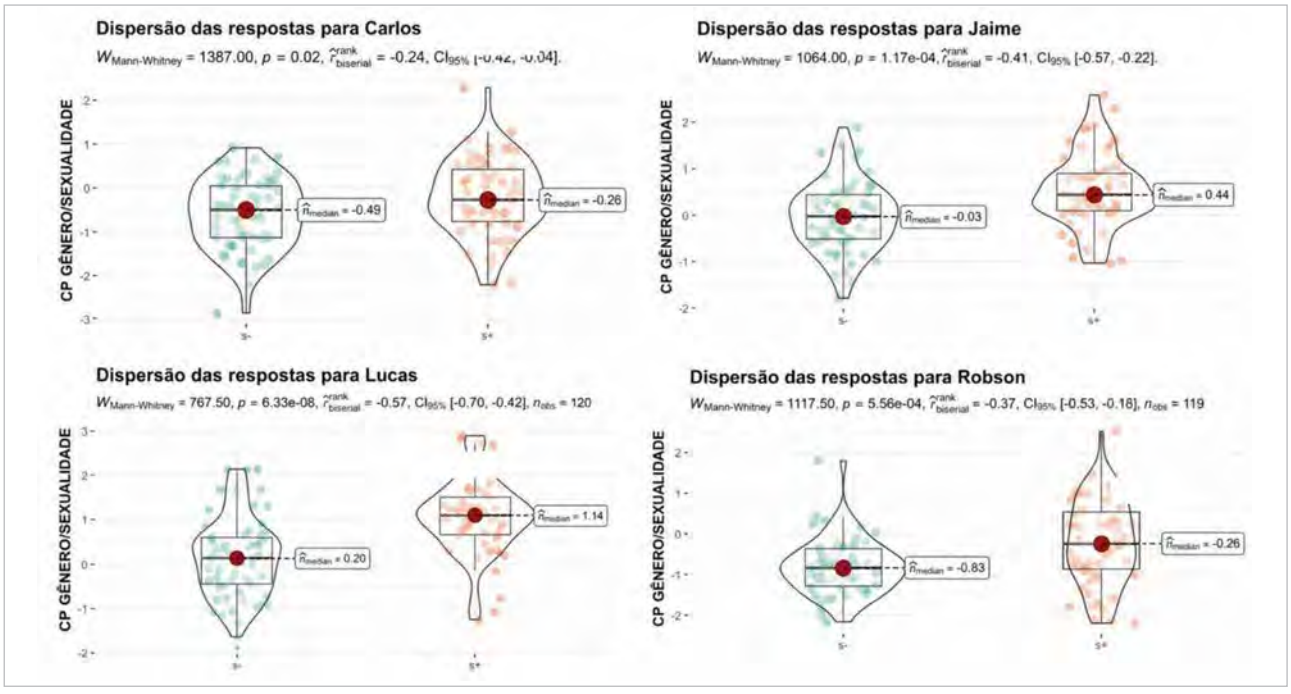
	CP1 – Competência	**CP2 – Gênero/Sexualidade**
Escolarizado	**0.82**	0.04
Amigável	0.50	0.31
Gayness	0.05	**0.85**
Formalidade	0.70	0.02
Inteligência	**0.81**	0.10
Classe	**0.69**	-0.19
Masculinidade	0.02	**-0.84**
Eigenvalue	2.57	1.55
% de variância	37%	22%
% Acumulativa	37%	59%

Fonte: Elaboração dos autores

A tabela 1 apresenta os escores dos dois componentes principais retidos na ACP. Foram considerados apenas carregamentos iguais ou superiores a 0,70, conforme a literatura da área, e os demais foram descartados (Malhotra, 2001). O primeiro componente denomina-se "Competência" em virtude do fato de que, dele, participam as respostas nas escalas de escolaridade, inteligência e formalidade (em correlação positiva): no geral, quando os falantes foram percebidos como mais escolarizados, também foram avaliados como mais inteligentes e mais formais. Devido ao nosso presente foco em percepções de gênero/sexualidade associadas à duração de /-s/, não desenvolvemos análises ulteriores acerca deste primeiro componente.

Como as respostas nas escalas para percepção de "soar masculino" e "soar gay" integram um componente no conjunto de respostas dadas pelos ouvintes, em vez de analisá-las individualmente, no sentido de verificar os padrões para cada um dos falantes, utilizamos os scores gerados pela ACP, para o CP2 Gênero/Sexualidade, como variável dependente (em vez, portanto, dos valores correspondentes às escalas de 6 pontos, apresentadas aos ouvintes). Assim, enquanto nos gráficos do gráfico 1, temos os valores correspondentes às escalas propriamente (1 a 6) no eixo y, nos gráficos a seguir temos valores de -2 a 2, correspondentes aos scores gerados pela ACP. Interessa examinar as respostas para cada um dos quatro falantes no sentido de verificar se o padrão geral que vimos anteriormente - s+ leva à percepção geral de que os falantes soam "mais gay" e "menos masculinos" - se revela também para cada um dos quatro falantes.

Gráfico 3: Dispersão das respostas dos ouvintes para cada falante, nos disfarces s+ e s- (eixo *y* = *scores* da ACP)



Fonte: Elaboração dos autores

Embora se verifique variação nas respostas para cada falante, o padrão geral anteriormente verificado se reatesta nas respostas para Carlos, Jaime, Lucas e Robson, quando analisadas individualmente: cada um deles é percebido como alguém que soa "menos masculino/mais gay" (o componente Gênero/Sexualidade) quando ouvido no disfarce s+ (dispersão de respostas à direita, em cada par). A significância estatística da diferença nas respostas baseadas em s+ e s- para cada falante é confirmada pelo teste de Mann-Whitney, com valores $p < 0{,}005$.

Campbell-Kibler (2007, 2011) e Levon (2014) mostram, como vimos anteriormente, que ouvintes não respondem a um único traço isolado, mas a configurações de pistas que evocam estereótipos de tipos de homens, como o "gay educado e afeminado" *versus* o "heterossexual másculo e pouco escolarizado". Nesse quadro, a associação encontrada neste experimento entre s+ e avaliações de "mais gay/menos masculino", em contraste com o s- alinhado à persona masculina hegemônica, é compatível com um repertório estereotípico já documentado: a variante s+ reforça, para os ouvintes, um estilo enquadrado como menos masculino conforme as normas de masculinidade dominante, enquanto a variante s- indicia um modelo de "homem padrão" previsto nos modelos ocidentais do que se entende por "ser/ soar masculino".

O fato de que cada um dos quatro falantes é percebido na mesma direção – "menos masculinos/ mais gay" – no disfarce s+ se apresenta como uma evidência forte de que a duração de /-s/ funciona como índice relativamente estável de gênero/sexualidade. As diferenças de mediana e de tamanho de efeito entre eles – com Jaime e Lucas recebendo *scores* mais altos para "soar gay", enquanto a percepção de como soam Carlos e Robson é menos afetada pela manipulação – mostram que o índice não atua de modo uniforme. Esses contrastes sugerem que o s+ se combina, em cada voz, com outros traços segmentais e prosódicos próprios do indivíduo, de modo que alguns falantes já "se situam" (previamente às manipulações digitais referentes a /s/, na construção do experimento) mais proximamente à persona "homem mais masculino", enquanto outros se aproximam mais de uma persona mais compatível com estereótipos acerca do que significa, sociolinguisticamente, "soar gay" (Levon, 2007; Podesva, 2007).

Esses achados dialogam com estudos perceptuais que identificam a duração de /s/ (bem como o contorno de *pitch*) como detalhes acústicos relevantes na avaliação do que soa gay e/ou masculino (Levon, 2007; Callaway, 2021). Nesse sentido, a variável duração de /s/, pelos resultados que aqui se relatam, ilustra bem a ideia de que as formas linguísticas fazem parte de um sistema semiótico socialmente complexo (Eckert, 2012), no qual pequenos deslocamentos acústicos podem reorganizar a leitura que os ouvintes fazem do falante.

Com o objetivo de examinar os efeitos da duração da coda /-s/ na percepção de gênero e sexualidade em vozes masculinas em suas relações com variáveis sociais que identificam os ouvintes, ajustou-se um modelo linear de efeitos mistos que toma como variável dependente os escores do segundo componente principal. Esse modelo inclui como efeitos fixos o disfarce (s- *vs.* s+) em interação com: o falante, o sexo do ouvinte, sua orientação sexual, sua escolaridade, sua faixa etária, sua rede de amigos (poucos ou muitos amigos gays), bem como ouvinte como efeito aleatório.

Tabela 2: Resumo do Modelo de Regressão Linear para o CP2 Gênero e Sexualidade (Efeitos Mistos – função lmer)

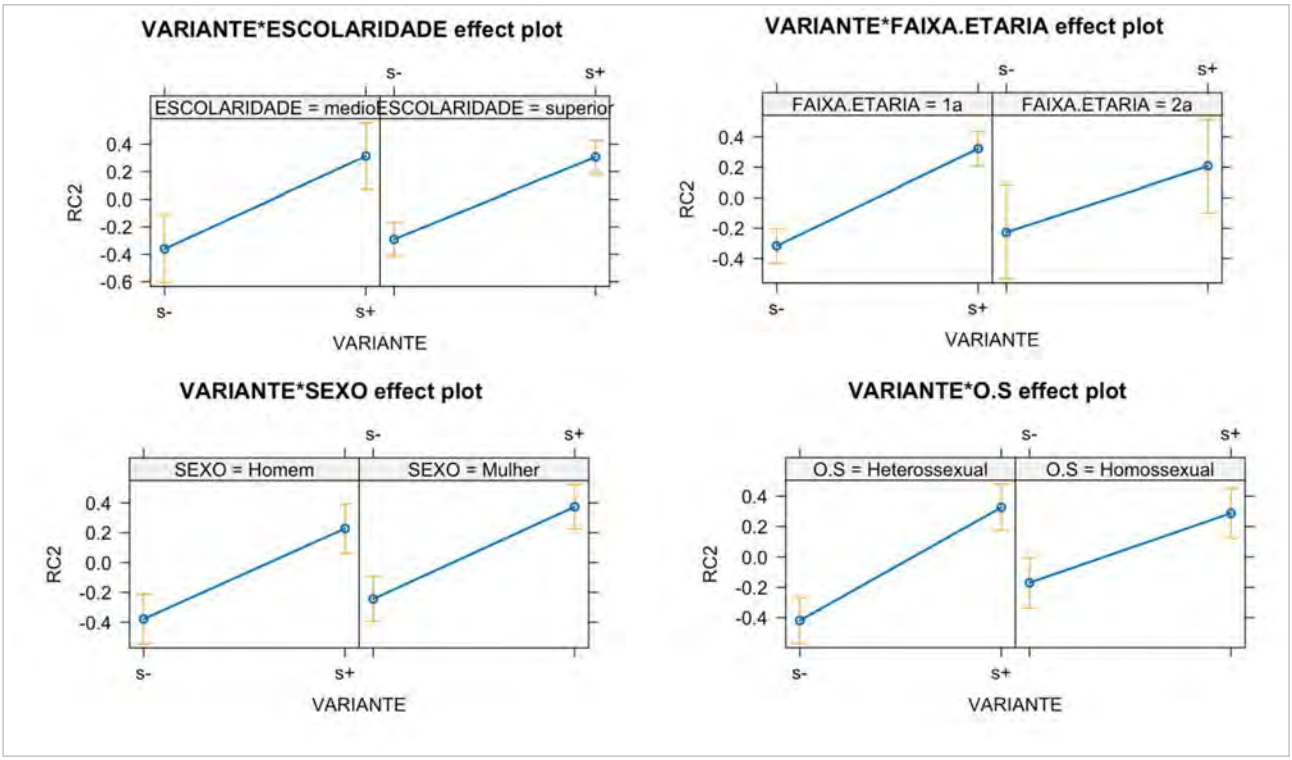
Preditores	Estimativas	Intervalo de confiança	p
(Intercept)	-0.93	-1.30 – -0.55	**<0.001**
[s+]	0.73	0.23 – 1.22	**<0.004**
falante [Jaime]	0.59	0.29 – 0.89	**<0.001**
falante [Lucas]	0.84	0.55 – 1.14	**<0.001**
falante [Robson]	-0.21	-0.50 – 0.09	0.164
Sexo [Feminino]	0.13	-0.10 – 0.36	0.253
Escolaridade [superior]	0.07	-0.20 – 0.34	0.616
Faixa etária [2a]	0.09	-0.24 – 0.42	0.594
Amigos gays [sim.poucos]	0.19	-0.05 – 0.43	0.118
Orientação Sexual [Homossexual]	0.25	0.02 – 0.48	**0.037**
[s+] × falante [Jaime]	0.18	-0.25 – 0.62	0.404
[s+] × falante [Lucas]	0.49	0.06 – 0.92	**0.027**
[s+] × falante [Robson]	0.21	-0.20 – 0.62	0.314
[s+] × Sexo [Feminino]	0.01	-0.30 – 0.32	0.943
[s+] × Escolaridade [superior]	-0.08	-0.44 – 0.29	0.681
[s+] × Faixa etária [2a]	-0.20	-0.64 – 0.24	0.365
[s+] × Amigos gays [sim.poucos]	-0.38	-0.70 – -0.06	**0.019**
[s+] × Orientação sexual [Homossexual]	-0.29	-0.60 – 0.02	0.071

Fórmula: CP2 ~ /s/*FALANTE + /s/*SEXO + /s/*ESCOLARIDADE + /s/*FAIXA.ETARIA + /s/*AMIGOS.GAYS + /s/*ORIENTAÇÃO SEXUAL + (1| OUVINTE)

Fonte: Elaboração dos autores

Esse modelo mostra um efeito geral da variável duração de /s/ na percepção dos ouvintes sobre gênero/sexualidade: s+, relativamente a s-, leva à percepção de que os falantes soam menos masculinos e mais gay. Esse efeito se mantém (comparativamente aos padrões apresentados anteriormente) mesmo com a inclusão das variáveis relativas ao ouvinte. Por outro lado, tais variáveis sociais – sexo, escolaridade, faixa etária e rede de amigos gays – não têm efeito, elas próprias, nas respostas que integram o CP2, com exceção da orientação sexual do ouvinte: aqueles que se identificam como gays, em comparação àqueles que se identificam como heterossexuais, tendem a atingir escores mais altos no componente Gênero/Sexualidade. Em outras palavras, no geral, ouvintes gays perceberam os falantes como "menos masculinos/mais gay" – mas esse efeito independe do disfarce (s+ ou s-): a interação entre orientação sexual e duração de /s/ não atinge significância estatística.

Gráfico 4: Gráficos de efeitos da variável /-s/ nas respostas de diferentes grupos de ouvintes (de acordo com seu sexo, sua escolaridade, sua faixa etária e sua orientação sexual)

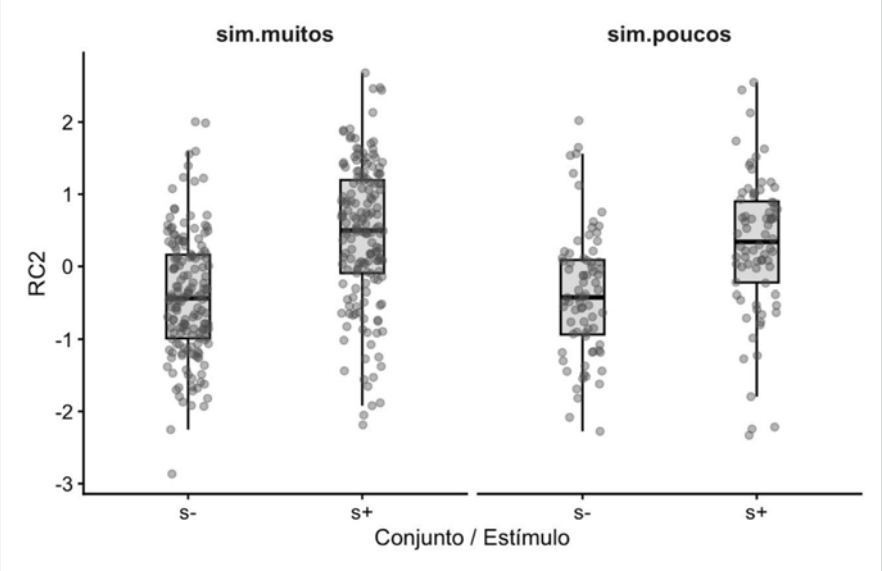


Fonte: Elaboração dos autores

No que toca às demais interações – entre a variável linguística e as outras variáveis sociais que identificam o ouvinte, vemos que os efeitos de s+ *vs.* s- também não dependem do sexo, da escolaridade e da faixa etária do ouvinte. Isso reforça o padrão geral que vimos anteriormente: a avaliação de que os falantes soam "menos masculinos/mais gay" no disfarce s+ é compartilhada pelos variados grupos de ouvintes: homens e mulheres, mais e menos escolarizados, mais jovens e mais velhos. Os gráficos de efeitos com linhas paralelas (para cada grupo de ouvintes) e intervalos de confiança sobrepostos (gráfico 4) indicam isto: que as respostas de homens e mulheres, mais e menos escolarizados, mais jovens e mais velhos, ouvintes gays e heterossexuais não variam significativamente a depender do disfarce s+ ou s-.

O modelo de regressão prevê ainda uma interação significativa entre duração de /-s/ e a variável "amigos gays". A interação [s+] × "poucos amigos gays" apresenta efeito negativo e estatisticamente significativo (β = −0,38; p = 0,019; IC 95% = −0,70 a −0,06), comparativamente ao seu *intercept* ([s-] x muitos amigos gays). O gráfico 4 auxilia na interpretação do modelo, uma vez que ilustra a dispersão das respostas para a variável "amigos gays" a depender do estímulo ouvido.

Gráfico 5: Dispersão dos escores para Gênero/Sexualidade (CP2) de acordo com a duração de /-s/ final (disfarce s- *vs.* s+), para ouvintes que relatam ter muitos ou poucos amigos gays em seus círculos de amizades



Fonte: Elaboração dos autores

O gráfico 5 mostra que, tanto para ouvintes que relatam ter muitos amigos gays (par à esquerda) quanto para aqueles que relatam ter poucos amigos gays (par à direita), os escores do CP2 (gênero/sexualidade) são mais altos para s+ e relativamente mais baixos para s-. Entretanto, a diferença entre respostas estimuladas por s- e s+ é maior nas respostas dos ouvintes que dizem ter muitos amigos gays. É isto que a interação significativa entre "/-s/ x amigos gays" (na penúltima linha da tabela 2) indica: comparativamente aos ouvintes que relatam ter poucos amigos gays, aqueles que têm muitos amigos gays em suas redes sociais percebem os falantes disfarçados de s+ no experimento como homens que soam menos masculinos/mais gay mais vezes.

Cabe considerar que ouvintes que relataram possuir muitos amigos gays possivelmente experimentam, em sua vida social, maior exposição à variação na produção da coda /-s/ final, especialmente no que se refere à duração do segmento, enquanto recurso estilístico empregado por homens gays – um aspecto que, até o momento, não foi explorado de forma sistemática em estudos de produção, menos ainda por estudos de produção que se valem de observação etnográfica de práticas sociolinguísticas. Por outro lado, nossos dados perceptuais sugerem que ouvintes mais familiarizados com práticas sociolinguísticas de homens que se identificam como gays tendem a associar, com maior frequência, a realização alongada s+ à percepção de que os falantes soam menos masculinos/mais gay.

5. À guisa de conclusão

Este estudo objetivou investigar se a duração da coda /-s/ em final de palavra, em português brasileiro, exerce efeito na percepção sociolinguística de gênero e sexualidade em vozes masculinas. Os dados fornecem a clara resposta de que sim: s+ experimentalmente leva à percepção de que as vozes de quatro homens soam menos masculinas e mais gays, tanto em análises globais quanto no exame de cada falante, individualmente. Esse efeito se mantém mesmo quando se examinam os efeitos de variáveis sociais relativas aos ouvintes, o que indica que a associação entre duração de /s/ e percepções de gênero/sexualidade é amplamente compartilhada no repertório perceptual do conjunto de ouvintes que participaram do experimento. Nesse sentido, os achados do presente estudo ampliam evidências reportadas na literatura sociolinguística: a duração de /s/ constitui um recurso simbólico relevante na construção de significados sociais relacionados a gênero e sexualidade, em várias línguas, inclusive no português brasileiro.

Interessa reiterar a importância da generalidade desse efeito. De um lado, os resultados indicam que a influência da duração de /-s/ não varia em função do sexo, da escolaridade, da faixa etária ou da orientação sexual dos ouvintes, o que sugere que essa variável linguística opera, da perspectiva social, como um índice relativamente estável e amplamente difundido no repertório perceptual dos falantes. Tal estabilidade aponta para uma significação social compartilhada, cuja interpretação não se restringe a subgrupos específicos, mas circula de forma generalizada entre as categorias sociais de ouvintes que participaram do experimento. Adicionalmente, nossos dados reforçam a ideia, defendida em estudos característicos da terceira onda da sociolinguística, de que variantes linguísticas podem funcionar como recursos semióticos integrantes de campos de significados sociais potenciais – e não como marcadores categóricos e invariáveis de identidades sociais.

Já a interação estatisticamente significativa entre duração da coda final /-s/ e a variável social "rede de amigos gays" sugere que o efeito de s+ na percepção de que falantes soam mais gay/menos masculinos não é homogênea entre os ouvintes, do ponto de vista de sua inserção social e de suas experiências de convivência com homens gays – algo ainda não explorado em termos de práticas sociolinguística e ainda pouco explorado na literatura sobre percepção sociolinguística. Tal como vimos no modelo resumido na tabela 2 e ao final da seção anterior, o efeito negativo para s+ no caso de ouvintes que afirmam ter "poucos amigos gays" indica que, nesse grupo, o alongamento do segmento não aciona com a mesma força os significados sociais associados a soar menos masculino e mais gay, comparativamente ao que se observa para ouvintes que se identificam como pessoas que têm muitos amigos gays.

Em termos teóricos, os resultados do experimento aqui em foco ilustram o conceito de indicialidade de segunda ordem (Silverstein, 2003; Eckert, 2012, 2018): s+ não "significa" "soa gay" de forma direta ou essencial, mas participa de um campo indicial no qual determinados padrões acústicos são reiteradamente associados a estilos menos hegemônicos de masculinidades, especialmente na experiência social dos ouvintes de cuja rede social participam "muitos amigos gays".

Empiricamente, este estudo insere o português brasileiro no conjunto de línguas nas quais já se documentaram efeitos perceptuais da variável /-s/ em percepções de gênero e sexualidade, como o inglês, o espanhol, o dinamarquês, e variedades do afrikaans. Trata-se, portanto, de um avanço para a sociolinguística da percepção no Brasil, ainda caracterizada por lacunas nesse domínio específico de investigação.

Ao mesmo tempo, os resultados aqui apresentados permitem retomar um problema central ainda em aberto, que motivou a própria concepção deste experimento: a relação entre os efeitos perceptuais da duração de /-s/ e os efeitos previamente documentados acerca da concordância nominal (CN) em percepções de gênero e sexualidade (Mendes, 2016, 2018). Embora o presente estudo traga evidências acerca dos efeitos de /-s/ nessas percepções, permanece a questão sobre até que ponto os efeitos observados em trabalhos anteriores dizem respeito à concordância propriamente, à realização mais curta ou mais longa de /-s/ ou ainda à interação entre ambas. Trabalhos futuros poderão se dedicar a examinar se (CN) e /-s/ se potencializam mutuamente ou se os efeitos de uma dessas variáveis podem ser mascarados ou bloqueados pelos da outra.

Referências

BASSI, Alessandra; SEARA, Izabel Christine. A produção das fricativas alveolar, ápico-alveolar e palato-alveolar em coda silábica no PB e no PE. *Letras de Hoje*, Porto Alegre, v. 52, n. 1, pp. 77-88, 2017.

BEKKER, Ian; LEVON, Erez. The embedded indexical value of /s/-fronting in Afrikaans and South African English. *Linguistics*, Berlin, v. 55, n. 5, p. 1109-1139, 2017. DOI: https://doi.org/10.1515/ling-2017-0022.

BIASIBETTI, Ana Paula Correa da Silva. *Produção e percepção das fricativas sibilantes em Porto Alegre/RS e Florianópolis/SC*. 2018. Tese (Doutorado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2018.

BOERSMA, Paul; WEENINK, David. *Praat*: doing phonetics by computer. 2014

BUCHOLTZ, Mary; HALL, Kira. *Theorizing identity in language and sexuality research*. Language in Society, Cambridge, v. 33, n. 4, pp. 469-515, 2004.

CALLAWAY, Alec. Power mediates the processing of gender during sibilant categorization. *Laboratory Phonology*, v. 12, n. 1, pp. 1-36, 2021.

CAMERON, Deborah. Performing gender identity: young men's talk and the construction of heterosexual masculinity. *In*: JOHNSON, Sally; MEINHOF, Ulrike H. (ed.). *Language and masculinity*. Oxford: Blackwell, 1997. p. 47-64.

CAMPBELL-KIBLER, Kathryn. *Listener perceptions of sociolinguistic variables*: the case of (ing). 2006. 264 f. Tese (Doutorado em Linguística) – Stanford University, Stanford, 2006.

CAMPBELL-KIBLER, Kathryn. Accent, (ING), and the social logic of listener perceptions. *American Speech*, Durham, v. 82, n. 1, p. 32-64, 2007.

CAMPBELL-KIBLER, Kathryn. The nature of sociolinguistic perception. *Language Variation and Change*, Cambridge, v. 21, n. 1, pp. 135-156, 2009. DOI: https://doi.org/10.1017/S0954394509000052.

CAMPBELL-KIBLER, Kathryn. Intersecting variables and perceived sexual orientation in men. *American Speech*, Durham, v. 86, n. 1, pp. 52-68, 2011.

CRIST, Sean. Duration of onset consonants in gay male stereotyped speech. *University of Pennsylvania Working Papers in Linguistics*, Philadelphia, v. 4, n. 3, pp. 53-70, 1997.

DRAGER, Katie. *Experimental research methods in sociolinguistics*. London: Bloomsbury Academic, 2018.

ECKERT, Penelope. Variation and the indexical field. *Journal of Sociolinguistics*, Oxford, v. 12, n. 4, pp. 453-476, 2008.

ECKERT, Penelope. Three waves of variation study: the emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, Palo Alto, v. 41, pp. 87-100, 2012.

ECKERT, Penelope. *Meaning and linguistic variation*: the third wave in sociolinguistics. Cambridge: Cambridge University Press, 2018.

ECKERT, Penelope; MCCONNELL-GINET, Sally. Think practically and look locally: language and gender as community-based practice. *Annual Review of Anthropology*, Palo Alto, v. 21, pp. 461-490, 1992.

HALL, Kira. Lip service on the fantasy lines. *In*: HALL, Kira; BUCHOLTZ, Mary (ed.). *Gender articulated: language and the socially constructed self*. New York: Routledge, 1995. pp. 183-216.

HALL, Kira. Intertextual sexuality: parodies of class, identity, and desire in liminal Delhi. *Journal of Linguistic Anthropology*, Arlington, v. 15, n. 1, pp. 125-144, 2005.

HALL, Kira; O'DONOVAN, Veronica. Shifting gender positions among Hindis-peaking hijras. *In*: BERKLEY WOMEN AND LANGUAGE GROUP (org.). *Gender and belief systems: Proceedings of the fourth Berkeley Women and Language Conference*. Berkeley: University of California, 1997. pp. 152-167.

HAUPT, Carine. As fricativas [s], [z], [ʃ] e [ʒ] do português brasileiro. *Estudos Linguísticos*, Araraquara, v. 36, n. 1, pp. 37-46, 2007.

HEFFERNAN, Kevin. Evidence from HNR that /s/ is a social marker of gender. *Toronto Working Papers in Linguistics*, Toronto, v. 23, pp. 71-84, 2004.

HUSSON, François; LÊ, Sébastien; PAGÈS, Jérôme. *Exploratory multivariate analysis by example using R*. 2. ed. Boca Raton: CRC Press, 2017.

KENT, R. D.; READ, C. *Análise Acústica da Fala*. São Paulo – SP: Editora Cortez, 2015. Tradução de Alexsandro Meireles. 2015.

LAMBERT, W. E.; HODGSON, R. C.; GARDNER, R. C.; FILLENBAUM, S. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, v. 60, n. 1, pp. 44-51, 1960.

LEVON, Erez. Hearing "gay": prosody, interpretation and the affective judgments of men's speech. *American Speech*, Durham, v. 81, n. 1, pp. 56-78, 2006.

LEVON, Erez. Sexuality in context: variation and the sociolinguistic perception of identity. *Language in Society*, Cambridge, v. 36, n. 4, pp. 533-554, 2007.

LEVON, Erez. Categories, stereotypes and the linguistic perception of sexuality. *Language in Society,* Cambridge, v. 43, n. 5, pp. 539-566, 2014.

MACK, Sara. Phonetic variation and perception of sexual orientation in Caribbean Spanish. *In*: Levon, E.; Mendes, R. *Language, sexuality and power*: Studies in intersectional sociolinguistics. Oxford University Press. 2016.

MAEGAARD, Marie; PHARAO, Nicolai. /s/ variation and perceptions of male sexuality in Denmark. *In*: LEVON, Erez; MENDES, Ronald Beline (ed.). *Language, sexuality and power*: intersectional studies in sociolinguistics. Oxford: Oxford University Press, 2016. pp. 88-104.

MALHOTRA, Naresh K. *Pesquisa de marketing*: uma orientação aplicada. 3. ed. Porto Alegre: Bookman, 2001.

MENDES, Ronald Beline. 'Nonstandard plural noun phrase agreement as an index of masculinity'. *In*: Levon, E. and Mendes, R. B. (eds) *Language, sexuality, and power*: studies in intersectional sociolinguistics. Oxford: Oxford University Press, 2016, pp. 105-130.

MENDES, Ronald Beline. *Percepção e performance de masculinidades: efeitos da concordância nominal de número e da pronúncia de /e/ nasal*. 236 f. Tese de Livre Docência. Universidade de São Paulo, 2018.

MENDES, Ronald Beline; OUSHIRO, Lívia. O paulistano no mapa sociolinguístico brasileiro. *Alfa: Revista de Linguístic*a, São Paulo, v. 56, n. 3, pp. 973-1001, 2012.

MCDONOUGH, Kim; TROFIMOVICH, Pavel. *Using priming methods in second language research*. New York: Routledge, 2008.

OLIVEIRA, Flávio Ricardo Medina de. *Análise acústica de fricativas e africadas produzidas por japoneses aprendizes de português brasileiro*. 2011. Dissertação (Mestrado em Letras) – Universidade Federal do Paraná, Curitiba, 2011.

OUSHIRO, Lívia. *Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo*. 2015. Tese (Doutorado em Linguística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2015.

PATIL, Indrajeet. Visualizations with statistical details: the 'ggstatsplot' approach. *Journal of Open Source Software*, v. 6, n. 61, 3167, 2021. DOI: https://doi.org/10.21105/joss.03167.

PLECK, Joseph H.; SONENSTEIN, Freya L.; KU, Leighton C. Masculinity ideology: its impact on adolescent males' heterosexual relationships. *Journal of Social Issues*, Malden, v. 49, n. 3, pp. 11-29, 1993. DOI: https://doi.org/10.1111/j.1540-4560.1993.tb01166.x

PLECK, Joseph H.; SONENSTEIN, Freya L.; KU, Leighton C. Attitudes toward male roles among adolescent males: a discriminant validity analysis. *Sex Roles*, v. 30, n. 7-8, pp. 481-501, 1994. DOI: https://doi.org/10.1007/BF01420798.

PODESVA, Robert J. Phonation type as a stylistic variable: the use of falsetto in constructing a persona. *Journal of Sociolinguistics*, v. 11, n. 4, pp. 478-504, 2007. DOI: https://doi.org/10.1111/j.1467-9841.2007.00334.x.

PODESVA, Robert J. Constraints on the social meaning of released /t/. *Language Variation and Change*, Cambridge, v. 27, n. 3, pp. 387-420, 2015.

R CORE TEAM. *R*: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2025.

SANTOS, Wendel Silva dos. *Percepções sociolinguísticas acerca da variação subjuntivo/indicativo em São Luís e São Paulo*. 2020. Tese (Doutorado em Linguística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2020.

SENE, Marcus Garcia de. *A percepção sociolinguística de gênero e sexualidade: efeitos da duração de /s/ e do pitch médio*. 2022. 216 f. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista "Júlio de Mesquita Filho", Araraquara, 2022.

SENE, Marcus Garcia de. "Eu soo gay ou masculino?" – tarefa de julgamento sociolinguístico de gênero e sexualidade: uma investigação sociofonética do *pitch* médio. *Acta Scientiarum. Language and Culture*, v. 46, n. 2, p. e71959, 2 dez. 2024.

SILVERSTEIN, Michael. Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, Amsterdam, v. 23, n. 3-4, pp. 193-229, 2003.

SILVA, Thaïs Cristófaro; SEARA, Izabel; SILVA, Adelaide; RAUBER, Andreia Schurt; CANTONI, Maria. *Fonética acústica*: os sons do português brasileiro. São Paulo: Contexto, 2019.

SMYTH, R.; JACOBS, G.; ROGERS, H. Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society*, v. 32, n. 3, pp. 329-350, 2003.

SMYTH, Ron; ROGERS, Henry. Do gay-sounding men speak like women? *Toronto Working Papers in Linguistics*, Toronto, v. 27, pp. 129-144, 2008.

STUART-SMITH, Jane; TIMMINS, Claire; WRENCH, Alan. Sex and gender differences in Glaswegian /s/. *In*: I*NTERNATIONAL CONGRESS OF PHONETIC SCIENCES*, *15*., 2003, Barcelona. Proceedings [...]. Barcelona: Universitat Autònoma de Barcelona, 2003. pp. 1851-1854

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72236

Pajubá como laboratório sociolinguístico no século XXI

Pajubá as a sociolinguistic laboratory for the twenty-first century

Renato Régis Barroso[1]

Yonara Cristina de Souza dos Santos[2]

RESUMO

Este artigo propõe compreender o Pajubá como um laboratório sociolinguístico privilegiado para a investigação da linguagem no século XXI. Em oposição às leituras que o reduzem à condição de gíria, código restrito ou repertório marginal, defende-se que o Pajubá constitui um sistema linguístico socialmente regulado, cuja organização permite observar, de forma integrada, processos de variação, performatividade, indexicalidade, legitimidade linguística e resistência discursiva. Fundamentado na Sociolinguística Variacionista, especialmente nas contribuições de William Labov e Anthony Julius Naro, em diálogo com a Linguística Queer e a Análise de Discurso Crítica, o estudo parte da hipótese de que o Pajubá amplia as possibilidades analíticas da Sociolinguística contemporânea ao evidenciar a indissociabilidade entre língua, corpo, identidade e poder. A pesquisa adota abordagem qualitativa, baseada em questionário eletrônico de caráter descritivo e entrevistas semiestruturadas realizadas com cinquenta participantes LGBTQIAPN+ distribuídos entre cinco metrópoles brasileiras. A interpretação do *corpus* articula pressupostos da Sociolinguística Variacionista, da Linguística Queer e da Análise de Discurso Crítica, privilegiando os sentidos atribuídos pelos participantes às suas práticas linguísticas. Os resultados demonstram que o Pajubá apresenta regularidades de uso, mecanismos de alternância sensíveis ao contexto interacional e elevado potencial de construção de pertencimento social, funcionando simultaneamente como repertório linguístico, prática performativa e tecnologia de resistência ao preconceito linguístico e à violência simbólica. Conclui-se que o Pajubá constitui um espaço privilegiado para compreender as transformações da linguagem contemporânea e reafirma a necessidade de ampliar os horizontes epistemológicos da Sociolinguística por meio de repertórios historicamente marginalizados.

PALAVRAS-CHAVE: Pajubá. Laboratório Sociolinguístico. Sociolinguística Variacionista.

ABSTRACT

This article proposes an understanding of Pajubá as a privileged sociolinguistic laboratory for investigating language in the twenty-first century. Contrary to perspectives that reduce it to slang, a restricted code or a marginal linguistic repertoire, the study argues that Pajubá constitutes a socially regulated linguistic system whose organization enables the integrated observation of linguistic variation, performativity, indexicality, linguistic legitimacy, and discursive resistance. Grounded in Variationist Sociolinguistics, particularly the contributions of William Labov and Anthony Julius Naro, and developed in dialogue with Queer Linguistics and Critical Discourse Analysis, the study is based on the hypothesis that Pajubá broadens the analytical possibilities of contemporary Sociolinguistics by demonstrating the inseparability of language, body, identity, and power. The research adopts a qualitative approach based on a descriptive online questionnaire and semi-

[1] Instituto Federal de Educação, Ciência e Tecnologia do Amazonas (IFAM), renato.regis@ifam.edu.br, https://orcid.org/0009-0000-2727-0148.

[2] Universidade Federal do Amazonas (UFAM), yonara.criss@gmail.com, https://orcid.org/0000-0003-1961-8911.

Recebido em 20/01/2026 - Aceito em 10/07/2026

structured interviews conducted with fifty LGBTQIAPN+ participants from five Brazilian metropolitan areas. The *corpus* was interpreted through the combined theoretical frameworks of Variationist Sociolinguistics, Queer Linguistics, and Critical Discourse Analysis, with particular attention to the meanings participants attribute to their own linguistic practices. The findings demonstrate that Pajubá displays systematic patterns of use, context-sensitive mechanisms of linguistic variation, and a strong capacity to construct social belonging, functioning simultaneously as a linguistic repertoire, a performative practice, and a technology of resistance against linguistic prejudice and symbolic violence. The study concludes that Pajubá constitutes a privileged site for understanding contemporary transformations in language and reinforces the need to broaden the epistemological horizons of Sociolinguistics through the study of historically marginalized linguistic repertoires.
KEYWORDS: Pajubá. Sociolinguistic Laboratory. Variationist Sociolinguistics.

1. Introdução

Há repertórios linguísticos que apenas circulam. Outros, no entanto, condensam em sua própria existência os conflitos, as disputas e as transformações que atravessam uma sociedade. O Pajubá pertence à segunda categoria. Forjado nas margens, entre experiências de violência simbólica, criatividade coletiva e resistência cotidiana, esse repertório não emerge como simples conjunto de expressões identitárias, tampouco como um código circunstancial empregado por pessoas LGBTQIAPN+. Trata-se de uma prática linguística produzida historicamente por sujeitos que precisaram reinventar a linguagem para construir reconhecimento, proteção e pertencimento em contextos marcados pela vigilância social. Sua permanência ao longo de décadas demonstra que sua vitalidade não decorre do acaso, mas da capacidade de responder às necessidades comunicativas de uma comunidade que transformou a linguagem em espaço de memória, afeto e sobrevivência. Ainda assim, o Pajubá continua sendo frequentemente reduzido, tanto no imaginário social quanto em parte da literatura acadêmica, à condição de gíria, jargão ou curiosidade lexical. Essa leitura obscurece sua organização interna e impede que ele seja reconhecido como objeto capaz de contribuir para questões centrais da Linguística contemporânea.

É justamente dessa lacuna que emerge o presente estudo. Mais do que defender o Pajubá como objeto legítimo da investigação linguística, propõe-se compreendê-lo como um laboratório sociolinguístico para pensar a linguagem no século XXI. A noção de laboratório empregada neste artigo não possui caráter metafórico, mas epistemológico. Assim como um laboratório científico possibilita observar fenômenos em funcionamento, o Pajubá reúne, em uma única prática linguística, processos que ocupam posição central nas discussões contemporâneas da Sociolinguística: variação, indexicalidade, alternância contextual, performatividade, construção identitária, circulação de repertórios, ideologias linguísticas e resistência discursiva. Longe de representar um caso excepcional, ele concentra processos que caracterizam inúmeras práticas linguísticas contemporâneas e permite observá-los em permanente interação. Nesse sentido, estudar o Pajubá significa deslocar a atenção da língua abstrata para a linguagem efetivamente produzida em situações concretas de uso, reafirmando um dos princípios fundamentais da tradição sociolinguística.

Essa perspectiva ancora-se na Sociolinguística Variacionista, particularmente nas contribuições de William Labov (1972) e Anthony Julius Naro (1992), que compreenderam a variação como parte constitutiva da competência linguística dos falantes. Ao reconhecer que diferentes escolhas linguísticas respondem a condicionamentos sociais, discursivos e pragmáticos, essa tradição rompe definitivamente com a ideia de uma língua homogênea e invariável. Entretanto, os desafios colocados pela linguagem no século XXI exigem ampliar esse horizonte interpretativo. Por essa razão, este estudo estabelece diálogo com a Linguística Queer e com a Análise de Discurso Crítica, compreendendo que as escolhas linguísticas não apenas refletem estruturas da língua, mas produzem identidades, negociam pertencimentos e materializam relações de poder. Nessa articulação teórica, o Pajubá deixa de ser visto como repertório periférico para tornar-se um espaço privilegiado de observação das relações entre linguagem, corpo e sociedade.

Sob essa perspectiva, a circulação do Pajubá revela um sistema regulado por convenções compartilhadas entre seus falantes. Seu emprego não ocorre de maneira aleatória, mas é condicionado pela avaliação constante do contexto de interação, do interlocutor e das consequências sociais produzidas por determinadas escolhas linguísticas. A variação deixa de significar instabilidade para evidenciar competência sociolinguística; a alternância entre repertórios revela estratégias de adaptação comunicativa; e a performatividade manifesta-se como prática cotidiana por meio da qual sujeitos constroem pertencimentos, estabelecem alianças e enfrentam processos de invisibilização. O Pajubá evidencia, assim, que compreender a linguagem implica compreender também as formas pelas quais comunidades historicamente marginalizadas produzem conhecimento, elaboram mecanismos de proteção simbólica e transformam experiências sociais em práticas discursivas.

Diante desse cenário, este artigo procura responder à seguinte questão: de que maneira o Pajubá pode ser compreendido como um laboratório sociolinguístico no século XXI e o que sua análise revela sobre os processos de variação, performatividade, legitimidade e resistência que caracterizam a linguagem em uso? Parte-se da hipótese de que o Pajubá reúne, em um único repertório linguístico, fenômenos frequentemente estudados de forma fragmentada pela Sociolinguística, permitindo analisá-los de maneira integrada. O objetivo geral consiste em compreender como sua organização linguística e seus usos sociais ampliam as possibilidades analíticas da Sociolinguística contemporânea. Para isso, o estudo busca descrever seus padrões de organização e variação, investigar os mecanismos de alternância contextual empregados pelos falantes, analisar os processos de performatividade e construção identitária produzidos por seu uso e interpretar, à luz da Análise de Discurso Crítica, os discursos de legitimidade, estigmatização e resistência que atravessam sua circulação.

A relevância deste trabalho manifesta-se em três dimensões complementares. No plano teórico, amplia o campo da Sociolinguística ao propor o Pajubá como um objeto capaz de reunir diferentes fenômenos linguísticos em uma mesma prática discursiva, fortalecendo o diálogo entre a Sociolinguística Variacionista, a Linguística Queer e a Análise de Discurso Crítica. No plano epistemológico, questiona critérios históricos de legitimidade científica que determinaram quais repertórios poderiam ser reconhecidos como objetos da Linguística, reafirmando que toda prática

linguística sistemática constitui conhecimento digno de descrição e interpretação. No plano social, evidencia que investigar a linguagem significa investigar também os modos pelos quais sujeitos constroem pertencimentos, enfrentam estigmas e produzem formas de resistência por meio da palavra. Ao integrar o dossiê dedicado às contribuições da Sociolinguística no século XXI, este artigo presta homenagem ao legado de Anthony Julius Naro ao reafirmar que levar a sério a variação implica, necessariamente, levar a sério os sujeitos que a produzem. O Pajubá deixa, assim, de ocupar a periferia dos estudos linguísticos para assumir a condição de um laboratório privilegiado, a partir do qual é possível compreender, em toda a sua complexidade, como a linguagem varia, produz identidades, negocia poderes e reinventa formas de existência.

2. Referencial teórico

2.1. O Pajubá como laboratório sociolinguístico: variação, performatividade e construção de identidades

Durante grande parte do século XX, a Linguística foi orientada pela busca de uma língua abstrata, homogênea e idealizada, concebida como um sistema autônomo e relativamente independente das condições concretas de uso. Nesse paradigma, a heterogeneidade era frequentemente interpretada como imperfeição, desvio ou ruído em relação a uma norma considerada legítima. A emergência da Sociolinguística Variacionista representa uma ruptura decisiva com essa tradição ao demonstrar que a variação constitui um princípio organizador do próprio funcionamento da língua. William Labov inaugura essa mudança de perspectiva ao afirmar que "a variabilidade é parte integrante do sistema linguístico e não pode ser ignorada sem empobrecer a análise" (Labov, 1972, p. 188). A língua deixa, então, de ser compreendida como uma entidade homogênea e passa a ser analisada como prática social permanentemente atravessada por condicionamentos históricos, culturais e interacionais. A variação deixa de representar instabilidade para constituir evidência da competência linguística dos falantes. Esse deslocamento epistemológico amplia significativamente o objeto da Linguística e cria as condições para que repertórios historicamente marginalizados possam ser reconhecidos como espaços legítimos de investigação científica.

No contexto brasileiro, essa mudança de paradigma encontra importante desenvolvimento nos trabalhos de Anthony Julius Naro, cuja contribuição ultrapassa a descrição estatística da variação para demonstrar que as chamadas regras variáveis integram o próprio conhecimento linguístico dos falantes. Segundo o autor, "as regras probabilísticas fazem parte do conhecimento linguístico do falante, sendo acionadas de forma sistemática em contextos específicos" (Naro, 1992, p. 23). Em diálogo com Scherre (2005), essa perspectiva evidencia que as escolhas linguísticas não decorrem de improvisação nem de ausência de regras, mas de competências construídas socialmente e internalizadas pelos sujeitos ao longo de suas experiências de interação. A língua em uso organiza-se por condicionamentos linguísticos, históricos e sociais que orientam continuamente as decisões dos falantes. Esse pressuposto constitui um dos pilares deste estudo, pois permite compreender o Pajubá

como repertório linguístico estruturado, aprendido na convivência comunitária e mobilizado de maneira estratégica em diferentes situações comunicativas. Seu funcionamento evidencia regularidades próprias, sensibilidade ao contexto e mecanismos de alternância que afastam definitivamente a ideia de um conjunto disperso de gírias ou expressões ocasionais.

Entretanto, limitar o Pajubá à condição de mais um objeto da Sociolinguística seria reduzir significativamente sua potência analítica. A hipótese defendida neste artigo é mais abrangente: o Pajubá constitui um verdadeiro laboratório sociolinguístico para compreender a linguagem no século XXI. A noção de laboratório não possui aqui sentido metafórico, mas epistemológico. Em um laboratório científico observam-se, simultaneamente, diferentes fenômenos que permitem compreender determinado objeto em funcionamento. De maneira semelhante, o Pajubá concentra, em uma única prática linguística, processos que tradicionalmente aparecem dispersos em diferentes objetos de investigação sociolinguística. Em seu funcionamento convivem mecanismos de variação, alternância contextual, indexicalidade, performatividade, circulação de repertórios, construção identitária, negociação de pertencimentos e resistência discursiva. Essa convergência confere ao Pajubá um estatuto singular dentro dos estudos da linguagem. Em vez de ilustrar isoladamente um conceito da Sociolinguística, ele permite observar, de forma integrada, diferentes dimensões que estruturam a linguagem em uso, tornando-se um espaço privilegiado para compreender os desafios teóricos colocados à Sociolinguística contemporânea.

Essa interpretação aproxima-se das contribuições de Eckert (2000, p. 128), para quem a variação linguística deve ser analisada a partir das comunidades de prática, deslocando o foco das grandes categorias sociais para os processos cotidianos de interação. A autora demonstra que as escolhas linguísticas participam da construção de estilos sociais e da produção de pertencimentos, deixando de funcionar apenas como indicadores de classe, idade ou escolaridade. Falar passa a significar posicionar-se socialmente. Essa perspectiva amplia o alcance da Sociolinguística Variacionista e permite compreender o Pajubá para além da dimensão lexical. Seu uso depende da participação em redes específicas de sociabilidade, da partilha de experiências coletivas e do reconhecimento de normas construídas no interior da própria comunidade LGBTQIAPN+. Cada escolha realizada nesse repertório produz efeitos de aproximação, cumplicidade, humor, acolhimento ou resistência, demonstrando que a variação linguística não pode ser dissociada das práticas sociais que lhe atribuem sentido.

Essa leitura é aprofundada pela noção de indexicalidade proposta por Silverstein (2003, p. 198), segundo a qual as formas linguísticas carregam significados sociais que ultrapassam sua função referencial. As palavras não apenas nomeiam objetos ou acontecimentos; elas apontam para identidades, valores, posições sociais e experiências compartilhadas. No Pajubá, esse processo manifesta-se de maneira particularmente intensa. Expressões, metáforas e formas de tratamento não apenas comunicam informações, mas ativam memórias coletivas, experiências de exclusão, práticas de solidariedade e modos específicos de reconhecimento comunitário. O repertório transforma-se, assim, em um sistema de índices sociais que permite aos sujeitos reconhecerem uns aos outros,

estabelecerem alianças e produzirem sentidos compartilhados sobre suas experiências. O que circula por meio do Pajubá não é apenas informação linguística, mas um complexo patrimônio simbólico continuamente reconstruído nas interações sociais.

Os desdobramentos contemporâneos da Sociolinguística aproximam-se, nesse ponto, das formulações da Teoria Queer ao compreender que a linguagem não apenas representa identidades previamente constituídas, mas participa ativamente de sua produção. Butler (1990, p. 67) rompe com as perspectivas essencialistas ao afirmar que gênero constitui efeito de práticas performativas reiteradas no tempo. A identidade deixa de ser compreendida como essência para tornar-se processo continuamente produzido por meio de ações, gestos e práticas discursivas. A linguagem ocupa posição central nesse movimento, pois oferece os recursos simbólicos necessários para nomear experiências, produzir pertencimentos e reorganizar relações sociais. Sob essa perspectiva, o Pajubá deixa de ser compreendido como consequência de identidades LGBTQIAPN+ previamente estabelecidas. Ao contrário, participa diretamente da produção dessas identidades, oferecendo aos falantes um repertório por meio do qual diferentes formas de existência podem ser continuamente performadas, negociadas e reconhecidas.

No contexto brasileiro, Melo (2024, p. 16) amplia essa discussão ao definir a Linguística Queer como um campo comprometido com a investigação dos discursos produzidos por corpos dissidentes de gênero e sexualidade, compreendendo-os como espaços privilegiados para questionar as normatividades sociais. Essa formulação desloca definitivamente a análise da linguagem para o campo das disputas epistemológicas, permitindo reconhecer sujeitos historicamente marginalizados não apenas como objetos, mas como produtores de conhecimento sobre suas próprias práticas linguísticas. Borba (2021) reforça essa perspectiva ao afirmar que a linguagem "não apenas expressa identidades, mas as produz performativamente na interação" (Borba, 2021, p. 63). O Pajubá evidencia esse processo de maneira exemplar. Cada ato de fala realizado nesse repertório constitui simultaneamente uma escolha linguística, um gesto identitário e uma prática política de reconhecimento coletivo. A linguagem transforma-se em espaço de criação de mundos possíveis, de negociação permanente das diferenças e de afirmação de formas de existência historicamente invisibilizadas.

É justamente essa capacidade de reunir, em uma única prática linguística, processos de variação, indexicalidade, performatividade e construção identitária que autoriza compreender o Pajubá como um laboratório sociolinguístico para a investigação da linguagem no século XXI. Sua relevância ultrapassa a descrição de um repertório específico e alcança questões centrais da Sociolinguística contemporânea, permitindo observar, em funcionamento, fenômenos que frequentemente aparecem dispersos em diferentes objetos de investigação. O Pajubá demonstra que compreender a linguagem implica compreender também os modos pelos quais sujeitos produzem pertencimentos, negociam identidades e reorganizam continuamente suas formas de existência por meio da palavra. Mais do que confirmar pressupostos clássicos da Sociolinguística Variacionista, ele amplia seus horizontes analíticos ao evidenciar que língua, corpo, experiência social e poder constituem dimensões indissociáveis da vida linguística contemporânea.

2.2. A *Gramática Descritiva do Pajubá* como evidência empírica de um laboratório sociolinguístico

As discussões desenvolvidas na seção anterior convergem para uma proposição fundamental deste artigo: compreender o Pajubá como laboratório sociolinguístico exige demonstrar empiricamente que sua organização ultrapassa a condição de repertório lexical disperso e constitui um sistema linguístico suficientemente estruturado para revelar, simultaneamente, diferentes fenômenos investigados pela Sociolinguística contemporânea. É precisamente nesse ponto que a *Gramática Descritiva do Pajubá* (Régis, 2025) assume papel central na presente discussão.

Mais do que representar um inventário lexical ou um exercício de sistematização terminológica, a obra constitui a primeira descrição linguística abrangente do Pajubá em seus diferentes níveis de organização, reunindo análises fonéticas, fonológicas, morfológicas, sintáticas, semânticas, pragmáticas e discursivas. Essa descrição desloca definitivamente o Pajubá do lugar da curiosidade sociocultural para o campo da investigação linguística rigorosa, oferecendo evidências de que seu funcionamento responde a regularidades observáveis e socialmente compartilhadas. Nesse sentido, a gramática deixa de ser apenas um produto da pesquisa para converter-se em uma demonstração concreta da hipótese central defendida neste estudo: a de que o Pajubá constitui um laboratório sociolinguístico capaz de condensar, em uma única prática linguística, múltiplos processos de variação, indexicalidade, performatividade e resistência.

Ao dialogar diretamente com os pressupostos da Sociolinguística Variacionista, a *Gramática Descritiva do Pajubá* confirma empiricamente a proposição de Anthony Julius Naro (1992, p. 144) segundo a qual as regras variáveis integram o conhecimento linguístico dos falantes. A análise do repertório evidencia que suas escolhas lexicais, morfológicas e pragmáticas não decorrem de improvisação ou instabilidade, mas de mecanismos sistemáticos aprendidos no interior da própria comunidade linguística. Assim como qualquer outra variedade, o Pajubá apresenta padrões recorrentes de formação lexical, processos relativamente estáveis de criação vocabular, mecanismos de adaptação semântica e estratégias pragmáticas condicionadas pelo contexto de interação. Sua circulação revela que os falantes compartilham um conhecimento coletivo que orienta tanto a produção quanto a interpretação dos enunciados. Essa constatação dialoga diretamente com Labov (1972, p. 71), ao reafirmar que a heterogeneidade não representa ausência de organização, mas uma das formas pelas quais a língua estrutura seu funcionamento. A *Gramática Descritiva do Pajubá* oferece, portanto, um campo empírico privilegiado onde as proposições clássicas da Sociolinguística podem ser observadas em funcionamento, reforçando a ideia de que a descrição da variação exige observar a língua onde ela efetivamente acontece: nas práticas sociais de seus falantes.

Não obstante, limitar a contribuição da obra à confirmação da sistematicidade do Pajubá significaria reduzir significativamente seu alcance epistemológico. Sua principal contribuição reside justamente em demonstrar que um repertório produzido nas margens da norma concentra processos que extrapolam a descrição estrutural da língua e alcançam dimensões identitárias, performativas e políticas. Cada unidade lexical descrita na gramática carrega consigo trajetórias históricas, relações de

pertencimento, estratégias de proteção e formas específicas de produzir reconhecimento comunitário. A descrição linguística passa, assim, a dialogar continuamente com a Linguística Queer, a Análise de Discurso Crítica e os estudos da performatividade, evidenciando que as regularidades observadas no Pajubá não podem ser dissociadas das experiências sociais de seus falantes. A gramática demonstra que estrutura e experiência não constituem esferas independentes, mas dimensões profundamente articuladas de um mesmo fenômeno linguístico. O laboratório sociolinguístico revela-se precisamente nessa capacidade de permitir que diferentes níveis de análise — fonologia, morfologia, pragmática, identidade, discurso e poder — sejam observados simultaneamente em um mesmo repertório. Essa perspectiva torna-se particularmente evidente quando o autor afirma:

> A Gramática do Pajubá é, portanto, o reflexo de uma coletividade que, mesmo historicamente silenciada, ressignificou a língua para narrar sua própria existência. Para alguns seria como uma 'minigramática', visto que tudo que chega para disputar com a gramática normativa é visto de forma ofensiva. Mas é importante reiterar que além de descritiva, nossa obra é aplicada (por isso esse tom mais subversivo). A gramática do Pajubá tem de ser assim, porque estamos falando de um povo que teve que se encolher para sobreviver, que se travestiu de linguagem cifrada para não morrer — e, ao mesmo tempo, é gigante em sua capacidade de criar, reinventar e performar mundos possíveis. (Régis, 2025, p. 71-72).

Essa passagem ultrapassa a função tradicional de um comentário metodológico e assume caráter epistemológico. Ao antecipar a crítica que poderia reduzir a obra à condição de "minigramática", o autor evidencia que o próprio conceito de gramática se encontra atravessado por disputas de legitimidade científica. A classificação depreciativa não decorre da estrutura do objeto descrito, mas da dificuldade histórica em reconhecer como dignas de sistematização linguística práticas produzidas por sujeitos socialmente marginalizados. Nessa perspectiva, a gramática deixa de representar apenas um exercício de descrição e transforma-se em um gesto de enfrentamento às ideologias linguísticas que condicionam quais repertórios merecem reconhecimento acadêmico. O diálogo com Bourdieu (1991, p. 78) torna-se inevitável: o que está em disputa não é apenas uma descrição gramatical, mas o próprio direito de determinados grupos produzirem conhecimento sobre suas formas de linguagem.

A metáfora do "travestir-se de linguagem cifrada para não morrer" constitui talvez a síntese mais contundente da dimensão sociolinguística do Pajubá. Ela evidencia que a organização do repertório não pode ser compreendida separadamente das condições históricas que produziram sua existência. A gramática registra estruturas linguísticas, mas essas estruturas emergem de experiências concretas de vigilância, exclusão, solidariedade e resistência. O Pajubá nasce como resposta à violência, mas permanece porque produz comunidade. Essa dimensão aproxima diretamente a obra das reflexões de Fairclough (2001) e Van Dijk (1998), ao demonstrar que a linguagem constitui simultaneamente prática social, espaço de disputa ideológica e instrumento de transformação das relações de poder. Descrever o Pajubá significa, portanto, descrever também as formas pelas quais sujeitos historicamente silenciados produziram estratégias discursivas capazes de garantir reconhecimento, pertencimento e continuidade histórica.

É justamente essa convergência entre descrição linguística, performatividade, indexicalidade, resistência e poder que transforma a *Gramática Descritiva do Pajubá* em muito mais do que um trabalho de descrição lexical. Ela representa a comprovação empírica de que o Pajubá reúne, em um único repertório, processos centrais para compreender a Sociolinguística contemporânea. Se o tópico anterior demonstrou teoricamente que a variação, a identidade e as relações de poder podem ser observadas nesse repertório, a gramática oferece a evidência concreta de que tais fenômenos se articulam de maneira sistemática e observável. O Pajubá deixa, assim, de ocupar a posição de objeto periférico para converter-se em um verdadeiro laboratório sociolinguístico do século XXI, no qual a descrição da língua coincide com a descrição das formas pelas quais sujeitos produzem identidades, negociam legitimidades, enfrentam ideologias e reinventam continuamente suas possibilidades de existência por meio da linguagem.

3. Metodologia

Este estudo adota uma abordagem qualitativa, por compreender que os fenômenos investigados não podem ser explicados apenas pela descrição de formas linguísticas ou pela quantificação de ocorrências, mas exigem a interpretação dos sentidos produzidos pelos sujeitos em suas práticas sociais. O Pajubá é compreendido, nesta pesquisa, como um repertório linguístico socialmente regulado, cuja organização resulta da articulação entre linguagem, identidade, performatividade e relações de poder, exigindo um percurso metodológico capaz de apreender essa complexidade em contextos reais de interação. Nessa perspectiva, o estudo insere-se no paradigma interpretativista das Ciências Humanas, segundo o qual a realidade social é construída pelos próprios sujeitos e compreendida a partir dos significados atribuídos às suas experiências (Denzin; Lincoln, 2018). Tal posicionamento metodológico permite compreender o Pajubá não como um conjunto isolado de itens lexicais, mas como prática linguística situada, atravessada por processos históricos, sociais e discursivos que produzem pertencimento, proteção e resistência.

O delineamento metodológico articula pressupostos da Sociolinguística Variacionista, da Linguística Queer e da Análise de Discurso Crítica, possibilitando observar o Pajubá simultaneamente como sistema linguístico, prática performativa e espaço de disputa simbólica. Da Sociolinguística Variacionista incorporam-se, sobretudo, as contribuições de William Labov (1972) e Anthony Julius Naro (1992), que compreendem a variação como parte constitutiva da competência linguística dos falantes. Da Linguística Queer, especialmente a partir de Butler (1990), Borba (2021) e Melo (2024), adota-se a compreensão da linguagem como prática performativa por meio da qual identidades são produzidas, negociadas e ressignificadas nas interações sociais. Por sua vez, a Análise de Discurso Crítica, representada principalmente por Fairclough (2001) e Van Dijk (1998), fornece os instrumentos necessários para compreender como as escolhas linguísticas se articulam às relações de poder, aos processos de estigmatização e às estratégias de resistência que atravessam a circulação do Pajubá.

A produção do *corpus* envolveu cinquenta participantes autodeclarados LGBTQIAPN+, distribuídos igualmente entre cinco metrópoles brasileiras: São Paulo, Rio de Janeiro, Salvador, Belém e Manaus. A escolha dessas cidades fundamentou-se em sua relevância histórica e sociocultural para a circulação do Pajubá, contemplando diferentes contextos urbanos das regiões Sudeste, Nordeste e Norte do país. Não se buscou representatividade estatística, mas diversidade sociolinguística suficiente para compreender distintas formas de apropriação, circulação e ressignificação desse repertório. Os participantes possuíam idade igual ou superior a dezoito anos, diferentes níveis de escolarização e trajetórias variadas de inserção em redes de sociabilidade LGBTQIAPN+. A seleção ocorreu por amostragem intencional, utilizando-se a técnica de bola de neve (*snowball sampling*), complementada pela divulgação do estudo em grupos digitais voltados à comunidade LGBTQIAPN+. Como critérios de inclusão, consideraram-se a autodeclaração como pessoa LGBTQIAPN+, o contato recorrente com o Pajubá e a disponibilidade para participar voluntariamente da pesquisa. Antes da realização das entrevistas, foi conduzida uma breve conversa exploratória destinada a confirmar a familiaridade dos participantes com o repertório, assegurando a pertinência do *corpus* analisado.

A produção dos dados ocorreu em duas etapas complementares. Inicialmente, os participantes responderam a um questionário eletrônico elaborado na plataforma Google Forms, composto por perguntas de caracterização sociodemográfica e por afirmações organizadas em escala Likert de cinco pontos, abordando aspectos relacionados ao reconhecimento do Pajubá, aos contextos de uso, à construção de pertencimento, ao preconceito linguístico, à legitimidade científica e à circulação social do repertório. Embora estruturado em escala Likert, esse instrumento foi utilizado exclusivamente com finalidade descritiva, contribuindo para organizar tendências gerais observadas entre os participantes e orientar a etapa seguinte da investigação, sem qualquer pretensão inferencial ou estatística.

Em consonância com os princípios da transparência científica e da reprodutibilidade da pesquisa, o questionário permanece disponível para consulta pública por meio do endereço eletrônico[3]. Na segunda etapa, foram realizadas entrevistas semiestruturadas, consideradas o principal instrumento de produção do *corpus* qualitativo. As entrevistas tiveram duração média entre trinta e quarenta e cinco minutos e ocorreram presencialmente ou por videoconferência, conforme a disponibilidade dos participantes. O roteiro contemplou questões relacionadas às trajetórias de contato com o Pajubá, aos contextos de uso, às estratégias de alternância linguística, às experiências de pertencimento, aos episódios de preconceito linguístico, às relações entre linguagem e identidade e às percepções acerca da legitimidade científica do repertório. Mediante autorização prévia, todas as entrevistas foram gravadas em áudio e posteriormente transcritas integralmente, preservando pausas, hesitações e demais marcas discursivas relevantes para a interpretação dos dados.

Após a transcrição, cada participante recebeu um código alfanumérico, preservando-se integralmente seu anonimato durante todas as etapas da investigação. A análise do *corpus* desenvolveu-se mediante articulação entre os referenciais teóricos adotados, buscando inicialmente

[3] Disponível em: https://docs.google.com/forms/d/e/1FAIpQLSellVbWCKvMLlPvxW42nNSnpO8upETwI_YWgpReDuZ-LJEwgw/viewform. Acesso em: 28 jul. 2026.

identificar padrões de variação, alternância contextual e regularidades de uso do Pajubá. Em seguida, examinaram-se os processos de performatividade, construção identitária e pertencimento produzidos pelas práticas linguísticas dos participantes. Finalmente, os fragmentos discursivos foram interpretados à luz da Análise de Discurso Crítica, privilegiando a compreensão das relações entre linguagem, poder, preconceito linguístico, resistência simbólica e legitimidade científica. As categorias analíticas foram construídas em diálogo permanente entre teoria e *corpus*. Inicialmente, definiram-se categorias derivadas do referencial teórico, como variação linguística, performatividade, indexicalidade, legitimidade e preconceito linguístico. Durante a leitura sucessiva das entrevistas, entretanto, emergiram categorias diretamente relacionadas às experiências narradas pelos participantes, entre elas "Pajubá como abrigo", "medo institucional", "família como espaço de vigilância", "humor como estratégia de resistência" e "circulação digital do repertório", permitindo que a interpretação permanecesse aberta às especificidades do *corpus*.

O rigor científico da pesquisa foi assegurado pela coerência entre o problema de pesquisa, os referenciais teóricos, os procedimentos metodológicos e a interpretação dos dados, bem como pela articulação entre os diferentes instrumentos de produção do *corpus*. As tendências identificadas no questionário eletrônico foram continuamente confrontadas com os sentidos produzidos nas entrevistas, fortalecendo a consistência analítica do estudo. Todos os participantes foram previamente informados sobre os objetivos da pesquisa, a utilização exclusivamente acadêmica dos dados e o direito de interromper sua participação a qualquer momento, mediante consentimento livre e esclarecido. Considerando a natureza sensível do objeto investigado, buscou-se desenvolver uma pesquisa comprometida não apenas com os princípios éticos formais, mas também com uma ética epistemológica, reconhecendo os participantes como sujeitos produtores de conhecimento sobre suas próprias práticas linguísticas. Por fim, reconhece-se que a natureza intencional da amostra e o recorte geográfico adotado não permitem generalizações estatísticas. Todavia, tais limites não comprometem os objetivos da investigação, uma vez que seu propósito consiste em compreender, em profundidade, os processos sociolinguísticos que fazem do Pajubá um laboratório privilegiado para a investigação da linguagem no século XXI.

4. Resultados e discussão

Os resultados apresentados nesta pesquisa são interpretados a partir da compreensão de que o Pajubá não constitui apenas um repertório lexical utilizado por sujeitos LGBTQIAPN+, mas um sistema linguístico em funcionamento, regulado por convenções compartilhadas, escolhas contextuais e processos permanentes de negociação identitária. Diferentemente de estudos que privilegiam exclusivamente a descrição quantitativa de atitudes linguísticas, esta investigação compreende que a principal contribuição do *corpus* reside nas práticas discursivas produzidas pelos próprios participantes. Por essa razão, os fragmentos das entrevistas constituem o eixo central da análise, enquanto os dados provenientes do questionário eletrônico são mobilizados apenas como recurso descritivo para evidenciar tendências recorrentes observadas ao longo da investigação. Essa opção

metodológica é coerente com a abordagem qualitativa adotada e com o objetivo deste estudo, que busca compreender de que maneira o Pajubá opera como um verdadeiro laboratório sociolinguístico para pensar a variação linguística no século XXI.

Nessa perspectiva, a análise foi organizada em torno de três grandes eixos interpretativos. O primeiro concentra-se na descrição do Pajubá como sistema linguístico estruturado, evidenciando regularidades, mecanismos de variação e formas de indexicalidade presentes no *corpus*. O segundo examina o repertório como prática performativa de produção de pertencimento, mostrando como as escolhas linguísticas participam da construção das identidades LGBTQIAPN+. O terceiro investiga os processos de estigmatização e resistência discursiva que atravessam a circulação do Pajubá em diferentes espaços sociais. Em todos os eixos, os fragmentos das entrevistas constituem a principal evidência empírica da pesquisa, sendo interpretados à luz da Sociolinguística Variacionista, da Linguística Queer e da Análise de Discurso Crítica.

4.1. O Pajubá como sistema linguístico em funcionamento

A primeira evidência produzida pelo *corpus* diz respeito ao reconhecimento do Pajubá como um repertório linguisticamente organizado. Ao longo das entrevistas, tornou-se recorrente a percepção de que o uso desse repertório não ocorre de maneira improvisada ou aleatória, mas segundo regularidades compartilhadas pelos próprios falantes. Essa constatação desloca o Pajubá do lugar tradicionalmente ocupado pelas "gírias" ou pelos códigos efêmeros e o aproxima daquilo que Labov (1972) e Naro (1992) compreendem como sistemas variáveis organizados por regras probabilísticas internalizadas. Em outras palavras, a variação observada não representa ausência de estrutura; ao contrário, constitui precisamente a manifestação do funcionamento da língua em uso. Essa percepção aparece de forma explícita na fala da participante E12, ao afirmar:

> "Não é bagunça, não. A gente sabe quando pode falar, com quem pode falar e até como falar. Se errar, o povo percebe." (E12, Salvador)

Esse fragmento sintetiza um dos principais argumentos deste artigo. Ao negar que o Pajubá seja "bagunça", o participante antecipa um discurso social historicamente dirigido ao repertório e produz um contradiscurso que reafirma sua legitimidade linguística. A expressão "a gente sabe" não se refere apenas ao conhecimento de determinadas palavras, mas à existência de uma competência sociolinguística compartilhada, capaz de orientar escolhas adequadas conforme interlocutor, contexto e finalidade comunicativa. O verbo "saber" passa a indexar pertencimento, experiência e domínio das regras internas do repertório, evidenciando aquilo que Naro (1992) denomina conhecimento linguístico internalizado.

Sob a perspectiva da Linguística Queer, esse conhecimento ultrapassa a dimensão estrutural. Saber utilizar o Pajubá significa também saber ocupar determinados espaços de sociabilidade, reconhecer marcadores identitários e compartilhar formas específicas de produção do pertencimento.

A competência linguística descrita pela participante é, simultaneamente, competência social. Não basta conhecer o léxico; é necessário compreender quando, como e diante de quem determinadas escolhas podem ser mobilizadas. Nesse sentido, a língua deixa de funcionar apenas como instrumento de comunicação e passa a operar como marcador de comunidade, reconhecimento e proteção coletiva.

A Análise de Discurso Crítica permite ampliar ainda mais essa interpretação. Ao afirmar que "se errar, o povo percebe", a participante revela a existência de um regime normativo próprio da comunidade. Diferentemente da representação externa que associa o Pajubá à ausência de regras, o discurso evidencia que há mecanismos internos de validação, correção e reconhecimento. O que muda não é a existência da norma, mas sua origem. Em vez de derivar da gramática normativa institucional, ela emerge da experiência coletiva dos próprios falantes. O fragmento desmonta, portanto, uma das principais ideologias linguísticas que historicamente justificaram a exclusão do Pajubá do campo científico: a crença de que apenas repertórios legitimados pelas instituições seriam capazes de apresentar regularidade linguística. Essa competência compartilhada também se manifesta na própria organização lexical do repertório. Durante as entrevistas, diversos participantes recorreram espontaneamente a expressões cujo funcionamento depende de conhecimento comunitário prévio, demonstrando que determinados itens do Pajubá possuem sentidos relativamente estáveis e socialmente compartilhados.

Quadro 1: Pajubá em uso

Fragmento	Fenômeno sociolinguístico observado	Interpretação
"Aquenda esse babado."	Especialização lexical	O verbo "aquendar" assume função pragmática específica, compreendida apenas por falantes competentes do repertório.
"Ela fez a egípcia."	Idiomatização	A expressão apresenta sentido não composicional, funcionando como unidade fraseológica cristalizada.
"Mapô!"	Africanismo lexical	O item preserva memória afro-diaspórica e produz indexicalidade identitária.

Fonte: Elaboração dos autores

O quadro evidencia que o funcionamento do Pajubá ultrapassa a simples acumulação de palavras isoladas. Cada item lexical participa de uma rede de relações semânticas, pragmáticas e discursivas cuja interpretação depende do compartilhamento de experiências socioculturais específicas. Esse comportamento confirma empiricamente a hipótese defendida ao longo deste artigo: o Pajubá constitui um repertório estruturado, organizado por regularidades próprias e permanentemente atualizado por seus falantes. Nesse sentido, o repertório confirma-se como verdadeiro laboratório sociolinguístico, permitindo observar, em tempo real, processos de variação, mudança, indexicalidade e construção identitária que dificilmente aparecem de forma tão explícita em variedades mais estabilizadas da língua portuguesa.

4.2. Pajubá, performatividade e produção de pertencimento

Se o primeiro eixo demonstrou que o Pajubá constitui um sistema linguístico organizado, o segundo desloca o olhar para uma dimensão igualmente decisiva: sua capacidade de produzir pertencimento social. O *corpus* evidencia que o repertório não funciona apenas como meio de comunicação entre sujeitos LGBTQIAPN+, mas como prática performativa por meio da qual identidades são continuamente produzidas, reconhecidas e legitimadas na interação. Essa constatação aproxima-se diretamente da concepção defendida por Butler (1990, p. 67), segundo a qual identidades não antecedem a linguagem; elas são reiteradamente produzidas por atos performativos. No caso do Pajubá, falar deixa de ser apenas transmitir informações para tornar-se um gesto de reconhecimento mútuo, uma forma de localizar o outro dentro de uma comunidade de experiências compartilhadas. Essa dimensão aparece de maneira particularmente sensível no fragmento produzido pelo participante E04:

> "Quando eu falo, eu sei que tô entre os meus. É outra respiração." (E04, Belém)

A escolha da metáfora "outra respiração" desloca imediatamente o Pajubá do campo lexical para o campo da experiência corporal. Respirar constitui uma das condições fundamentais da existência; ao associar o uso do repertório a uma mudança na própria respiração, o participante sugere que falar Pajubá produz uma sensação de alívio, segurança e reconhecimento impossível de ser alcançada em contextos marcados pela vigilância. O repertório deixa de ser apenas linguagem e passa a constituir ambiente afetivo.

Sob a perspectiva da Linguística Queer, esse enunciado ilustra aquilo que Borba (2021) descreve como a capacidade da linguagem de produzir identidades durante a interação. O pertencimento não aparece como consequência do grupo; é o próprio ato de falar que cria esse pertencimento. O Pajubá, portanto, não representa uma identidade previamente existente, mas participa ativamente de sua constituição performativa. Na perspectiva da Análise de Discurso Crítica, o fragmento também revela a existência de dois espaços discursivos distintos. O primeiro corresponde aos ambientes nos quais o sujeito precisa monitorar permanentemente sua fala. O segundo emerge quando ele afirma estar "entre os meus". Nesse espaço, desaparece a necessidade de vigilância constante, permitindo que o repertório funcione como tecnologia de cuidado coletivo. O discurso constrói, assim, uma oposição entre exterior e interior, exclusão e acolhimento, vigilância e liberdade. Outro fragmento amplia essa compreensão:

> "Nem sempre eu preciso perguntar se a pessoa é LGBT. Quando ela solta um 'babado', um 'mapô', eu já entendo que existe alguma conexão." (E18, Rio de Janeiro)

Esse excerto evidencia um fenômeno central da Sociolinguística contemporânea: a indexicalidade. Conforme Silverstein (2003, p. 196), determinadas formas linguísticas ultrapassam seu conteúdo referencial e passam a indexar pertencimentos, trajetórias sociais e posições identitárias. Nesse caso, palavras como 'babado' e 'mapô' deixam de funcionar apenas como itens lexicais. Elas operam como marcadores de reconhecimento comunitário. O significado deixa de estar apenas na palavra; passa a residir na rede de relações sociais que ela ativa. Esse comportamento linguístico pode ser sintetizado no quadro a seguir:

Quadro 2: Microlaboratório sociolinguístico: linguagem como marcador de pertencimento

Fragmento	Fenômeno observado	Interpretação sociolinguística
"É outra respiração"	Metáfora identitária	O Pajubá é representado como experiência corporal de pertencimento.
"Babado"	Indexicalidade	O item lexical funciona como marcador de reconhecimento entre interlocutores.
"Mapô"	Africanismo reindexado	O empréstimo lexical passa a produzir identidade queer contemporânea.

Fonte: Elaboração dos autores

O quadro demonstra que o funcionamento do Pajubá depende menos da definição dicionarizada de cada palavra do que da capacidade desses itens de produzir reconhecimento social. O repertório constitui uma rede de índices culturais cuja interpretação pressupõe experiência compartilhada. Essa característica reforça a hipótese defendida ao longo deste estudo de que o Pajubá deve ser compreendido como prática linguística situada, e não como simples inventário lexical. Outro aspecto recorrente nas entrevistas refere-se ao caráter afetivo do repertório.

> "Quando uma travesti mais velha fala comigo em Pajubá, parece que ela tá dizendo: 'você pode ficar tranquila aqui'." (E31, Salvador)

Esse fragmento desloca novamente a análise para uma dimensão pouco explorada pela Sociolinguística tradicional: a linguagem como prática de cuidado. O Pajubá deixa de operar apenas como marcador identitário para funcionar como dispositivo de acolhimento comunitário. A expressão "você pode ficar tranquila aqui" não aparece verbalizada pelo interlocutor; ela é inferida pelo próprio uso do repertório. O significado emerge precisamente da competência sociolinguística compartilhada entre os falantes.

Na perspectiva da Linguística Queer, esse fenômeno evidencia que determinadas escolhas linguísticas produzem condições específicas de existência social. Falar Pajubá torna-se um gesto de hospitalidade discursiva, por meio do qual sujeitos historicamente submetidos à violência simbólica constroem espaços temporários de proteção e reconhecimento. Além disso, o uso do Pajubá também

coloca em foco o questionamento da própria comunidade em usar o código linguístico como um recurso contra-hegemônico:

> "Se estudam o português errado dos ricos, por que não o nosso? Certo?" (E19, Rio de Janeiro)

Esse fragmento evidencia uma crítica direta às ideologias linguísticas hegemônicas. Ao inverter a lógica do "erro", o sujeito questiona os critérios de legitimidade científica e denuncia a seletividade epistemológica do campo. Na perspectiva da ADC, trata-se de um discurso contra-hegemônico que expõe a dimensão política da produção do conhecimento linguístico.

4.3. O Pajubá como prática de resistência: preconceito linguístico, vigilância e sobrevivência

Os dados produzidos nesta pesquisa evidenciam que o Pajubá não pode ser compreendido apenas como um repertório destinado à construção do pertencimento comunitário. Sua emergência histórica e sua permanência contemporânea estão profundamente relacionadas às múltiplas formas de violência simbólica dirigidas aos corpos LGBTQIAPN+. Em praticamente todas as entrevistas realizadas, o repertório aparece associado a experiências de vigilância, medo, discriminação e autocensura. Isso significa que o Pajubá não nasce apenas do desejo de criar uma linguagem própria; ele nasce, sobretudo, da necessidade de continuar existindo em espaços nos quais determinadas formas de falar, desejar e performar gênero permanecem socialmente sancionadas. Nesse sentido, compreender o Pajubá implica compreender também os mecanismos pelos quais a linguagem se converte em estratégia de sobrevivência diante do preconceito linguístico e da violência institucional. Essa dimensão emerge de forma particularmente expressiva no relato do participante E27:

> "No trabalho eu travo. Não é que eu não saiba falar. É que eu sei demais o risco que corremos." (E27, São Paulo)

O fragmento desloca imediatamente a interpretação da competência linguística para o campo das relações de poder. O verbo "travar" não descreve uma incapacidade comunicativa, mas um processo consciente de monitoramento discursivo. O participante afirma possuir domínio do repertório, porém opta por silenciá-lo diante da possibilidade de sofrer sanções sociais. O problema, portanto, não reside na língua, mas nas condições sociais que regulam sua circulação.

Sob a perspectiva da Sociolinguística Variacionista, esse comportamento confirma que a alternância entre repertórios linguísticos constitui uma estratégia sistemática de adaptação aos diferentes contextos de interação. Conforme defendem Labov (1972) e Naro (1992), as escolhas linguísticas realizadas pelos falantes respondem a condicionamentos sociais específicos, sendo parte integrante da competência comunicativa. No caso do Pajubá, entretanto, essa alternância assume uma

dimensão adicional: ela deixa de representar apenas adequação estilística e passa a funcionar como mecanismo de autoproteção.

A Análise de Discurso Crítica permite aprofundar essa interpretação. Fairclough (2001) demonstra que as instituições regulam não apenas comportamentos, mas também formas legítimas de falar. O silêncio relatado pelo participante revela precisamente esse processo. O sujeito não abandona o Pajubá porque deixa de reconhecê-lo como parte de sua identidade; ele o suspende temporariamente porque compreende que determinados espaços institucionais continuam operando segundo regimes discursivos que associam determinadas formas linguísticas à inadequação profissional, à falta de credibilidade ou ao desvio moral. O preconceito, portanto, não se dirige às palavras em si, mas aos sujeitos que essas palavras tornam visíveis. Essa dinâmica reaparece em outro fragmento:

> "Na frente da minha família eu penso duas vezes antes de falar qualquer palavra do Pajubá. Às vezes nem é preconceito direto. É medo de explicar quem eu sou." (E09, Manaus)

Nesse enunciado, o repertório passa a funcionar como índice identitário. O participante reconhece que determinadas palavras possuem capacidade suficiente para revelar pertencimentos sociais que nem sempre deseja tornar públicos. O Pajubá deixa de ser apenas um código linguístico e transforma-se em marcador de identidade, fazendo com que cada escolha lexical carregue consigo implicações afetivas, familiares e políticas. O silêncio, nesse contexto, constitui uma prática discursiva tão significativa quanto a própria fala.

Ao prisma da Linguística Queer, esse comportamento evidencia que linguagem e identidade são inseparáveis. Como argumenta Butler (1990), os sujeitos negociam continuamente sua inteligibilidade social por meio de performances reiteradas. O participante, ao controlar sua fala, controla também a forma como sua identidade poderá ser lida pelos interlocutores. O repertório, portanto, participa diretamente da produção da própria existência social.

Outro aspecto recorrente nas entrevistas refere-se ao processo de ressignificação do preconceito. Em vez de abandonarem o repertório diante da desqualificação externa, muitos participantes relatam exatamente o movimento contrário: quanto maior o estigma, maior o valor simbólico atribuído ao Pajubá no interior da comunidade:

> "Falam que é feio, que é vulgar..., mas é nosso. E isso dá força." (E33, Manaus)

Esse pequeno enunciado concentra uma das principais contribuições deste estudo. O adjetivo "nosso" funciona como marcador de apropriação simbólica. O estigma deixa de produzir apenas exclusão e passa a alimentar processos de fortalecimento identitário. O repertório torna-se patrimônio coletivo justamente porque foi historicamente rejeitado pelos espaços de poder. O que para determinados grupos representa inadequação linguística, para seus falantes converte-se em memória, pertencimento e resistência.

A lógica discursiva observada nesse fragmento aproxima-se daquilo que Van Dijk (1998) descreve como produção de contradiscursos. Os sujeitos deixam de aceitar passivamente as classificações impostas pelas ideologias dominantes e passam a produzir novas interpretações sobre sua própria linguagem. O preconceito deixa de definir o valor do repertório; passa a revelar apenas as relações de poder que procuram deslegitimá-lo. Esse movimento pode ser observado de maneira sintética no quadro a seguir:

Quadro 3: Microlaboratório sociolinguístico: resistência discursiva e circulação do Pajubá

Fragmento	Fenômeno observado	Interpretação
"No trabalho eu travo."	Monitoramento discursivo	Alternância linguística motivada por vigilância institucional.
"Penso duas vezes antes de falar."	Autocensura linguística	Gestão da identidade em contextos familiares.
"Mas é nosso."	Ressignificação do estigma	Apropriação simbólica do repertório como prática de resistência

Fonte: Elaboração dos autores

Os três fragmentos permitem observar que o funcionamento do Pajubá não pode ser explicado exclusivamente por fatores linguísticos internos. Cada escolha lexical encontra-se atravessada por relações de poder que regulam quem pode falar, onde pode falar e quais consequências acompanham esse ato. O repertório passa a funcionar, simultaneamente, como mecanismo de proteção, marcador de reconhecimento comunitário e estratégia de enfrentamento às ideologias linguísticas que historicamente procuraram reduzir sua legitimidade científica e social.

Essa interpretação encontra respaldo nas tendências identificadas pelo questionário eletrônico. Conforme sintetizado no quadro acima, observa-se elevada percepção simultânea de estigmatização social e valorização interna do Pajubá. Embora os participantes reconheçam que o repertório continua sendo alvo frequente de preconceito em diferentes espaços institucionais, atribuem-lhe elevado valor afetivo, identitário e político. Os dados quantitativos, portanto, não inauguram essa conclusão, mas corroboram aquilo que as entrevistas evidenciaram de maneira consistente: o Pajubá sobrevive justamente porque seus falantes recusam transformá-lo em objeto de vergonha. Ao contrário, convertem-no em patrimônio linguístico, memória coletiva e prática cotidiana de resistência.

Assim, os três eixos analisados ao longo desta seção convergem para uma mesma compreensão. O Pajubá constitui um sistema linguístico estruturado, socialmente regulado e profundamente atravessado por relações de poder. Sua variação não representa instabilidade; expressa competência. Sua performatividade não constitui mero recurso estilístico; produz pertencimento. Sua circulação desigual não evidencia fragilidade; revela os mecanismos pelos quais determinadas formas de linguagem continuam sendo controladas pelas instituições. É precisamente nesse ponto que o Pajubá se consolida como um laboratório sociolinguístico do século XXI: um espaço privilegiado para

observar, em tempo real, como língua, identidade, poder e resistência se entrelaçam na experiência concreta dos sujeitos. Estudá-lo, portanto, não significa apenas ampliar o inventário de objetos da Sociolinguística; significa repensar a própria disciplina a partir das margens que historicamente ela ainda escutou pouco.

4.4. O Pajubá como laboratório sociolinguístico no século XXI: síntese interpretativa dos resultados

Os resultados apresentados ao longo desta investigação permitem afirmar que o Pajubá ultrapassa amplamente a condição de repertório lexical utilizado por sujeitos LGBTQIAPN+ em situações específicas de interação. As evidências produzidas pela corpas demonstram que se trata de um sistema linguístico socialmente organizado, regulado por convenções compartilhadas e permanentemente atualizado pelos próprios falantes. Mais do que isso, revelam que sua circulação oferece à Sociolinguística contemporânea um espaço privilegiado para observar, simultaneamente, fenômenos de variação, indexicalidade, performatividade, preconceito linguístico, monitoramento discursivo e resistência simbólica. É precisamente essa convergência de processos que justifica compreendê-lo como um verdadeiro laboratório sociolinguístico do século XXI.

Ao longo da análise verificou-se que a variação observada no Pajubá confirma um dos princípios fundamentais da Sociolinguística Variacionista: a heterogeneidade não representa desorganização do sistema linguístico, mas constitui seu próprio modo de funcionamento. Conforme demonstraram Labov (1972) e Naro (1992), as escolhas linguísticas realizadas pelos falantes respondem a condicionamentos sociais, pragmáticos e discursivos relativamente estáveis. Os fragmentos analisados nesta pesquisa corroboram essa perspectiva ao evidenciarem que o repertório é acionado de maneira sensível ao contexto, ao interlocutor e aos diferentes graus de risco social. A alternância entre Pajubá e português hegemônico não decorre de improvisação, mas de uma competência sociolinguística refinada, construída na experiência cotidiana dos próprios sujeitos.

Não obstante, os dados também revelam que limitar a análise do Pajubá aos modelos clássicos da Sociolinguística seria insuficiente para compreender sua complexidade. A investigação evidenciou que cada escolha lexical produz simultaneamente efeitos identitários, afetivos e políticos que extrapolam a descrição variacionista tradicional. Nesse ponto, a Linguística Queer amplia significativamente o alcance interpretativo da pesquisa ao demonstrar que linguagem, corpo e identidade não constituem dimensões independentes. O Pajubá não apenas acompanha identidades dissidentes; participa ativamente de sua produção performativa. Cada palavra mobilizada pelos participantes reorganiza relações de pertencimento, estabelece fronteiras comunitárias e produz formas específicas de reconhecimento entre sujeitos historicamente submetidos à invisibilização social.

Ao mesmo tempo, a Análise de Discurso Crítica permitiu compreender que os conflitos observados em torno do Pajubá não se originam em sua estrutura linguística, mas nas relações de poder que regulam sua circulação. Os discursos que historicamente classificaram esse repertório como

"gíria", "erro", "linguagem vulgar" ou "código marginal" revelam menos características da língua e muito mais os mecanismos ideológicos responsáveis por hierarquizar determinadas formas de falar. O preconceito linguístico identificado ao longo das entrevistas não opera sobre palavras isoladas; dirige-se aos corpos que essas palavras tornam socialmente visíveis. Dessa forma, investigar o Pajubá significou investigar também os processos por meio dos quais a linguagem participa da produção da exclusão e, simultaneamente, da construção da resistência.

Essa articulação entre Sociolinguística Variacionista, Linguística Queer e Análise de Discurso Crítica constitui, talvez, uma das principais contribuições deste estudo. Em vez de tratar essas perspectivas como campos concorrentes, os resultados demonstram que elas se complementam na compreensão de fenômenos linguísticos contemporâneos. A Sociolinguística explica a organização da variação; a Linguística Queer evidencia sua dimensão performativa; a Análise de Discurso Crítica revela as relações de poder que condicionam sua circulação. O Pajubá emerge exatamente no ponto de encontro dessas três perspectivas, funcionando como objeto empírico capaz de colocar em diálogo tradições teóricas que, durante muito tempo, percorreram caminhos paralelos.

Os próprios participantes da pesquisa parecem reconhecer essa complexidade ao descreverem o repertório como lugar de proteção, memória, reconhecimento e cuidado coletivo. Em praticamente todos os relatos analisados, o Pajubá aparece menos como simples instrumento comunicativo e mais como prática cotidiana de existência. Essa constatação desloca significativamente o modo como a Linguística pode compreender seus próprios objetos de investigação. Se a língua é também espaço de produção de pertencimento, de negociação identitária e de enfrentamento das desigualdades, então a descrição linguística deixa de limitar-se às estruturas formais e passa necessariamente a incorporar as condições históricas, sociais e políticas que tornam determinadas formas de falar possíveis ou impossíveis.

Os dados analisados também evidenciam que o Pajubá permanece em permanente processo de transformação. Sua intensa circulação em ambientes digitais, redes sociais, plataformas audiovisuais e espaços culturais indica que o repertório continua ampliando seus domínios de uso, incorporando novas formas lexicais, reorganizando significados e produzindo diferentes modalidades de indexicalidade. Essa dinâmica confirma que o Pajubá não representa um objeto encerrado ou estabilizado, mas um repertório vivo, cujas transformações acompanham as próprias mudanças experimentadas pelas comunidades LGBTQIAPN+ brasileiras ao longo das últimas décadas. Nesse sentido, o laboratório sociolinguístico permanece aberto: cada nova interação produz novas possibilidades de análise, novos processos de variação e novas formas de resistência linguística.

É justamente essa vitalidade que torna o Pajubá particularmente relevante para a Sociolinguística do século XXI. Ao observá-lo em funcionamento, torna-se possível acompanhar, em um mesmo fenômeno linguístico, processos de mudança lexical, reorganização pragmática, construção performativa de identidades, circulação desigual dos repertórios, produção de ideologias linguísticas e elaboração de estratégias de sobrevivência discursiva. Poucos objetos contemporâneos reúnem, de

maneira tão evidente, tantas dimensões da linguagem em uso. Por essa razão, o Pajubá não constitui apenas mais um objeto possível para a Sociolinguística. Ele oferece condições privilegiadas para que a própria disciplina reflita criticamente sobre seus conceitos, seus métodos e, sobretudo, sobre os critérios historicamente utilizados para definir quais práticas linguísticas merecem ser reconhecidas como legítimas.

Assim, a análise desenvolvida neste artigo permite responder ao problema de pesquisa que orientou toda a investigação. O Pajubá configura-se como um sistema linguístico variável e performativo cuja descrição revela não apenas os limites das abordagens tradicionais da Sociolinguística, mas, principalmente, suas enormes possibilidades de renovação quando colocada em diálogo com perspectivas críticas contemporâneas. Nesse horizonte, compreender o Pajubá significa compreender que a língua vive precisamente onde as relações sociais se tornam mais intensas, mais desiguais e mais criativas. É nas margens, frequentemente transformadas em espaços de silêncio pela norma, que a linguagem revela com maior nitidez sua extraordinária capacidade de reinventar sujeitos, comunidades e modos de existir.

5. Conclusão

Este estudo partiu de uma pergunta aparentemente simples, mas epistemologicamente profunda: o que o Pajubá revela sobre a Sociolinguística do século XXI? Ao longo da investigação, tornou-se evidente que responder a essa questão significava deslocar o próprio olhar da Linguística sobre seus objetos tradicionais de estudo. Mais do que descrever um repertório produzido por sujeitos LGBTQIAPN+, este artigo procurou compreender como uma prática linguística historicamente marginalizada pode iluminar fenômenos centrais da linguagem contemporânea. Nesse percurso, o Pajubá deixou de ocupar o lugar de objeto periférico para assumir a condição de laboratório sociolinguístico privilegiado, capaz de evidenciar que a variação linguística, a performatividade, a indexicalidade e as relações de poder não constituem dimensões isoladas da linguagem, mas processos profundamente articulados na experiência cotidiana dos falantes.

Os resultados permitiram demonstrar que o Pajubá apresenta regularidades suficientes para ser compreendido como um repertório linguisticamente estruturado e socialmente regulado. As análises evidenciaram que sua circulação não ocorre de maneira aleatória, mas responde a condicionamentos pragmáticos, identitários e contextuais compartilhados pelos próprios falantes. A alternância entre o Pajubá e o português hegemônico revelou uma competência sociolinguística refinada, orientada pela avaliação permanente do interlocutor, do ambiente de interação e dos riscos implicados em cada escolha linguística. Longe de representar instabilidade ou improvisação, a variação observada confirma um dos princípios fundamentais da Sociolinguística Variacionista: a heterogeneidade constitui o próprio funcionamento da língua. Nesse sentido, o Pajubá reafirma que nenhuma prática linguística pode ser considerada cientificamente irrelevante apenas porque se desenvolve fora dos espaços tradicionalmente legitimados pela norma.

Ao mesmo tempo, esta pesquisa evidenciou que limitar a interpretação do Pajubá aos modelos clássicos da variação seria insuficiente para compreender sua complexidade. Os fragmentos analisados demonstraram que cada escolha lexical participa da construção performativa das identidades LGBTQIAPN+, produzindo pertencimento, reconhecimento e cuidado coletivo. O repertório não apenas comunica experiências; ele cria condições específicas de existência social. Falar Pajubá significa produzir vínculos, reconhecer trajetórias comuns e construir espaços simbólicos de acolhimento em contextos frequentemente marcados pela exclusão. A Linguística Queer mostrou-se, portanto, indispensável para compreender que determinadas formas de linguagem não apenas representam sujeitos, mas participam ativamente da produção de suas identidades e de seus modos de existir.

A Análise de Discurso Crítica ampliou ainda mais essa compreensão ao evidenciar que o preconceito dirigido ao Pajubá não decorre de características internas do repertório, mas das relações de poder que historicamente regulam quem pode falar, quais variedades podem circular com prestígio e quais práticas linguísticas permanecem associadas ao desvio. Os discursos que classificam o Pajubá como "gíria", "erro" ou "linguagem vulgar" não descrevem propriedades da língua; descrevem projetos de controle social sobre determinados corpos e determinadas formas de existência. O *corpus* revelou que o repertório é permanentemente atravessado por processos de vigilância institucional, autocensura, resistência e reapropriação simbólica. Nesse cenário, o Pajubá emerge simultaneamente como alvo do preconceito linguístico e como uma das mais sofisticadas estratégias de enfrentamento a esse mesmo preconceito.

Uma das contribuições mais relevantes desta investigação consiste justamente em demonstrar que o Pajubá permite colocar em diálogo diferentes tradições da Linguística contemporânea. A Sociolinguística Variacionista explica sua organização estrutural; a Linguística Queer evidencia sua dimensão performativa; a Análise de Discurso Crítica revela as relações de poder que condicionam sua circulação. Em vez de campos concorrentes, essas perspectivas mostraram-se complementares diante de um objeto cuja complexidade exige abordagens igualmente complexas. O Pajubá transforma-se, assim, em um espaço privilegiado para observar, simultaneamente, processos de mudança linguística, indexicalidade, construção identitária, resistência discursiva e negociação permanente entre linguagem e poder. Poucos repertórios contemporâneos oferecem condições tão fecundas para compreender a língua em seu funcionamento social mais intenso.

Ao defender o Pajubá como laboratório sociolinguístico, este artigo não propõe apenas a ampliação do inventário de objetos legítimos da Linguística. Propõe, sobretudo, uma ampliação do próprio horizonte epistemológico da disciplina. Se a Sociolinguística nasceu para demonstrar que toda língua varia, talvez o desafio do século XXI seja reconhecer que toda variação também produz efeitos políticos, afetivos e identitários. A língua não varia apenas porque diferentes grupos falam de maneiras distintas; ela varia porque diferentes sujeitos ocupam posições desiguais na estrutura social e transformam essa experiência em prática discursiva. O Pajubá evidencia com clareza que estudar a variação linguística implica, inevitavelmente, estudar as formas pelas quais os sujeitos negociam pertencimentos, enfrentam violências e reinventam possibilidades de existência.

Os resultados apresentados, entretanto, não encerram o debate. Ao contrário, evidenciam a amplitude de um campo investigativo ainda em construção. Permanecem abertas questões fundamentais relacionadas às transformações do Pajubá em ambientes digitais, às diferenças geracionais de uso, aos processos de interiorização do repertório, às suas variações regionais e às novas formas de circulação produzidas pelas redes sociais e pelas plataformas digitais. Também se mostram particularmente promissoras investigações sobre o funcionamento do Pajubá em contextos escolares, institucionais e jurídicos, espaços nos quais a vigilância linguística assume formas específicas e ainda pouco exploradas pela literatura. Da mesma maneira, a articulação entre o Pajubá e outros repertórios produzidos por populações historicamente marginalizadas — como comunidades indígenas, quilombolas e periféricas — constitui uma agenda de pesquisa capaz de ampliar significativamente os horizontes da Sociolinguística brasileira. Por fim, este artigo defende que o maior desafio da Linguística contemporânea talvez não seja desenvolver novas metodologias ou novos modelos analíticos, mas reaprender a escutar. Durante muito tempo, determinados repertórios permaneceram invisíveis não porque lhes faltasse estrutura, mas porque lhes faltava legitimidade perante as instituições que produzem conhecimento. O Pajubá demonstra que as margens nunca estiveram em silêncio; elas apenas foram pouco ouvidas. Reconhecê-lo como laboratório sociolinguístico significa admitir que algumas das questões mais relevantes para compreender a linguagem contemporânea talvez não estejam nos centros de prestígio linguístico, mas justamente nos territórios onde sujeitos historicamente subalternizados reinventaram, pela linguagem, novas formas de existir, resistir e produzir conhecimento. Nesse sentido, estudar o Pajubá não representa apenas ampliar os limites da Sociolinguística; representa ampliar, sobretudo, os limites daquilo que estamos dispostos a reconhecer como ciência da linguagem.

Referências

BORBA, Rodrigo (org.). *Discursos transviados:* por uma linguística queer. 1. ed. São Paulo: Cortez, 2021.

BOURDIEU, Pierre. *Language and symbolic power.* Cambridge: Polity Press, 1991.

BUTLER, Judith. *Gender trouble:* feminism and the subversion of identity. New York: Routledge, 1990.

DENZIN, Norman K.; LINCOLN, Yvonna S. (org.). *The SAGE handbook of qualitative research.* 5. ed. Thousand Oaks: SAGE Publications, 2018.

ECKERT, Penelope. *Linguistic variation as social practice:* the linguistic construction of identity in Belten High. Oxford: Blackwell, 2000.

FAIRCLOUGH, Norman. *Discurso e mudança social.* Brasília: Editora Universidade de Brasília, 2001.

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

MELO, Iran Ferreira de. *Linguística queer*. Campinas, SP: Pontes Editores, 2024.

NARO, Anthony Julius. Idade. *In*: MOLLICA, Maria Cecília (org.). *Introdução à sociolinguística variacionista*. Rio de Janeiro: UFRJ, 1992.

RÉGIS, Renato. *Gramática descritiva do pajubá:* em diálogo com as linguísticas queer, africana e antropológica. São Paulo: Parábola Editorial, 2025.

SCHERRE, Maria Marta Pereira. *Doa-se lindos filhotes de poodle*: variação linguística, mídia e preconceito. São Paulo: Parábola Editorial, 2005.

SILVERSTEIN, Michael. Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, [*s. l.*], v. 23, n. 3-4, pp. 193-229, 2003.

VAN DIJK, Teun A. *Ideology:* a multidisciplinary approach. London: SAGE, 1998.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72240

Entre ø e forma explícita como marcador relacional em orações relativas: reorganização do sistema no PB[1]

Between ø and the overt form as a relational marker in relative clauses: system reorganization in BP

Cecilia Mollica[2]

RESUMO

Este artigo investiga a alternância entre o apagamento (Ø) e a retenção de preposições como marcadores relacionais em orações relativas do português brasileiro (PB). A partir de dados de fala espontânea da amostra Recontato/2000 (Projeto PEUL), a pesquisa adota uma abordagem descritivo-qualitativa, revelando que fatores extralinguísticos, como gênero, idade e escolaridade, não apresentam efeito determinante sobre o fenômeno. Os resultados demonstram que a variação é governada por restrições de natureza interna e organiza-se ao longo de um *continuum* semântico-funcional. A realização nula atua como valor estrutural *default* em contextos de alta previsibilidade e esvaziamento nocional, ocorrendo de forma categórica com as preposições *a* e *de*, enquanto a forma explícita resiste como mecanismo de resgate argumental em nexos de maior densidade, como *com* e *em*. Sob uma perspectiva funcional e com base em evidências diacrônicas, os dados indicam que o padrão atestado não configura déficit gramatical, mas reflete a reorganização do sistema relacional do PB, impulsionada pela economia estrutural, pela reanálise verbal e pela rotinização do *que* como operador universal de subordinação.
PALAVRAS-CHAVE: Orações relativas. Marcador relacional. Marcação nula (Ø). Contínuo semântico. Português brasileiro.

ABSTRACT

This article investigates the alternation between the deletion (Ø) and the retention of prepositions as relational markers in Brazilian Portuguese (BP) relative clauses. Based on spontaneous speech data from the Recontato/2000 sample (PEUL Project), the research adopts a descriptive-qualitative approach, revealing that extralinguistic factors—such as gender, age, and education level—do not have a determining effect on the phenomenon. The results demonstrate that the variation is governed by internal constraints and is organized along a semantic-functional continuum. Null realization acts as a default structural value in contexts of high predictability and notional bleaching, with the categorical deletion of *a* and *de*, while the explicit form persists as an argument rescue mechanism in more dense syntactic links, such as *com* and *em*. From a functional perspective and based on diachronic evidence from the literature, the data indicate that the attested pattern does not constitute a grammatical deficit, but rather reflects a reorganization of the BP relational system driven by structural economy, verbal reanalysis, and the routinization of *que* as a universal subordination operator.

[1] Dedico este artigo à memória do Professor Anthony Julius Naro, figura central na consolidação da Linguística no Brasil e pilar do PEUL (UFRJ). Naro legou-nos uma ética de pesquisa pautada pelo rigor teórico e metodológico e pelo compromisso com a inserção da nossa produção acadêmica no cenário internacional. Entre os seus muitos ensinamentos, salientava que o linguista deve render-se aos dados e nunca o contrário. Defendia Naro que o avanço do conhecimento pressupõe que os fenômenos linguísticos sejam revisitados múltiplas vezes sob novos olhares. Este texto busca honrar seu legado, pautando-se pela honestidade intelectual e pela primazia do fato linguístico, marcas indissociáveis de sua prática acadêmica.

[2] Universidade Federal do Rio de Janeiro (UFRJ), ceciliamollica@letras.ufrj.br, https://orcid.org/0000-0001-6261-4663.

Recebido em 20/01/2026 - Aceito em 10/06/2026

KEYWORDS: Relative clauses. Relational marker. Zero marking (Ø). Semantic *continuum*. Brazilian Portuguese.

1. Introdução

As orações relativas no português brasileiro (doravante PB) constituem domínio linguístico privilegiado para a observação de processos de reorganização gramatical, particularmente no que diz respeito à codificação das relações entre predicadores e termos relativizados. Defendo neste artigo que a alternância entre a realização explícita da preposição e a marcação nula (Ø) nessas construções reflete um processo de reanálise estrutural em curso, associado à redefinição do papel das marcas relacionais no sistema do PB.

Os usos da língua falada mostram que a exigência da tradição gramatical (Bechara, 2009) com relação à presença de preposição como elemento obrigatório da regência verbal não se sustenta na gramática natural do falante. Os estudos voltados para o português brasileiro, como o de Bagno (2000, p. 185-198), demonstram que tanto os empregos das variedades [+urbanas] [+orais] [+/-monitoradas], quanto os das variedades [+urbanas] [+escritas] [+/-monitoradas], em falantes cultos, exibem a variação ora analisada sem prejuízo interpretativo. Nesta pesquisa, parto da hipótese de que a reorganização de estratégias de relativização no PB voltada para a oposição marcação explícita/ variante *default* (Ø) é principalmente regulada pelo grau de inferibilidade nos níveis sintático, semântico e pragmático na relação sentencial. Em linha com Camacho (2012), assumo que a persistência de nexos na fala espontânea ocorre em situação de menor transparência relacional, enquanto a opção não marcada situa-se em contextos de alta previsibilidade discursiva.

De acordo com Labov (1972), padrões linguísticos variáveis que se distribuem de forma sistemática e condicionada não devem ser tratados como desvios ou irregularidades, mas como parte integrante da competência gramatical dos falantes, de modo que a alternância entre marcação nula *versus* realização preposicionada nas orações relativas do PB não se distribui aleatoriamente, mas confirma a concepção de heterogeneidade estruturada. Gomes (1996) propõe que a marcação nula, licenciada pelo sistema do PB, configura o valor estrutural *default*, sensível às propriedades semânticas do nexo relacional. Defendo que, além dos princípios de economia e inferibilidade, a marcação nula opera em virtude da ressemantização do verbo da cláusula relativa e pode sofrer o efeito de traços do SN antecedente, assim como da distância entre o antecedente e a relativa, em nome da economia sintática.

Notem-se os exemplos de fala espontânea. Em (1), o verbo *assistir*, tradicionalmente associado à preposição *a*, ocorre em oração relativa sem marcação preposicional explícita; em (2), o verbo *gostar* é empregado sem o *de* como seria a exigência da regência normativa:

(1) Nem sei nem qual filme **que** ela assistiu. (Amostra Recontato, falante 2)

(2) Se fosse um lugar **que** eu não gostasse, eu não ia fazê. (Amostra Recontato, falante 4)

Em ambos os casos, a interpretação da relação entre o predicador e o termo relativizado é plenamente recuperável. Na oralidade, a recorrência desses padrões aponta para uma redistribuição funcional da codificação relacional, na qual a dependência argumental deixa de ser obrigatoriamente expressa por meio da preposição.

Araújo (2025) demonstra que a variação nas estratégias de relativização acompanha toda a história do português, com ocorrência persistente de construções com e sem preposição, inclusive com padrões equivalentes às relativas cortadoras. Essa continuidade diacrônica sustenta a interpretação da marcação *default* Ø no PB como resultado de rearranjo interno em função da natureza relacional dedutível do ponto de vista semântico e não como inovação recente. De acordo com Santos (2025), Bagno (2000), dentre outros, a análise das orações relativas no português revela um sistema em constante tensão entre norma padrão e uso efetivo. Os estudos demonstram que, enquanto a gramática tradicional prescreve o uso de pronomes relativos específicos, como *cujo* ou *o qual*, a realidade descritiva dos vários estados do português aponta para a simplificação das estruturas, com predominância do *que* e o surgimento de estratégias como as relativas cortadoras e copiadoras.

A recuperabilidade semântica da relação entre cláusulas é favorecida por processo de ressemantização, no qual também o verbo, ao abdicar de sua regência estrita, passa a funcionar como moldura de cenário ou ancoragem discursiva, permitindo que a conexão entre as orações seja estabelecida por inferência direta.

(3) Tem que ver o pessoal, pra dá u:m pouco de dinheiro pra fazer filme e chega umas merdas dessa aí: faz qualquer coisa assi:m, e tira a credibilidade dos investidores, **que** já é muito difícil de captar dinheiro. (Amostra Recontato, falante 1)

Em *tira a credibilidade dos investidores, que já é muito difícil de captar dinheiro*, no trecho em (3), a relativa não restringe o referente, mas sustenta a progressão argumentativa, o que explica a preferência pelo *que* como marcador relacional *default* 0. No padrão culto, teríamos *investidores, de quem já é muito difícil captar dinheiro,* tratando-se de uma relativa explicativa. O padrão recorrente com *que* decorre de estratégia de baixo custo formal, favorecida pela depreensão do significado e compatível com a reorganização em curso do sistema de relativização no PB defendida nesta pesquisa.

No processo de enfraquecimento da marcação preposicional, observa-se a expansão funcional do *que*, conforme atestado em Lopes e Silva (2007), que indicam o afastamento progressivo do *que* de estatuto estritamente pronominal em direção ao de operador relacional pouco especificado. A alta frequência do que é o motor da gramaticalização em direção ao marcador *default*, pois a repetição constante promove a rotinização de estruturas que conduz ao esvaziamento do valor pronominal do item e sua transformação em complementizador universal. Tal constatação corrobora a hipótese de Tarallo (1993), que defende a recategorização do *que* como o complementizador nas orações relativas não padrão do português brasileiro, atuando como operador de subordinação desprovido de traços-phi, sem marcas morfológicas de concordância de gênero, número e pessoa. Esse processo justifica a predominância das relativas cortadoras na fala, em que a busca por eficiência e fluidez

discursiva resulta na marcação *default* e na consequente simplificação do sistema de relativização. Essa trajetória gera esvaziamento da carga semântica da preposição *em*, situação que se torna evidente quando analisamos contextos de fala espontânea. Nos exemplos (4) e (5), observa-se como núcleos nominais de tempo pontual favorecem ainda mais a omissão preposicional:

(4) Tem dias **que** eu chego, isso aqui tá um auê, eu não tô a fim de fazê nada, que eu tô morrendo de cansada, dor de cabeça, eu deito e quero dormí, quero descansá, não quero sabê de casa, não quero sabê de comida, não quero sabê de nada. (Amostra Recontato, falante 1)

(5) Porque a Luana é uma dificuldade pra comê assim uma ("folha") aí eu pego bota couve, aí ela não gosta de abóbora, bota abóbora dentro do feijão, na hora **que** vai pro prato, eu já amasso, então ela come que nem sente (Amostra Recontato, falante 9)

Em (4) e (5), não se trata de omissão da preposição subjacente recuperável *em que*, mas de construções já gramaticalizadas, nas quais os nomes núcleo *dias* e *hora* atuam como pivô relacional, e a relação semântica temporal é integralmente inferida. Estudos sobre a reanálise de sequências do tipo *(prep)+(det)+N + que*, como os de Paiva e Pereira (2004) e de Pereira, Braga e Paiva (2010) mostram que tais estruturas passam a funcionar como unidades relacionais abstratas, dispensando a codificação preposicional explícita. No exemplo (4), o núcleo *dias* atua em uma estrutura existencial onde a relação temporal é tão intrínseca, que a presença da preposição *nos quais* ou *em que* é sentida como excesso formal desnecessário ao processamento imediato. Já no exemplo (5), a construção *na hora que* ilustra perfeitamente o estatuto híbrido discutido por Paiva e Pereira (2004), segundo o qual, embora *na* esteja preservado, a ausência da preposição interna *em* sinaliza que o sintagma está em vias de se converter em conectivo temporal equivalendo a *quando*. Nesses casos, a análise afasta-se da simples ideia de omissão da preposição subjacente recuperável e passa a situar-se no campo das construções já gramaticalizadas.

O mesmo item formal pode ainda assumir valores discursivos mais amplos, aproximando-se de conectores causais, como em:

(6) Eu não tô a fim de fazê nada, **que** eu tô morrendo de cansada. (Amostra Recontato, falante 1)

A coexistência desses usos de causalidade evidencia princípios de persistência e estratificação propostos por Hopper (1991), já que formas inovadoras se expandem sem eliminar imediatamente camadas mais antigas do sistema. No PB falado, a realização Ø estende-se a amplo conjunto de relações semânticas. Em contextos temporais, por exemplo, observa-se a neutralização sistemática da preposição:

(7) Quando nós compramos o microondas, na época, gente o Wanderley ficava fascinado, cada dia era uma coisa diferente, enquanto isso a casa ia ficando o caos, mas aquele momento **que** eu chegava, assim, tipo oito horas da noite, microondas é uma coisa, assim, muito rápida, então você oito horas tinha alguma coisa pronta pro jantar, uma coisa diferente (inint), agora não que num vira mais novidade, né? (Amostra Recontato, falante 4)

O processo de reanálise também se estende a núcleos que denotam duração ou períodos cronológicos como em (7). Nesse caso, a carga semântica de *momento* é suficientemente densa para ancorar a relação de subordinação, tornando a preposição *em* semanticamente redundante. Examine-se o exemplo (8):

(8) Nessa oficina **que** eu tô hoje que é a de trem de pouso, eu tô colocando pra eles porque a Varig começou. (Amostra Recontato, falante 9)

No exemplo (8), a função locativa é recuperada não pela sintaxe da oração relativa, mas pela semântica do núcleo *oficina* aliada à regência do verbo. Conforme postula Mollica (2003), o apagamento da preposição nesse contexto locativo demonstra que, no contínuo de carga semântica, *em* ocupa posição de vulnerabilidade. De acordo com Santos (2020) e Araújo (2025), a ausência do nexo conta com a capacidade do interlocutor de inferir a relação espacial a partir do núcleo nominal. Para Camacho (2012), a transparência funcional permite que o falante economize o esforço de alçamento (*pied-piping*),

A transição do domínio temporal para o locativo (*cf.* exemplos 7 e 8) ratifica a tese de que o PB caminha para uma gramática de marcação mínima, priorizando núcleos que atuam como orientadores semânticos. Além de tempo e lugar, nomes relacionais de modo e predicadores que selecionam complementos preposicionados também favorecem a marcação zero:

(9) A maneira **que** a mulher fala é mais delicada. (Amostra Recontato, falante 9)

(10) Eu com 15 anos falei: “mãe, tu tem que separá do meu pai”, meu pai num é (inint) a coisa é muito chata, sabe? Mas isso é uma entrevista, acredito que ela tenha o seu fim, né? Não tenha abertura assim geral, [né? Totalmente (inint)] então (inint) e tal e ela muito não, meio que entrou e tal, muito pela imagem da forma **que** ela tinha sido educada, mais pelos valores [que era dos pais dela do que realmente da prática]. (Amostra Recontato, falante 9)

(11) Elas ficam amigas, porque ela descobre que a mulher do, a ex-mulher do homem **que** ela está vivendo estava com câncer. Aí, ela fica amiga pra dar força a ela e tudo. Mas eu não me lembro desse filme. (Amostra Recontato, falante 14)

Nos casos (9) e (10), os núcleos *maneira* e *forma* operam de modo análogo aos nomes de tempo, porque a semântica do substantivo anuncia a relação de modo, facilitando a marcação *default*. Já o exemplo (11) ilustra a relativa cortadora em sua essência sintática. Nele, o corte do nexo evidencia a produtividade da estratégia cortadora no PB, ocorrendo mesmo quando o antecedente *homem*, de alta carga lexical, não atue como pivô relacional óbvio de tempo ou modo.

A hipótese central que norteia esta investigação é a de que a alternância entre retenção e apagamento preposicional é governada por um *continuum* semântico. Proponho que os nexos preposicionais se distribuem ao longo de uma escala de densidade lexical e previsibilidade argumental.

Em uma extremidade, situam-se as preposições puramente gramaticais e altamente inferíveis, como *a* e *de*, enquanto na outra extremidade localizam-se nexos de maior carga semântica como *em* e *com*, que podem resistir ao apagamento, em alguns contextos, por serem relevantes à interpretação do enunciado. Nessa perspectiva, assumo que o PB opera em um gradiente de marcação que vai da retenção obrigatória a total elisão da preposição. Assim, mesmo em ambientes argumentais mais estritos, a fala espontânea admite o zero, como se observa em (12):

(12) Então, hoje em dia, a televisão, ela vai em busca de programas **que** a população fique entorpecida. (Amostra Recontato, falante 11)

Em (12), diferente das relativas temporais, a elisão da preposição em contextos de regência oblíqua, como *ficar entorpecida com*, exige maior esforço inferencial. O apagamento do nexo prepositivo em contextos de antecedentes complexos corrobora a tendência do PB em tratar o *que* como conectivo universal, desvinculado da marcação de caso.

Os exemplos apresentados indicam que a variante não marcada depende tanto da natureza do antecedente quanto da grade argumental do predicador. Como se verá adiante, o sistema caminha para a neutralização da regência, processo em que o verbo é ressemantizado para funções de ancoragem discursiva. Nesse quadro, analiso a distribuição de *a*, *de*, *em* e *com*, interpretando a ausência desses nexos não como omissões aleatórias, mas como parte de uma reorganização funcional célere e sistemática do sistema de relativização no PB. Para testar essa hipótese e observar o comportamento dessa reorganização na fala espontânea, a seção a seguir detalha o percurso metodológico da presente pesquisa e a análise empírica dos dados extraídos da Amostra Recontato.

2. Análise empírica

O exame dos empregos linguísticos reais partiu do pressuposto de que a variante Ø e a variante com marcador preposicional não constituem fenômeno categórico, mas se distribuem de forma gradual, consoante variáveis parametrizadoras dos empregos de preposições nas relativas cortadoras. Os dados empíricos permitem observar em que medida a especialização funcional do operador *que* se articula a diferentes graus de transparência relacional, refletidos nos padrões de uso das preposições *a*, *de*, *em* e *com* ilustrados anteriormente.

Assim, a presente análise examina a realização das preposições *a*, *de*, *em* e *com* em orações relativas em alternância à marcação *default* Ø, com base em dados da Amostra Recontato do PEUL, constituída de indivíduos regravados no ano 2000, que constam originalmente na Amostra Censo /80. A grade inicial da pesquisa, conforme detalhado na tabela 1, rastreou dados de 16 indivíduos. O grupo abrange uma ampla faixa etária, com falantes entre 25 e 77 anos, e contempla uma diversidade de níveis de escolaridade, que se estende desde o Ensino Fundamental 1 até o Ensino Superior.

Tabela 1: Perfil social da grade inicial de falantes da amostra Censo 2000/Recontato

Falante	Sexo	Idade	Escolaridade
R01 (Eri)	Feminino	25	Ensino Médio
R02 (AdrL)	Feminino	26	Fundamental 2 (inc.)
R03 (AdrR)	Feminino	28	Superior
R04 (Fat)	Feminino	33	Superior (inc.)
R05 (SanR)	Feminino	33	Superior
R06 (Jup)	Feminino	36	Fundamental 2 (inc.)
R07 (Leo)	Masculino	35	Superior (inc.)
R08 (Lei)	Feminino	43	Fundamental 2 (inc.)
R09 (Dav)	Masculino	47	Ensino Médio
R10 (Vas)	Masculino	48	Fundamental 2
R11 (Eve)	Feminino	58	Ensino Médio (inc.)
R12 (Mgl)	Feminino	70	Ensino Médio
R13 (Jan)	Masculino	74	Fundamental 1 (inc.)
R14 (Nad)	Feminino	74	Fundamental 1
R15 (Ago)	Masculino	77	Fundamental 1
R16 (Jos)	Feminino	77	Fundamental 1

Fonte: Elaboração da autora

A escolha da *data base* se justifica em função de alguns falantes terem galgado níveis mais altos de escolaridade em relação à amostra Censo/80, de modo que suspeitei, inicialmente, ser o letramento formal relevante para a emergência da marcação com nexos prepositivos, conforme demonstrado por Melo (2016). Na realidade, na amostra aqui trabalhada, o fenômeno mostrou-se governado primordialmente por restrições de natureza interna, senão vejamos.

A análise partiu do levantamento de todas as orações relativas na fala dos 16 informantes rastreados, contudo, após a filtragem metodológica, apenas 10 deles apresentaram dados válidos do fenômeno para a pesquisa (7 mulheres e 3 homens). Embora o *corpus* contemple estratificação social diversificada, a observação qualitativa revelou que as variáveis extralinguísticas gênero/sexo, idade e escolaridade não exibem efeito determinante em relação à marcação nula (Ø) ou à presença do nexo prepositivo, por algumas razões: (1) número insuficiente de sujeitos homens em relação ao de mulheres; (2) faixa etária extremamente alargada; (3) número insuficiente de evidências quanto à variável escolarização.

O estudo de Rocha (2011, p. 97-121) compara as relativas na fala popular e culta de Salvador com uma massa robusta de dados, dos quais foram consideradas também as preposições *a*, *de*, *em* e *com* para a análise. O *input* da regra de apagamento dos nexos se mostrou muito alto, em torno de 94,9% a 96,6%, em relação à escolaridade nos *corpora* culto e popular considerados. Nos dados, apenas o nexo *em* remetendo a espaço abstrato foi utilizado, no sentido de *meio em que uma pessoa vive*. Rocha (2011) entende, assim, que se trata de emprego residual de uma mudança praticamente operada no sistema do PB, ponto a partir do qual a variação já não é mais sensível ao efeito da escolarização, especificamente, sofrendo igualmente pouca influência de fatores extralinguísticos em

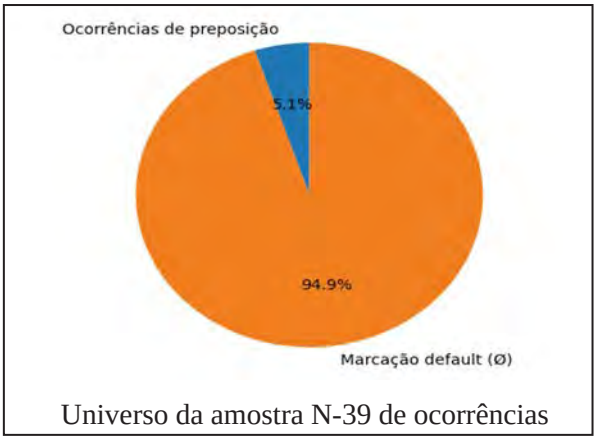
geral. Trata-se de casos em que a variante padrão vai emergir sob pressão escolar somente, mesmo assim, sem muito êxito das estratégias de ensino.

Para responder à sistematicidade do fenômeno, a análise aqui empreendida baseou-se em parâmetros predefinidos, estabelecendo os contextos favorecedores para a retenção ou apagamento da preposição. Em diálogo com as restrições apontadas por Gomes (1996) para a perda e a aquisição das preposições no PB, a avaliação dos dados operou sobre três eixos centrais: (i) a natureza semântica da preposição, opondo nexos funcionais, como *a/de*, a lexicais, como *com*; (ii) o traço semântico do antecedente/núcleo, que corresponde aos contextos locativos, temporais e de modo; (iii) o grau de inferibilidade garantido pela matriz verbal, vale dizer, a capacidade de o verbo atuar e outros índices contextuais operarem como ancoragem discursiva, dispensando a preposição.

Cabe ressaltar que, devido à filtragem que resultou no N=39, os resultados configuram-se como tendências descritivas qualitativas. Certamente, nova verificação deve ser realizada em novas amostras ora constituídas no âmbito do PEUL, que fornecerão escopo adequado e necessário para atestar a operação da mudança em curso no PB.

Os resultados detalham a distribuição de ocorrências relativa a fatores estruturais e semânticos e revelam, de um lado, como o apagamento é alavancado em contextos de alta previsibilidade, e, de outro, como a retenção concentra-se em nexos de maior densidade. Os percentuais apresentados revelam padrão altamente assimétrico. O gráfico 1 mostra que somente 5,1% das ocorrências exibem preposição realizada, enquanto 94,9% correspondem a construções com marcação Ø.

Gráfico 1: Índice de ocorrência de preposição em orações relativa



Universo da amostra N-39 de ocorrências

Fonte: Elaboração da autora

Apenas 2 ocorrências preposicionadas em 39 relativas reforçam a leitura de Ø como valor estrutural *default*, padrão que vem sendo interpretado à luz da gramaticalização do *que*, como proposto por Lopes e Silva (2007) e explicado em Martelotta *et al*. (1996). À medida que o *que* se consolida como operador relativo generalizado, torna-se reduzida a necessidade de marcação explícita das relações sintáticas por meio de preposições, sobretudo em contextos em que o vínculo argumental é facilmente recuperável por inferência discursiva.

Sob perspectiva formal, os resultados dialogam também com a proposta de Kato e Nunes (2014), segundo a qual as diferentes estratégias de relativização no PB convergem para estruturas menos complexas do ponto de vista sintático. De acordo com os autores, a predominância de relativas sem preposição pode ser interpretada como reflexo da preferência por derivações mais econômicas.

O padrão da cortadora caracteriza-se pela omissão da preposição regida pelo verbo da oração relativa, resultando em estrutura na qual apenas o pronome relativo *que* atua como conectivo. Mollica (1977, 1997) demonstra que não se trata de erro aleatório, mas de regra variável produtiva no PB, em que a cortadora alterna com as estratégias copiadora e padrão, sendo a cópia incidindo predominantemente como sujeito no papel de realce, de topicalização. À época, eu já observei que a preposição é apagada com mais frequência quando a função sintática do relativo é de um complemento menos central, como adjuntos adverbiais de lugar ou tempo.

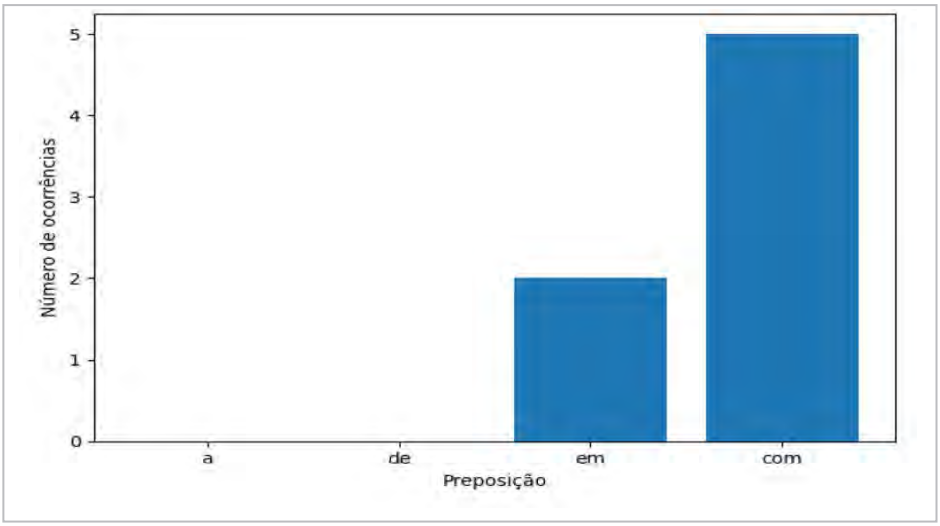
Mollica *et al*. (2018) se voltam para produções em mídias sociais e textos escolares. Os autores afirmam que a variante *default Ø* se estabilizou no PB e continua sendo a deriva natural da língua também em ferramentas tecnológicas e na escrita digital, por serem estruturas mais ágeis. O PB deixa a forma padrão restrita a contextos de monitoramento ou a textos literários clássicos.

Melo (2016) e Mollica (2003) se referem à variação preposicional como sensível a fatores extralinguísticos, o que não se confirmou no *corpus* desta pesquisa em virtude de serem tão poucos os dados examinados. Os estudos com amostras maiores, como o de Bagno (2000), apontam também para baixa incidência de marcação, o que sugere a prevalência da marcação *deflault* como padrão típico de contextos menos monitorados.

Assim, a omissão do termo regido se impõe como opção funcional, fenômeno que sintetiza convergência entre mudança gramatical e economia estrutural. Como medida no âmbito da pedagogia da variação linguística, Castilho (2010) e Lima e Mollica (2026) discutem a incorporação da oralidade no ensino do português mediante a constatação de que a escola tenta impor a relativa padrão, no entanto o sistema mental do falante brasileiro já assimilou a relativa cortadora como forma natural de processamento sintático. Na mesma direção, mas voltando-se para o sistema do português europeu (PE), Duarte (2011, p. 15-30) reflete que, ao longo da escolaridade, os adolescentes vão convergindo para um sistema com sentenças relativas preposicionadas, embora mantenham níveis elevados de aceitação de relativas cortadoras.

Os dados aqui considerados atestam a relação *retenção/apagamento* com base na densidade do nexo. No gráfico 2, observa-se a retenção pontual das preposições *com* e *em*, contrastando com o apagamento categórico de *a* e *de*.

Gráfico 2: Distribuição das preposições em relativas (N = 39)



Fonte: Elaboração da autora

Pela configuração disposta no gráfico 2, a relação interclausal é sensível à carga nocional do nexo, o que permite dialogar com os estudos funcionalistas diacrônicos que associam a variante *default* Ø à gramaticalização do *que* e à reconfiguração da marcação relacional no PB. Em linha com Tarallo (1983), confirma-se que o apagamento do nexo nas relativas cortadoras é governado pela natureza da preposição. Nessa mesma perspectiva, Tarallo (1993) lança a hipótese do estatuto do *que* como um complementizador nas relativas não padrão.

Fica, então, estabelecido o gradiente que vai das preposições puramente funcionais *a* e *de*, como atestado nos exemplos (1) e (2), *filme Ø que ela assistiu* e *lugar Ø que eu não gostasse*, até *em* e *com*, que resistem ao corte em contextos de maior carga nocional nas relações de meio/ferramenta e de companhia/adjacência.

A materialidade do fenômeno é atestada pelo exemplo (13). Nele, observa-se a variação intrapessoal em tempo real, no mesmo turno em que o falante alterna a retenção e o apagamento do nexo *com* ao utilizar o predicador *casar*:

(13) [...] ela conheceu essa pessoa **com quem** ela se casou e regularizou a situação dele [no]... no país (Estados Unidos). E essa minha segunda filha ela conheceu (hes) esse marido agora **que [Ø]** ela casou aqui no Brasil, ele é americano. (Amostra Recontato, falante 11)

O nexo *em*, por sua vez, pode resistir como mecanismo de segurança, de resgate argumental em contextos de maior esforço inferencial. Tal fato ilustra-se no processamento estrutural em transição, em ocorrências de refacção, como em (14):

(14) Uma coisa **em que** <a gen...> nós não concordamos, ela gosta, ela é (hes), vamos dizê, uma novela ou outra ela vê. (Amostra Recontato, falante 7)

Em (14), o falante aciona primeiramente a estrutura padronizada, hesita e recalcula a rota sintática. A construção com refacção *uma coisa em que nós não concordamos, ela gosta, ela é (hes), vamos dizê, uma novela* joga luz ao entendimento de que o apagamento constitui estratégia natural de coesão com o mínimo esforço e não consiste em erro.

Quadro similar se estende a contextos de adjuntos adverbiais de tempo e modo, cujos antecedentes possuem os traços [-humano] e [-animado]. O exemplo (15) demonstra que a exigência semântica de retenção é quase inexistente e o corte da preposição é favorecido pela rotinização:

(15) A maneira **que [Ø]** a mulher fala é mais delicada.

Por outro lado, o polo de maior retenção no *continuum* obedece a restrições semânticas com base na hierarquia de animacidade. Quando o antecedente é [+humano], o uso da variante *default que* compromete a clareza e gera ambiguidade, como ilustrado em (16):

(16) Tem um gerente de banco **que [Ø]** eu falo: "Ah, seu Jorge".

Em (16), a ausência da preposição obscurece a regência exata, *falar com* o gerente ou *falar do* gerente, e transfere a operação de desambiguação para o nível pragmático do discurso. A força do marcador universal Ø no PB é tão grande, que se manifesta em quadros densos, ancorando-se no contexto situacional, como exemplificado em (17):

(17) E: Você nunca teve problema com patrão?
F: Não. Uma **que [Ø]** eu trabalhei foi para os Estados Unido...

Aqui, a variante *default* sobrevive no lugar de *com quem* graças ao cenário compartilhado na interação entre entrevistador (E) e falante (F), quando fica claro que o sentido se sobrepõe à configuração estrutural por meio da devida carga nocional e discursiva de cada contexto.

O cenário desenhado pelos dados analisados ratifica o fato de que o apagamento preposicional consiste em estratégia preferencial do sistema do PB, entendendo-se a marcação *default* não como déficit ou empobrecimento gramatical, mas adaptação pragmática. Observe-se ainda o exemplo (18).

(18) Aí eles ficam lá cima, né? E eu fico em baixo na minha pedrinha, escuto, mas é uma brigalhada. Nesse final de semana, teve uma briga **que [Ø]** saiu até facada dos maridos lá, quase que um mata o outro lá, por causa de negócio de fofoca, negócio de água, brigalhada. (Amostra Recontato, falante 6)

Em (18), o relativizador *que* retoma *briga* como um cenário, cujo encadeamento narrativo *teve uma briga e saiu facada*, é autossuficiente. Na variedade padrão, a gramática exigiria o suporte de uma estrutura preposicionada *na qual*, mas a força lógica da relação locativa dispensa o andaime. A gramática permite que o falante economize no aparato formal sem prejuízo para a articulação do sentido. À luz de Martelotta (2011), essa escolha reflete a tendência para preferência de estratégias

de menor custo cognitivo, mecanismos que permitem forte conexão entre os eventos e dispensam marcação sintática plena ao deslocar a força interpretativa para a dimensão semântico-discursiva.

Assim, a distribuição das ocorrências revela a ecologia de sobrevivência das formas, pois nem todo elo gramatical resiste à erosão do tempo. O protagonismo de *com* nas retenções denuncia que a gramática não abre mão do que pesa, para evitar problemas de informação. Enquanto a volatilidade de *a* e *de* confirma que partículas sem *cor* lexical são as primeiras a serem dispensadas pela lei do menor esforço, a preposição *em* surge como fiel da balança, um elo em trânsito que ainda oscila para o desaparecimento. O exemplo (19) captura esse estado de suspensão, em que o marco temporal dispensa a marca explícita sem prejudicar a mensagem:

(19) A última vez **que [Ø]** eu trabalhei foi em outubro. (Amostra Recontato, falante 1)

Em (19), o *que* opera como vetor de ancoragem. O falante utiliza o marcador *default* para projetar as coordenadas de tempo e espaço, *última vez* e *outubro*, de modo a respeitar a primazia do sentido, porque a lógica do relato é autossuficiente. A rotinização de estruturas cria fórmulas prontas que o cérebro processa como pacotes únicos e não como peças isoladas. Nessas sequências, o *que* assume o papel de *cola universal*, um marcador automático que dispensa regências complexas em favor da velocidade de processamento. O exemplo (20) ilustra esse automatismo em que a estrutura N + *que* opera no piloto automático:

(20) Eu com 15 anos falei: "mãe, tu tem que separá do meu pai", meu pai num é (inint) a coisa é muito chata, sabe? Mas isso é uma entrevista, acredito que ela tenha o seu, né? Não tenha abertura assim geral, [né? Totalmente (inint)] então (inint) e tal e ela muito não, meio que entrou e tal, muito pela imagem da forma **que [Ø]** ela tinha sido educada, mais pelos valores que era dos pais dela do que realmente da prática. (Amostra Recontato, falante 7)

No exemplo (20), o falante opera com blocos de memória, e a preposição é descartada por constituir informação redundante diante do contexto. O relativizador atua, então, como nexo de mínima resistência e sinaliza apenas a conexão necessária para que a informação alcance o interlocutor com o menor custo de processamento possível. De acordo com Bybee (2003), a rotinização ocorre quando sequências de alta frequência perdem sua estrutura interna analítica. Assim, a estrutura *da forma que* em vez de *da forma com que* não constitui ausência de regra, mas a aplicação de uma regra rotinizada, por meio da qual o *que* atua como marcador de subordinação geral que otimiza a fluidez da fala espontânea.

Os dados atestam que a alternância entre preposição explícita e Ø se organiza ao longo de um contínuo, no qual a realização nula ocupa a posição menos marcada, enquanto a preservação da preposição emerge em contextos semanticamente mais exigentes. O padrão de apagamento predominante na amostra não aponta para perda ou empobrecimento estrutural, mas para a reorganização funcional da marcação relacional nas orações relativas, compatível com a tendência geral à economia estrutural e à transitividade direta do PB.

O gráfico 3 ilustra visualmente a materialização do *continuum* semântico. A disposição das preposições ao longo de uma escala parte do polo de maior retenção [+significado] em direção ao polo de esvaziamento dos nexos de alta previsibilidade estrutural [-significado]. Esse ordenamento espacial corrobora a tese de que a aplicação da regra de apagamento não opera como oposição binária, mas reflete o grau de transparência e recuperabilidade semântica demandado no contexto de cada verbo regente.

Gráfico 3: Distribuição da retenção preposicionada e do apagamento (Ø) conforme a densidade semântica da preposição.

[+ significado] <--- com --- em --- de --- a ---> [- significado]

Fonte: Elaboração da autora

Legenda: **vermelho** indica a preposição ***com*** (maior densidade lexical/retenção favorecida); **laranja** representa ***em*** (maior previsibilidade/alto índice de apagamento); **verde** refere-se à preposição ***de*** (previsibilidade de significado/maior apagamento); **azul** indica o nexo prepositivo ***a*** (maior previsibilidade semântica/alto índice de apagamento).

A descrição qualitativa dos dados revela o comportamento da gramática falada operando em tempo real. Com efeito, a alternância entre a preposição explícita e a realização nula Ø obedece a uma hierarquia sensível à densidade semântica. De acordo com Camacho (2013), a estabilização da omissão como opção estruturalmente não marcada no PB decorre do princípio de economia linguística vinculado à recuperabilidade, à transparência semântica. Em contextos de alta previsibilidade, o nexo preposicional atua apenas como marcador de junção, cuja função é dedutível a partir do significado do verbo regente e/ou de índices contextuais. Assim, o esvaziamento de preposições como *a* e *de* desonera o falante de sua realização explícita, uma vez que o valor relacional é processado mentalmente sem a necessidade de sinalização fonética.

Segundo Bybee (2010), as formas linguísticas associadas a valores semânticos mais específicos e menos frequentes tendem a resistir a processos de redução formal, justamente porque a interpretação depende menos de inferência contextual automática. Isso explica, de acordo com os resultados apresentados no gráfico 2, a menor incidência de Ø com *em* e, sobretudo, com *com*, cuja omissão compromete, de modo significativo em alguns contextos, a recuperação da relação argumental. Nesses casos, a preposição explícita passa a funcionar como recurso marcado, acionado quando a inferência discursiva não é suficiente para garantir a interpretação informacional.

Essa dinâmica pôde ser observada anteriormente na dualidade do verbo *casar* no exemplo (13), em que a falante alternou entre *com quem* e *que [Ø]* no mesmo turno. Tal comportamento ilustra perfeitamente o *continuum* semântico visualizado no gráfico 3. Quando o verbo assume a função de ancoragem do sintagma nominal como tópico atual, *o marido agora que ela casou*, o marcador zero emerge, priorizando a indexação do referente no cenário discursivo em detrimento da precisão de regência.

O gráfico 3 condensa, portanto, a geografia das preposições *a, de, em* e *com* nas relativas do PB, projetando-as em um gradiente de substância. Ao articular a força dos dados à tradição funcionalista, a visualização mapeia o limiar em que a densidade semântica de cada nexo passa a blindá-lo contra o apagamento. De acordo com essa hierarquia, a alternância entre preposição explícita e realização nula (Ø) reflete um processo de estratificação (Hopper, 1991).

Compare-se, no entanto, a força da reanálise verbal como motor de variação relacional, evidenciada em (21):

(21) Então, hoje em dia, a televisão, ela vai em busca de programas **que [Ø]** a população fique entorpecida. (Amostra Recontato, falante 11)

Em (21), observa-se o apagamento da preposição *com* que, embora ocupe o polo de maior carga nocional no *continuum* semântico (*cf.* gráfico 3**)**, é suprimida em favor do marcador zero. O verbo *ficar*, originalmente indicador de estado, sofre processo de reanálise funcional e deixa de exigir complementação regencial clássica para atuar como indexador de resultado dentro do cenário ancorado em *programas*. Ocorrências como essa demonstram que a recuperabilidade pode transcender a hierarquia da substância. Quando o quadro predicativo é robusto, o nexo torna-se tão previsível, que a realização fonética passa a ser opcional, porque o sistema prioriza a eficiência do processamento em relações sintáticas subentendidas.

A configuração das relativas observada nesta análise reflete a polarização sociolinguística discutida por Lucchesi (2015, 2025) e Mattos e Silva (2004), que situam o português brasileiro como fruto de contato histórico responsável pela simplificação de regências. Essa realidade é mapeada sob diversas abordagens teóricas em Cavalcante, Marcotulio e Lopes (2014), que sistematizam o status dessas construções na literatura nacional. A força empírica desse panorama manifesta-se nos dados de Araújo e Vieira (2021) com foco no português falado de Fortaleza e na análise de Santos e Silva (2023) sobre o falar de Seabra, cujos achados comprovam a hegemonia das estratégias não padrão em diferentes comunidades. A convergência desses estudos indica que o apagamento preposicional constitui de fato a estratégia predominante no vernáculo, sem, contudo, anular o papel da norma-padrão, que se mantém como referência em contextos de monitoramento estilístico nos termos de Bortoni-Ricardo (2004).

Apesar da forte tendência governada pela previsibilidade do nexo, a análise qualitativa revela dados nos quais a regência verbal é suprimida mesmo diante de preposições com forte carga nocional. Além do *caso do patrão* (*cf.* exemplo 17), já analisado no eixo da animacidade, nota-se o fenômeno no turno de fala a seguir:

(22) É a ex-mulher do homem **que [Ø]** ela está vivendo estava com câncer. (Amostra Recontato, falante 14)

O panorama delineado por esses dados sugere que a sintaxe das relativas no PB não é afetada apenas pelo esvaziamento das preposições, mas também pela ressemantização do predicador. O significado do verbo deve atuar como fator independente, já que verbos como *viver* e *trabalhar* passam a operar como indexadores de cenários ou estados existenciais ao abdicar de sua valência clássica. A perda da substância preposicional libera o verbo para assumir novas nuances funcionais, enquanto a nova carga semântica do predicador torna a preposição dispensável para a clareza do enunciado. O que se observa, portanto, é um ajuste mútuo que impulsiona o PB em direção a uma sintaxe mais econômica e pragmática.

Em uma perspectiva sistêmica, esta reorganização gramatical guarda estreita simetria com o uso estendido do relativo *onde*. Enquanto na estratégia de realização nula (Ø) o verbo assume a função de situar o cenário, como em *homem, programa, trabalho*, no *onde* estendido a responsabilidade é espacializar conceitos abstratos. Para Braga, Manflli e Oliveira (2007), itens como *situação* ou *idade* passam a ser tratados como ancoragem espacial, conforme ilustrado em (23) e (24):

(23) Vivemos uma situação **onde** não há recursos suficientes para atender a todos. (Exemplo adaptado)

(24) Chegou a uma idade **onde** já não aceitava imposições. (Exemplo adaptado)

Do ponto de vista descritivo-funcional, nesses contextos, *onde* opera como marcador relacional geral. O fenômeno articula-se diretamente à perda de transparência preposicional e ao *continuum* semântico proposto neste estudo. Coelho (2001) descreve os novos empregos como resultado da extensão semântica, que reinterpreta o valor espacial do item e permite que ele atue na conceptualização de referentes abstratos como espaços cognitivos. Portanto, seja pela reanálise do verbo ou pela substituição de preposições complexas pelo *onde* curinga, o PB demonstra preferência por estruturas nas quais a relação semântica é depreendida pelo arranjo discursivo.

Cabe notar, por fim, que o sentido do verbo e a natureza do conectivo operam em sincronia. Quanto mais o predicador se especializa na função de moldura situacional, tanto mais o sistema prescinde de marcas de subordinação explícitas, consolidando uma gramática mais ágil e cognitivamente transparente.

Considerações finais

Os resultados desta pesquisa demonstram que a alternância entre a presença explícita e a marcação (Ø) de nexos prepositivos nas fronteiras relacionais das estruturas relativas do português brasileiro não constitui fenômeno residual, mas padrão sistemático e produtivo, consolidado como valor estrutural *default*. Os dados ratificam a tese de que a variação é orientada por um *continuum* semântico e que fatores extralinguísticos perdem relevância diante da força da matriz estrutural, na qual a transparência semântica do nexo e a capacidade de recuperação de sentido ditam o ritmo da mudança.

A distribuição escalar das preposições oferece a chave interpretativa fundamental para os padrões observados. O contraste entre o apagamento quase categórico em contextos de alta inferibilidade, como nos núcleos temporais e na dessemantização de preposições gramaticais, e a resistência de preposições lexicalmente densas, como *com* em contexto específico, aponta para um sistema orientado para economia e previsibilidade.

A variação aqui examinada integra um conjunto mais amplo de ajustes graduais da gramática do PB, ainda que o foco empírico deste estudo tenha ficado restrito à presença ou ausência da preposição no âmbito das orações relativas. Essa dinâmica reflete a mudança em curso, observável inclusive na fala de sujeitos letrados, e evidencia a ecologia de sobrevivência de formas no idioma. A operação do *continuum* semântico revela-se particularmente na categoria dos adjuntos. Nexos de alta densidade lexical resistem ao apagamento por exigirem maior suporte formal, em nítido contraste com preposições em contextos de tempo e espaço, que sofrem apagamento quase categórico, reatestando a gramaticalização estrutural apontada pela literatura (Lima-Hernandes, 2010).

Em última análise, esta pesquisa ratifica a vitória do sentido sobre a configuração sintática. A tendência à transitividade direta não representa defeito ou empobrecimento linguístico, mas otimização gramatical. O fenômeno é orientado pelo princípio da compensação funcional, segundo o qual à medida que a sintaxe se retrai, a semântica se expande.

Referências

ARAÚJO, Aluiza Alves de; VIEIRA, Vinicius da Silva. Não sei onde está a preposição (de) que a oração relativa precisa: as estratégias de relativização no português popular de Fortaleza sob a ótica variacionista. *Confluência*, n. 61, 2021.

ARAÚJO, Edivalda Alves. As orações relativas da língua portuguesa: um estudo descritivo diacrônico. *In*: SANTOS, Wellington de Almeida (org.). *Linguística, letras e artes*: narrativas e discursos na construção cultural 2. Ponta Grossa: Atena, 2025.

BAGNO, Marcos. *Dramática da língua portuguesa:* tradição gramatical, mídia & exclusão social. São Paulo: Edições Loyola, 2000.

BECHARA, Evanildo. *Moderna gramática portuguesa*. 38. ed. Rio de Janeiro: Nova Fronteira, 2009.

BORTONI-RICARDO, Stella Maris. *Educação em língua materna:* a sociolinguística na sala de aula. São Paulo: Parábola Editorial, 2004.

BYBEE, Joan. Mechanisms of change in grammaticization: the role of frequency. *In*: JOSEPH, Brian D.; JANDA, Richard D. (org.). *The handbook of historical linguistics*. Oxford; Malden: Blackwell Publishing, 2003. pp. 602-623.

BYBEE, Joan. *Language, usage and cognition*. Cambridge: Cambridge University Press, 2010.

BRAGA, Maria Luiza; MANFILLI, Keylla; OLIVEIRA, Diego Leite de. É um problema de visão para perto, inevitável após os 40 anos de idade, onde o indivíduo tem dificuldade para ver imagens próximas: Onde em remitências anafóricas a categorias não-locativas. *Revista LinguíStica*, Rio de Janeiro, v. 3, n. 1, 2007. Disponível em: https://revistas.ufrj.br/index.php/rl/article/view/4397.

CAMACHO, Roberto Gomes. Transparência e opacidade na seleção de estratégias de relativização no Português. *Lingüística*, Montevideo, v. 27, n. 1, pp. 47-76, jun. 2012.

CAMACHO, Roberto Gomes. *Da Linguística formal à Linguística social*. São Paulo: Parábola Editorial, 2013.

CASTILHO, Ataliba de. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2010.

CASTILHO, Ataliba Teixeira. Incorporação da oralidade no ensino do português. *In*: LIMA-HERNANDES, Maria Célia; ABREU CHULATA, Kátia (org.). *Língua Portuguesa em foco:* ensino-aprendizagem, pesquisa e tradução. São Paulo: Editora Pensa Multimedia. 2010, pp.113-130.

CAVALCANTE, Silvia Rodrigues de Oliveira; MARCOTULIO, Leonardo Lennertz; LOPES, Célia Regina dos Santos. As orações relativas no português brasileiro: uma trajetória de 30 anos de estudos. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 22, n. 2, pp. 11-48, 2014.

COELHO, Sueli Maria. *Uma análise funcional do onde no português contemporâneo:* da sintaxe ao discurso. Dissertação de Mestrado. Universidade Católica de Minas Gerais, 2001.

DUARTE, Inês. Modo oral e modo escrito, estruturas sintácticas de desenvolvimento tardio e escolarização. *In*: VALENTE, André C.; PEREIRA, Maria Teresa G.(org.). *Língua Portuguesa:* descrição e ensino. São Paulo: Parábola Editorial. 2011.

GOMES, Christina Abreu. *Aquisição e perda da preposição no português do Brasil*. 1996. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1996.

HOPPER, Peter. On some principles of grammaticization. *In*: TRAUGOTT, E. C.; HEINE, B. (eds.). *Approaches to Grammaticalization*. Amsterdam: John Benjamins, 1991. pp. 17-35.

KATO, Mary Aizawa; NUNES, Jairo. Uma análise unificada dos três tipos de relativas restritivas do português brasileiro. *Web-Revista SOCIODIALETO*, Campo Grande: UEMS, v. 4, n. 12, pp. 575-590, 2014.

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

LIMA-HERNANDES, Maria Célia (org.). *Gramaticalização em perspectiva:* cognição, textualidade e ensino. São Paulo: Paulistana Editora, 2010.

LIMA, Elimaria Oliveira; MOLLICA, Maria Cecilia. Para além do erro e da insuficiência escolar: a arquitetura de pronomes e de cláusulas relativas 20 anos após Bortoni-Ricardo. *Revista Leitura*, *[S. l.]*, v. 1, n. 87, pp. 155-174, 2026. DOI: https://doi.org/10.28998/2317-9945.202687.155-174. Disponível em: https://periodicos.ufal.br/revistaleitura/article/view/20493. Acesso em: 14 abr. 2026.

LOPES, Célia Regina dos Santos; SILVA, Bianca Graziela Gomes da Silva. O papel da freqüência na gramaticalização do "que": análise das estratégias de relativização no português do Brasil. *Veredas*, vol.1, Juiz de Fora: UFJF, pp. 60-80, 2007.

MATTOS E SILVA, Rosa Virgínia. *O português são dois*: novas fronteiras entre a norma e o uso. Brasília: UnB, 2004

MARTELOTTA, Mário; VOTRE, Sebastião Josué; CEZARIO, Maria Maura. *Gramaticalização no Português do Brasil:* uma abordagem funcional. Rio de Janeiro: Editora Tempo Brasileiro, 1996.

MARTELOTTA, Mário Eduardo. *Mudança linguística* - uma abordagem baseada no uso. São Paulo: Editora Cortez, 2011.

MELO, Elaine Alves Santos. As estratégias de relativização do português brasileiro à luz da sociolinguística: um olhar para fatores intra e extralinguísticos e para as distintas normas linguísticas. *Scripta***,** v. 7, n. 2, 2016.

MOLLICA, Maria Cecilia. *Estudo da cópia nas construções relativas em português.* Dissertação (Mestrado em Letras) – Departamento de Letras, PUC-Rio, 1977.

MOLLICA, Maria Cecilia. Anáforas em relativas no português do Brasil. *Alfa*: Revista de Linguística, São Paulo, v. 41, n. especial, pp. 171-179, 1997.

MOLLICA, Maria Cecilia. Relativas em tempo real no português brasileiro contemporâneo. *In*: PAIVA, Maria da Conceição; DUARTE, Maria Eugenia Lamogia. (org.). *Mudança linguística em tempo real*. Rio de Janeiro: Contracapa/FAPERJ, 2003, pp. 129-138.

MOLLICA, Maria Cecilia; AVELAR, Daillane; BATISTA, Hadinei. Variação, mudança, letramento e tecnologia. *Diadorim: Revista de Estudos Linguísticos e Literários*, Rio de Janeiro, v. 20, n. especial, pp. 561-572, dez. 2018.

PAIVA, Maria da Conceição; PEREIRA, Maria Hermenegilda. Estatuto sintático das orações introduzidas pelas construções (prep)+det+N temporal+(prep)+que. *Veredas*: Revista de Estudos Linguísticos, Juiz de Fora, v. 8, n. 1 e 2, pp. 245-262, jan./dez. 2004.

PEREIRA, Maria Hermenegilda; BRAGA, Maria Luiza; PAIVA, Maria da Conceição. Gramaticalização das construções (prep1)+(det)+n+(prep2)+que. *In*: VITRAL, Lorenzo; COELHO, Sueli (org.). *Estudos de processos de gramaticalização*: metodologias e aplicações. São Paulo: Mercado de Letras, 2010. pp. 173-200.

ROCHA, Eva Maria Nery. Apagamento de preposições em orações completivas e relativas na fala popular e culta em Salvador. *In*: SILVA LOPES, Norma da (org.) *Tópicos em variação em Português*. Curitiba: Editora Appris e Editora Prismas. 2011.

SANTOS, André Poltronieri. *As construções relativas na fala de Vitória/ES:* uma perspectiva sociolinguística. 2020. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal do Espírito Santo, Vitória, 2020.

SANTOS, Wellington de Almeida. *As orações relativas da língua portuguesa:* um estudo descritivo-diacrônico. Ponta Grossa: Atena, 2025.

SANTOS, Elias de Souza; SILVA, Jéssica Carneiro da. As construções relativas no falar seabrense: uma análise multinomial (politômica). *Linguagem em (Re)vista*, v. 18, n. 2, pp. 110-134, 2023.

TARALLO, Fernando. *Relativization strategies in Brazilian Portuguese.* 1983. Tese (Doutorado) – University of Pennsylvania, Philadelphia, 1983.

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. *In*: ROBERTS, Ian; KATO, Mary A. (org.). *Português brasileiro:* uma viagem diacrônica (Homenagem a Fernando Tarallo). Campinas: Editora da Unicamp, 1993. pp. 69-105.

http://dx.doi.org/10.31513/linguistica.2026.v22n1a72437

Análise sociolinguística da lenição de /d/ em gerúndios na região metropolitana do Rio de Janeiro

Sociolinguistic analysis of the lenition of /d/ in gerunds in the metropolitan region of Rio de Janeiro

Matheus Victor Alves Pereira[1]

Eliete Figueira Batista da Silveira[2]

RESUMO

Este estudo investiga o processo de lenição da consoante /d/ no morfema de gerúndio na fala da região metropolitana do Rio de Janeiro, à luz da Sociolinguística Variacionista (Weinreich; Labov; Herzog, 2006 [1978]; Labov, 2008 [1972]). Embora o fenômeno já tenha sido abordado em pesquisas anteriores, este trabalho contribui ao analisá-lo a partir de um *corpus* de fala espontânea com ampla estratificação social, contemplando localidades estratégicas como Nova Iguaçu e Copacabana. O objetivo geral é examinar o processo de lenição de /d/ em formas gerundiais, analisando a influência de variáveis sociais e linguísticas, bem como comparar os resultados obtidos com estudos prévios, a fim de avaliar o status do fenômeno no português brasileiro (PB). Parte-se da hipótese de que o processo de abrandamento em gerúndios constitui uma mudança em curso nessa variedade. Os dados foram extraídos do *Corpus* Concordância, do Projeto CORPORAPORT (Vieira; Brandão, 2018), composto por entrevistas com falantes das duas localidades, Nova Iguaçu e Copacabana. A codificação seguiu os pressupostos metodológicos da Sociolinguística Variacionista e o tratamento estatístico foi realizado com o auxílio do software R (R Core Team, 2022). Os resultados revelam alta frequência de lenição, com 64% de redução e 36% de manutenção do segmento. A análise quali-quantitativa identificou como condicionamentos relevantes fatores como grau de escolaridade, contexto fonético-fonológico seguinte, frequência lexical e informante. Os achados evidenciam o papel de variáveis sociais e linguísticas nos processos de variação e possível mudança associados à lenição da consoante oclusiva alveolar /d/ em gerúndios na fala da região metropolitana do Rio de Janeiro.

PALAVRAS-CHAVE: Variação. Sociolinguística. Português Brasileiro. Redução de gerúndio.

ABSTRACT

This study investigates the process of lenition of the consonant /d/ in the gerund morpheme in the speech of the metropolitan region of Rio de Janeiro, within the framework of Variationist Sociolinguistics (Weinreich; Labov; Herzog, 2006 [1978]; Labov, 2008 [1972]). Although the phenomenon has been addressed in previous research, this study contributes by analyzing it based on a *corpus* of spontaneous speech with broad social stratification, including strategic localities such as Nova Iguaçu and Copacabana. The main objective is to examine the lenition of /d/ in gerund forms by analyzing the influence of social and linguistic variables, as well as to compare the results with those of previous studies in order to assess the status of the phenomenon in Brazilian Portuguese (BP). The study is based on the hypothesis that gerund lenition constitutes an ongoing change in this variety. The data was drawn from the Concordance *Corpus* of the CORPORAPORT Project

[1] Universidade Federal do Rio de Janeiro (UFRJ), matheusvictor347@letras.ufrj.br, https://orcid.org/0000-0003-1281-3618.

[2] Universidade Federal do Rio de Janeiro (UFRJ), elietesilveira@letras.ufrj.br, https://orcid.org/0000-0001-6928-2614.

Recebido em 05/02/2026 - Aceito em 10/06/2026

(Vieira; Brandão, 2018), which consists of interviews with speakers from the two locations, Nova Iguaçu and Copacabana. Data coding followed the methodological assumptions of Variationist Sociolinguistics, and statistical analysis was carried out using the R software (R Core Team, 2022). The results reveal a high frequency of lenition, with 64% reduction and 36% retention of the segment. The qualitative-quantitative analysis identified factors such as level of education, following phonetic-phonological context, lexical frequency, and speaker as relevant conditioning factors. These findings highlight the role of social and linguistic variables in processes of variation and possible change associated with the lenition of the alveolar plosive consonat /d/ in gerunds in the speech of the metropolitan region of Rio de Janeiro.
KEYWORDS: Variation. Sociolinguistics. Brazilian Portuguese. Gerund reduction.

1. Introdução

A redução em gerúndios é um fenômeno amplamente atestado no português brasileiro (PB), tendo sido investigado em diversos estudos sociolinguísticos (Mollica; Mattos, 1992; Costa, 2009; Sousa, 2009; Ferreira, 2010; Nascimento; Mota, 2004; Almeida; Oliveira, 2017; entre outros). O processo consiste na simplificação da sequência [ndo] para [no], como em *falando ~ falano*, e tem recebido diferentes denominações na literatura, tais como apagamento, supressão, redução, assimilação e lenição.

Neste trabalho, adota-se o termo lenição, por se tratar de uma designação mais abrangente, capaz de contemplar tanto o apagamento completo do segmento /d/ quanto seu abrandamento parcial. Tal escolha evita interpretações restritivas associadas a termos como apagamento ou supressão, que pressupõem a elisão total do segmento, não captando a complexidade do fenômeno sem o apoio de análises fonético-acústicas.

O objetivo geral deste estudo é analisar a lenição da consoante oclusiva alveolar /d/ no morfema de gerúndio na fala da região metropolitana do Rio de Janeiro. Especificamente, busca-se investigar a influência de variáveis sociais (sexo, faixa etária e escolaridade) e linguísticas (extensão do vocábulo e contexto fonético-fonológico seguinte), bem como discutir os condicionamentos que atuam sobre o fenômeno a partir de uma análise quali-quantitativa de dados de fala espontânea. Além disso, os resultados são comparados aos de estudos anteriores, com o intuito de avaliar o status da lenição em gerúndios no português brasileiro (PB), partindo-se da hipótese de que se trata de uma mudança em curso.

Apesar de pesquisas recentes sobre o tema no Rio de Janeiro (Melo; Silva, 2022; Gonçalves; Melo, 2024), ainda se mostra relevante a análise do fenômeno com base em *corpus* socialmente estratificado, que contemple localidades estratégicas da região metropolitana, como Nova Iguaçu e Copacabana. O estudo fundamenta-se nos pressupostos teórico-metodológicos da Sociolinguística Variacionista (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972]), abordagem que permite compreender a variação linguística a partir da interação entre fatores internos e externos ao sistema linguístico.

O trabalho inicia com a apresentação do processo de lenição de /d/ em gerúndios no português brasileiro (PB), incluindo uma discussão sobre a escolha da região metropolitana do Rio de Janeiro e

um panorama do fenômeno em outras variedades do português; em seguida, apresenta os fundamentos teóricos da Sociolinguística Variacionista e do paradigma quantitativo; posteriormente, descreve os procedimentos metodológicos adotados; e, por fim, expõe e discute os resultados relativos aos condicionamentos sociais e linguísticos da lenição.

2. Variação /ndo/ ~ /no/ no PB

A lenição da oclusiva /d/ em gerúndios resulta de um processo de assimilação, no qual duas consoantes passam a ser realizadas como um único som, de forma gradual (-nd > -nn > -n). Esse fenômeno ocorre devido à proximidade articulatória entre os fonemas alveolares ou dentais /n/ e /d/, que, por serem produzidos na mesma zona de articulação[3], se tornam iguais ou bastante semelhantes no modo de articulação (Ferreira, 2010, p. 28).

Cristófaro-Silva (1996) observa que as formas reduzidas como *comendo ~ comeno*, *falando ~ falano* e *mentindo ~ mentino* apresentam uma estrutura fonológica composta por uma raiz verbal seguida pela terminação de gerúndio /Ndo/, o que evidencia a operação desse processo fonológico.

> Quando temos [fa'lãdu] a consoante nasal cumpre o seu papel de nasalizar a vogal precedente e /do/ ocorre como a sílaba final. Quando temos [fa'lãnu] a consoante nasal cumpre o seu papel de nasalizar a vogal precedente, mas a consoante nasal irá também ocupar a posição de consoante inicial da sílaba final. (Cristófaro-Silva, 1996, p. 66)

As formas gerundiais formam palavras com travamento nasal. A nasalidade fonêmica das vogais no português brasileiro (PB) é um fenômeno obrigatório que se manifesta de forma uniforme em todo o território nacional.

Câmara Jr. (2015 [1970]) interpreta que as vogais nasais são uma combinação de dois elementos: uma vogal oral e um arquifonema nasal /N/. Nessa visão, o estruturalista compreende que a vogal antecedente é, na verdade, nasalizada pelo elemento nasal em coda silábica. Assim, o português apresentaria apenas vogais orais e vogais nasalizadas. Essa perspectiva é compartilhada por Cristófaro-Silva (1996).

Em visão contrária, Callou e Leite (1990) consideram as vogais nasais como fonemas distintos das vogais orais no português brasileiro (PB). Por meio da análise de pares mínimos, as autoras identificam a existência de cinco vogais nasais no sistema fonológico do português, sendo elas /ã, ẽ, ĩ, õ, ũ/, ao lado de 7 vogais orais /a, ɛ, e, i, ɔ, o, u/. Assim, de acordo com as linguistas, o inventário vocálico do português seria composto por 12 vogais: 7 orais e 5 nasais.

Tais considerações são motivadas pelo fato de o fenômeno de lenição nas formas de gerúndio provocar uma reestruturação silábica (CVC ~ CV), padrão mais recorrente no português. Além disso, a lenição da consoante oclusiva /d/ transforma a nasalidade fonológica em nasalidade fonética, ou

[3] A zona de articulação refere-se ao espaço na cavidade oral em que ocorre o contato ou a maior aproximação entre os articuladores (órgãos que produzem os sons da fala), como a língua, os lábios, os dentes e o palato, durante a produção de um som. Em termos mais específicos, trata-se do local em que um som é articulado dentro do trato vocal.

seja, a nasal heterossilábica [fɐ.ˈlã.**n**ʊ] ao mesmo tempo nasaliza a vogal antecedente (na estrutura de superfície), mas também funciona como ataque na sílaba formada, o que se entende por ambissilabicidade (na estrutura profunda).

Com o intuito de demonstrar que as formas gerundiais apresentam maior variabilidade, Mollica e Mattos (1992, p. 58) destacam que os verbos, de modo geral, constituem o grupo de itens mais sujeito a processos de inovação fonológica. A diferença observada entre a taxa de variação nos gerúndios e nos verbos no presente do indicativo pode ser atribuída ao fato de que o sufixo *-ndo* não integra o radical verbal nas formas gerundiais, sendo adicionado a ele. Nesse sentido, o fonema /d/, quando faz parte do radical – como ocorre em nomes, conjunções ou verbos no presente – tende a resistir à redução, por estar incorporado à unidade base da palavra. Já no caso da desinência *-ndo*, /d/ incide sobre um elemento afixal, o que favorece sua lenição.

Temos acima a menção a um aspecto relevante sobre a influência estrutural no processo de inovação fonológica, especialmente em relação à resistência do /d/ à queda. A distinção entre a base (raiz) do vocábulo e os elementos afixais explica, em parte, a maior suscetibilidade da desinência *-ndo* à redução. Por ser um elemento afixado, o /d/ no gerúndio está mais exposto às variações linguísticas do que quando integra o radical da palavra, como ocorre em nomes, conjunções ou formas do presente do indicativo.

Assim, de acordo com Cristófaro-Silva (1996, p. 67), a redução de gerúndios caracteriza-se por três aspectos principais. Primeiro, o fenômeno aplica-se exclusivamente à categoria gramatical do gerúndio. Por exemplo, é possível observar a redução em "Eu estou vendo você" (*vendo*, gerúndio do verbo *ver*), mas não em "Eu vendo livros" (*vendo*, presente do indicativo do verbo *vender*). Depois, a redução requer tanto informações fonológicas, relacionadas à sequência consonantal, quanto morfológicas, relacionadas à marcação do gerúndio. Por fim, o processo não resulta em uma reorganização lexical, mantendo-se limitado à dimensão fonológica e morfológica da língua.

2.1. Por que a região metropolitana do Rio de Janeiro?

A principal motivação para a escolha da região metropolitana do Rio de Janeiro foi a carência de estudos que tratam especificamente do fenômeno de lenição da consoante oclusiva /d/ em gerúndios nessa região. Parte-se do trabalho pioneiro de Mollica e Mattos (1992) sobre a assimilação *ndo* > *no*, com dados de 64 falantes não universitários residentes no município do Rio de Janeiro, participantes da Amostra Censo. Considerando outras categorias lexicais a fim de obter uma análise mais abrangente do fenômeno, as autoras, ao analisarem vocábulos invariáveis, destacaram nomes próprios, nomes comuns, a maioria dos adjetivos, bem como o numeral e a conjunção "segundo". Dados os resultados de Mollica e Mattos, observou-se uma maior incidência da assimilação em verbos no gerúndio, motivo pelo qual este trabalho se limita a essa categoria.

O *corpus* analisado neste trabalho busca contemplar a diversidade linguística ao incorporar dados de duas áreas distintas da região metropolitana do Rio de Janeiro: Nova Iguaçu, um município situado

na Baixada Fluminense, e Copacabana, um famoso bairro da Zona Sul carioca. Embora relativamente distantes, essas áreas mantêm uma forte conexão devido ao intenso trânsito diário de pessoas que se deslocam pela região metropolitana para trabalhar, estudar ou realizar atividades de lazer. Nesse sentido, entende-se que é possível realizar uma análise rica e representativa das particularidades linguísticas presentes no dialeto carioca, ressaltando como fatores sociais e linguísticos interagem na constituição da linguagem.

2.2. Panorama da lenição de /d/ em gerúndios em variedades do PB

No estudo pioneiro sobre a assimilação do /d/ em gerúndios, Mollica e Mattos (1992) verificaram que esse processo apresenta estabilidade no português brasileiro (PB). As pesquisadoras observaram, ainda, que o fenômeno não ocorre de maneira uniforme no léxico da língua, seja porque nem todos os itens suscetíveis à alteração são efetivamente afetados, seja porque o processo não abrange todas as categorias gramaticais.

A pesquisa foi baseada em dados coletados junto a 64 falantes não universitários residentes no município do Rio de Janeiro, participantes da Amostra Censo. Entre os vocábulos invariáveis analisados, destacam-se nomes próprios, nomes comuns, a maioria dos adjetivos, além do numeral e da conjunção "segundo". Incorporando itens de diferentes categorias gramaticais, as autoras articulam a descrição e análise sociolinguística laboviana com uma abordagem difusionista[4].

Mollica e Mattos (1992) propõem uma classificação dos dados relacionados ao contexto "ndo" com base em uma combinação de categoria morfológica e grau de variabilidade do fenômeno, em vez de seguirem critérios tradicionais. Na tabela apresentada pelas autoras (p. 55), embora os percentuais tratem da preservação de /d/, infere-se que a maior frequência de lenição do segmento /d/ ocorre nos gerúndios, com baixo percentual de conservação. O destaque vai para os gerúndios de caráter fático, ou seja, aqueles que podem fazer parte dos marcadores discursivos: E olha que eu conheço, tá *entendendo*? (p. 55).

Tabela 1: Percentual de preservação de /d/ em 'ndo' de acordo com o item.

Tipo de item gramatical	Quantidade	Percentual
Gerúndio	1.606/2.602	62%
Verbo no presente	56/62	90%
Substantivo comum	292/314	93%
Gerúndio fático	32/70	46%
Outros	858/1.053	81%

Fonte: Mollica e Mattos, 1992, p. 55.

[4] Diversas pesquisas variacionistas têm investigado a variação linguística no nível fonológico, baseando-se em duas hipóteses principais: a neogramática e a difusionista. Essas hipóteses divergem quanto ao entendimento da propagação da mudança ou variação sonora. Sob a ótica da difusão lexical, grosso modo, a mudança estaria centrada na palavra, em vez de no fonema.

No âmbito da análise variacionista, Mollica e Mattos (1992) identificaram a variável "extensão do vocábulo" (classificada em dissílabos, trissílabos e polissílabos) como a mais significativa para o fenômeno estudado. Os resultados indicaram que vocábulos mais extensos apresentam maior propensão à assimilação.

Quanto ao contexto fonético-fonológico seguinte, as autoras verificaram que o segmento /d/ é mais frequentemente preservado quando ocorre antes de uma pausa, já que, nesse ambiente, a ausência de um fonema adjacente não favorece a assimilação. Além disso, o condicionamento fonético-fonológico do contexto que promove a assimilação é aplicável apenas a uma categoria gramatical específica (o gerúndio).

Mollica e Mattos (1992) identificam as formas gerundiais como aquelas que exibem maior variabilidade no fenômeno analisado, sendo os verbos, de modo geral, os itens mais suscetíveis a processos de inovação fonológica. As autoras apontam que a diferença nas taxas de variação entre gerúndios e verbos no presente do indicativo pode ser atribuída ao fato de o morfema *-ndo* não fazer parte do radical nas formas gerundiais, embora se apresente como um elemento afixado. Nesse sentido, o fonema /d/, ao integrar a raiz da palavra, como ocorre em nomes, conjunções ou verbos no presente, tende a apresentar maior resistência à queda, por estar diretamente vinculado à base lexical. Já quando associado à desinência *-ndo*, /d/ passa a incidir sobre uma unidade afixal da palavra, o que pode influenciar sua maior suscetibilidade à variação (p. 58).

A preservação quase integral do segmento /d/ em nomes próprios está alinhada, como mostram as pesquisadoras, com as observações de Philips (1984) e Oliveira (1991), que destacam essa categoria como uma barreira lexical à entrada da variação na língua. Além disso, categorias como adjetivos, numerais e conectores (como "quando" e "segundo") também apresentam elevados índices de preservação, corroborando a hipótese de que algumas classes gramaticais resistem mais à inovação do que outras.

Focando a fala de João Pessoa, Martins (2004) analisou a lenição do fonema /d/ em final de palavra sob as perspectivas da Sociolinguística Variacionista e da Fonologia não linear. Os resultados indicaram predominância da lenição, especialmente em verbos no gerúndio. Entre as variáveis analisadas, destacaram-se a extensão do vocábulo, que favorece a lenição em palavras com mais de duas sílabas, e fatores sociais, sobretudo o sexo, com maior tendência à lenição entre os homens. Labov (2008 [1972]) explica essa diferença ao associá-la à maior percepção feminina dos sinais de estratificação social. Além disso, ele destaca que os papéis sociais distintos atribuídos a homens e mulheres geram expectativas diferentes quanto ao comportamento linguístico de cada grupo.

Sousa (2009) analisou a interferência da fala na escrita de alunos do 3º ao 5º ano do Ensino Fundamental, em Fortaleza, com foco na lenição da oclusiva /d/ e da vibrante final /r/. Os resultados indicaram influência parcial do sexo masculino e da extensão do vocábulo na lenição de /d/, enquanto o tipo verbal e o nível de escolaridade condicionaram a lenição de /r/. A orientação pedagógica pontual não reduziu significativamente as ocorrências, embora as médias de erro tenham sido baixas, sugerindo que tais fenômenos não configuram problemas ortográficos graves. O autor destaca a

relevância de fatores fonológicos e sugere a investigação desses fenômenos em séries iniciais, especialmente no 2º ano.

Ferreira (2010) investigou a lenição da consoante /d/ em formas de gerúndio no português falado em São José do Rio Preto (SP), com base na Sociolinguística Variacionista e na Fonologia Lexical. A análise de 999 ocorrências, extraídas do banco Iboruna, revelou lenição em 72% dos casos, mais frequente entre falantes jovens, do sexo masculino e com menor escolaridade. A autora demonstra, por meio de análise acústica, que a lenição envolve a perda da closura do /d/, com substituição por traços semelhantes aos do segmento nasal /n/, e conclui que o fenômeno ocorre exclusivamente em verbos com morfema de gerúndio.

Nascimento *et al.* (2013) analisaram a lenição de /d/ em gerúndios no falar popular de Fortaleza e identificaram como principais condicionadores a escolaridade, a faixa etária e os contextos fonéticos. A lenição foi favorecida por falantes com menor escolaridade e por determinados ambientes fonéticos, enquanto a extensão do vocábulo não se mostrou relevante. Os autores interpretam o fenômeno como variação estável, embora o comportamento feminino tenha divergido de tendências sociolinguísticas mais gerais.

No Pará, Araújo, Aragão e Almeida (2016) investigaram a realização de /d/ no contexto -ndo a partir de dados do ALiSPA. Os resultados indicaram baixa frequência de lenição, mais associada a falantes jovens e do sexo masculino, mas as autoras ressaltam limitações metodológicas do *corpus*, o que pode ter influenciado os índices observados.

Almeida e Oliveira (2017), ao analisarem o falar de Maceió/AL, verificaram que a lenição de /d/ em gerúndios é favorecida por falantes do sexo masculino, por vocábulos mais extensos e por contextos fonético-fonológicos específicos, evidenciando a atuação conjunta de fatores linguísticos e sociais.

Freitag, Cardoso e Pinheiro (2018), com dados do estado de Sergipe, confirmaram que a lenição de /d/ se restringe ao morfema de gerúndio e é condicionada principalmente por fatores sociais, como escolaridade, sexo e grau de monitoramento da fala. O estudo reforça que itens não verbais apresentam maior preservação do segmento do que formas em gerúndio.

Ferreira (2023), ao investigar a fala da Cidade de Goiás, observou predominância da forma lenizada em gerúndios, fortemente condicionada por fatores sociais, sobretudo escolaridade e faixa etária, enquanto variáveis linguísticas tiveram papel secundário, apontando para um fenômeno estável.

No Rio de Janeiro, Gonçalves e Melo (2024) identificaram predominância da variante nasal [n] no contexto -ndo em diferentes amostras sociolinguísticas, sugerindo um possível processo de mudança linguística em curso. Em estudo perceptual, Melo e Silva (2022) não encontraram avaliação social significativamente diferenciada entre as variantes [d] e [n], indicando que a variante nasal em contexto de gerúndio *-ndo* não parece ser percebida, nesse grupo social, como estereótipo associado à baixa escolaridade ou a falantes de classes sociais menos favorecidas[5].

[5] Acredita-se que seja necessário um aprofundamento da questão da avaliação social da lenição do /d/ em gerúndios, pois esse fenômeno linguístico muitas vezes sofre um julgamento estigmatizado, estereotipado e preconceituoso, especialmente por meio de sua representação em programas humorísticos.

3. Sociolinguística Variacionista

Este estudo se apoia na abordagem teórico-metodológica da Sociolinguística Variacionista, conhecida também como Teoria da Variação e Mudança (TVM), que teve sua concepção inicial na obra de Weinreich, Labov e Herzog (2006 [1968]). Nesse sentido, esta corrente teórica se preocupa em analisar as variações linguísticas a partir de contextos reais de uso, buscando entender as razões que podem motivar a variação/mudança linguística. Por essa razão, a concepção de língua parte de um caráter heterogêneo, que, segundo Labov (2008 [1972]), é uma característica inerente à língua. Assim, a Sociolinguística Variacionista entende que língua e sociedade são inseparáveis e se dedica a compreender os fatores essenciais que impulsionam a variação, focando na importância de cada um desses fatores para a construção do panorama da variação.

Labov (2008 [1972]) destaca a importância de estudar as regras linguísticas em sua forma, organização em sistemas e evolução ao longo do tempo, enfatizando a necessidade de situar a análise no contexto social. O autor questiona por que seria necessária uma abordagem linguística baseada em fatores sociais, argumentando que o ponto de partida natural para qualquer estudo da língua deve ser a análise da comunicação real entre falantes nativos em seu cotidiano. Para Labov, a interação social não é apenas um aspecto adicional, mas parte essencial para compreender plenamente o funcionamento e a dinâmica da língua.

As formas escolhidas por um falante na hora de falar ou escrever estão relacionadas não só ao próprio falante como também ao mundo que o cerca, isto é, ao contexto interacional de que ele está participando. Dessa relação intrínseca que se estabelece entre língua e sociedade, surge o conceito de heterogeneidade ordenada da língua. Por sua vez, a Sociolinguística compreende a língua como um sistema estruturado, que varia sem que as pessoas tenham problemas de comunicação. Trata-se, portanto, de um mecanismo heterogêneo e essa heterogeneidade deve ser buscada na própria comunidade de fala. Sendo assim, seria a noção de heterogeneidade o fator que distingue a corrente variacionista da estruturalista, que considerava a homogeneidade da língua falada.

A análise sociolinguística exige a investigação dos fatores linguísticos e socioculturais que influenciam o fenômeno estudado, ampliando o foco para incluir a influência dos aspectos sociais no processo de variação linguística. Assim, tal abordagem se baseia na premissa de que a mudança linguística só pode ser compreendida considerando-se o contexto social da comunidade de fala, uma vez que o sistema é constantemente moldado pelas dinâmicas e pressões dessa comunidade (Labov, 2008 [1972]). A partir disso, entende-se que o processo de variação e mudança linguística não pode ser compreendido sem considerar o contexto social da comunidade de fala, que exerce contínua pressão sobre o sistema linguístico. Nesse sentido, tais premissas teóricas orientam a organização metodológica adotada para o tratamento e a análise dos dados, apresentados na seção seguinte.

4. Aspectos metodológicos

Conforme discutido na seção anterior, a presente pesquisa fundamenta-se nos princípios da Sociolinguística Variacionista (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972]), quadro teórico que dispõe de bases consistentes para análise e descrição de um fenômeno linguístico em uma comunidade de fala. A Teoria da Variação e Mudança, por sua vez, apresenta uma metodologia sólida de pesquisa, uma vez que combina técnicas rigorosas de coleta e análise de dados, a partir da escolha de seus grupos de fatores, com uma compreensão profunda das interações entre língua e sociedade – o que permite revelar padrões de variação e mudança linguística em contextos reais de uso.

Nessa perspectiva, a Sociolinguística Variacionista parte do princípio de que língua e sociedade estão intrinsecamente conectadas, buscando compreender os fatores que influenciam a variação linguística e a relevância de cada um desses elementos na organização do cenário linguístico variável. Por isso, estudos dessa abordagem geralmente adotam uma perspectiva quantitativa, envolvendo etapas de coleta, codificação e análise de dados por meio de ferramentas computacionais como o *GoldVarb X* e o *R-brul*.

O principal objetivo desta investigação é compreender como variáveis sociais e linguísticas influenciam a aplicação da regra de lenição do fonema /d/ no morfema de gerúndio *-ndo* na região metropolitana do Rio de Janeiro.

4.1. Nova Iguaçu e Copacabana: descrição e estratificação da amostra

A coleta de dados desta pesquisa utilizou a Amostra Concordância (www.corporaport.ufrj.br), parte do Projeto CORPORAPORT (Vieira; Brandão, 2018), resultado da colaboração entre a UFRJ e a Universidade de Lisboa. As entrevistas foram realizadas entre 2008 e 2010 com 36 informantes distribuídos em duas áreas estratégicas da região metropolitana do Rio de Janeiro: Nova Iguaçu, na Baixada Fluminense, e Copacabana, na Zona Sul da capital.

Nova Iguaçu é o maior município da Baixada Fluminense, com cerca de 795 mil habitantes em 2010, crescimento industrial recente e acesso crescente ao Ensino Superior. Copacabana, bairro da Zona Sul do Rio, possui elevado Índice de Desenvolvimento Humano (IDH), refletindo melhores condições socioeconômicas (IBGE, 2012). As diferenças marcantes entre os locais – em termos de histórico socioeconômico, densidade populacional e infraestrutura urbana –, aliadas ao intenso fluxo de pessoas entre eles, tornam essas localidades estratégicas para o estudo de fenômenos linguísticos à luz de dinâmicas sociais.

Os falantes foram estratificados de acordo com a faixa etária (18-35, 36-55, 56-75 anos), o grau de escolaridade (Ensino Fundamental, Ensino Médio e Ensino Superior) e o sexo[6] (*cf.* quadro 1).

[6] Nesta pesquisa, considera-se o sexo biológico, sem abordar a discussão terminológica entre sexo e gênero.

Quadro 1: Estratificação da amostra

Sexo	Faixa etária	Nível de escolaridade
M - Masculino	A - 18 a 35 anos	1 - Ensino Fundamental
F - Feminino	B - 36 a 55 anos	2 - Ensino Médio
	C - 56 a 75 anos	3 - Ensino Superior

Fonte: Elaborado pelos autores.

Para cada dado exemplificado, será especificada a localidade de origem do informante: COP – Copacabana; NIG – Nova Iguaçu. Da mesma forma, as informações analisadas serão acompanhadas de indicações referentes ao perfil social dos participantes, seguindo os códigos apresentados no quadro 2 e conforme estratificação do próprio *corpus*.

Quadro 2: Estratificação da amostra.

Nova Iguaçu			
Escolaridade	**Nível 1**	**Nível 2**	**Nível 3**
Faixa A	NIG-A-1-M NIG-A-1-F	NIG-A-2-M NIG-A-2-F	NIG-A-3-M NIG-A-3-F
Faixa B	NIG-B-1-M NIG-B-1-F	NIG-B-2-M NIG-B-2-F	NIG-B-3-M NIG-B-3-F
Faixa C	NIG-C-1-M NIG-C-1-F	NIG-C-2-M NIG-C-2-F	NIG-C-3-M NIG-C-3-F
Copacabana			
Escolaridade	**Nível 1**	**Nível 2**	**Nível 3**
Faixa A	COP-A-1-M COP-A-1-F	COP-A-2-M COP-A-2-F	COP-A-3-M COP-A-3-F
Faixa B	COP-B-1-M COP-B-1-F	COP-B-2-M COP-B-2-F	COP-B-3-M COP-B-3-F
Faixa C	COP-C-1-M COP-C-1-F	COP-C-2-M COP-C-2-F	COP-C-3-M COP-C-3-F

Fonte: Elaborado pelos autores.

Cada entrevista seguiu um formato de diálogo entre entrevistador e entrevistado, abordando uma ampla gama de tópicos, incluindo opiniões sobre a cidade onde os informantes vivem, aspectos sociopolíticos e outros temas relacionados ao cotidiano e à experiência pessoal e profissional dos participantes.

4.2. Delimitação das variáveis e levantamento dos dados

O levantamento das ocorrências de gerúndio foi realizado em dois passos: seleção dos itens nos inquéritos escritos e escuta detalhada das gravações pelo método de análise de oitiva, amplamente usado em estudos sociolinguísticos para identificar diferentes realizações sonoras.

A variável dependente é a *lenição* ou *manutenção* do fonema /d/ nos gerúndios. As variáveis independentes incluem fatores linguísticos – extensão do vocábulo (dissílabo, trissílabo ou polissílabo) e contexto fonético-fonológico subsequente (/t/, /d/, /n/, outras consoantes, vogais ou pausas) – e sociais, como sexo, faixa etária, escolaridade e ponto de inquérito (Nova Iguaçu ou Copacabana).

Estudos anteriores indicam divergências quanto à influência do sexo, com maior lenição entre homens (Ferreira, 2010; Almeida; Oliveira, 2017) ou mulheres (Nascimento *et al*., 2013), e mostram que a escolaridade é um fator determinante: quanto maior o grau de escolaridade, menor a ocorrência da lenição (Ferreira; Tenani; Gonçalves, 2012). A extensão do vocábulo também se mostra relevante, com palavras mais longas, especialmente polissílabas, mais propensas à omissão do /d/ (Mollica; Mattos, 1992; Sousa, 2009; Almeida; Oliveira, 2017; Oliveira, 2014). A conjugação verbal foi descartada como variável, dada sua baixa relevância estatística em estudos anteriores.

4.3. Recolha de dados, critérios e modelo de análise.

Para a análise dos dados, foram considerados os 36 inquéritos (sendo 18 referentes à região de Nova Iguaçu e 18 de Copacabana), a partir dos quais se identificaram 2.392 ocorrências de gerúndio. Do total, 1.436 ocorrências de gerúndio foram encontradas na amostra de Nova Iguaçu e 956 nas entrevistas gravadas em Copacabana. Os dados foram organizados em planilha do Excel, permitindo uma melhor sistemática que contemplasse informações relevantes para a investigação. Tal organização incluiu os seguintes critérios: i) identificação do informante, ii) a palavra em que ocorreu o gerúndio (a partir da qual consideramos também a frequência do item), iii) a observação de lenição ou manutenção do fonema /d/, e as variáveis sociais e linguísticas, na seguinte ordem: iv) sexo, v) faixa etária, vi) escolaridade, vii) extensão do vocábulo, viii) contexto fonético-fonológico seguinte e ix) ponto de inquérito.

Os dados assim organizados foram submetidos ao pacote estatístico *R-brul*, uma ferramenta amplamente utilizada em análises sociolinguísticas e integrada ao software R. A metodologia estatística adotada incluiu a aplicação de regressão logística binária de efeitos mistos, na qual as variáveis *informante* e *frequência de item*[7] foram codificadas como efeito aleatório e as variáveis *escolaridade* e *contexto fonético-fonológico seguinte* foram selecionadas como efeitos fixos.

Efeitos fixos e aleatórios são conceitos estatísticos que tratam da relação entre os efeitos individuais específicos e as variáveis independentes. Os efeitos aleatórios assumem que esses efeitos individuais não estão correlacionados com as variáveis independentes. Já os efeitos fixos pressupõem uma correlação entre os efeitos individuais e as variáveis independentes, atribuindo a cada unidade um intercepto fixo, sendo mais adequados para analisar efeitos causais de variáveis explicativas ao longo do tempo. O modelo estatístico adotado se revela eficaz para a análise de variáveis categóricas dependentes, permitindo prever a probabilidade de ocorrência de um dado fenômeno linguístico.

[7] A variável frequência de item, neste estudo, é definida com base no total de ocorrências (*tokens*) de cada item presente na amostra, bem como no seu respectivo índice de aplicação da regra de lenição. Como etapa futura, pretende-se refinar essa análise, considerando não apenas a frequência bruta dos itens, mas também a frequência específica da lenição em cada um deles, o que possibilitará uma compreensão mais detalhada dos fatores que influenciam o processo variável em estudo.

A análise dos resultados deste trabalho será conduzida com base nos percentuais calculados, nos pesos relativos (PR) e nos *logodds*[8] gerados pelo programa, permitindo uma interpretação mais precisa e detalhada dos dados obtidos no estudo.

5. Análise dos resultados

A análise de 2.392 formas verbais no gerúndio, extraídas de falas espontâneas de 36 informantes, revelou que a regra de redução é amplamente aplicada na variedade linguística da região metropolitana do Rio de Janeiro. Do total de ocorrências, 1.537 (64%) apresentaram lenição do /d/, enquanto 855 (36%) mantiveram o segmento (*cf.* tabela 2). A análise que fazemos aqui tem por objetivo detalhar como o índice de aplicação da lenição se distribui na comunidade estudada, buscando observar os contextos sociais em que o fenômeno é mais recorrente. Para tanto, a exposição dos resultados será organizada com base na ordem de relevância dos grupos de fatores selecionados pelo programa *R-brul*.

Tabela 2: Frequência das variantes *manutenção de /d/* e *lenição de /d/* em gerúndios na região metropolitana do Rio de Janeiro.

Variável dependente	Ocorrências	%
Manutenção de /d/	855	36
Lenição de /d/	1.537	64
Total	2.392	100

Fonte: Elaborada pelos autores.

A regressão logística binária de efeitos mistos foi a abordagem estatística adotada para análise do *corpus*. Entre as variáveis analisadas, aquelas que se mostraram estatisticamente significativas foram *escolaridade* e *contexto fonético-fonológico seguinte*, como efeitos fixos, e *informante* e *frequência de item*, como efeitos aleatórios. As demais variáveis não apresentaram significância estatística.

5.1. Variável escolaridade

A variável *escolaridade* foi a primeira mais significativa selecionada pelo *R-brul*. A tabela abaixo (tabela 3) mostra os resultados obtidos.

[8] *Logodds* se refere a outra medida relevante, a chance, que representa a relação entre a probabilidade de um evento acontecer e a probabilidade de ele não acontecer.

Tabela 3: Resultados da variável escolaridade em relação à lenição de /d/ em gerúndios na região metropolitana do Rio de Janeiro.

Escolaridade	Lenição	Total	%	Logodds	Peso relativo
Ensino Médio	541	657	82	0.768	0.68
Ensino Fundamental	531	786	67	0.660	0.65
Ensino Superior	465	949	49	-1.428	0.19

Fonte: Elaborada pelos autores.

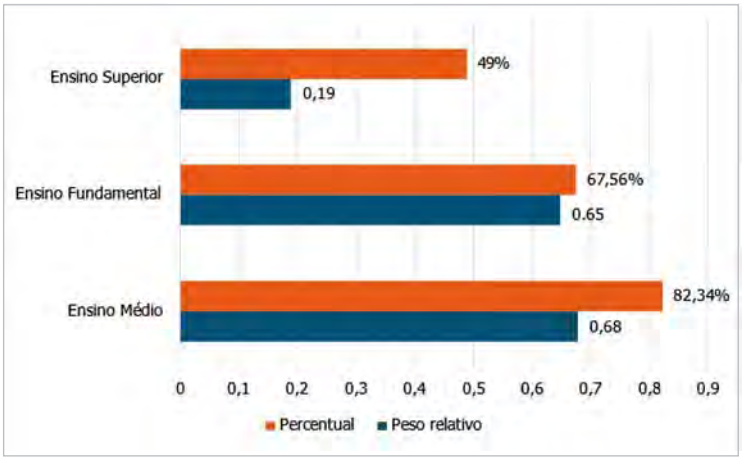
A seguir, observa-se alguns exemplos de lenição de /d/ em gerúndios considerando a *variável escolaridade*. Os dados são, respectivamente, de Ensino Médio, Fundamental e Superior:

(1) a. "... uma hora e meia pra tentar vender a minha ideia pra ele, do meu jeito ou tentan[Ø]o falar do jeito dele, né." (NIG-A-2-M)
b. "... a idade vai chegan[Ø]o a pilha vai ficando fraca." (NIG-B-1-F)
c. "... é lógico que eu não fico chaman[Ø]o a pessoa." (COP-A-3-M)

Com base na tabela 3, a lenição da consoante /d/ em gerúndios é mais frequente entre falantes com escolaridade Média (82%, PR = 0.68) e Fundamental (67%, PR = 0.65). Já os indivíduos com nível Superior tendem a preservar o /d/, como indicam PR inferiores a 0.5 e *logodds* negativos (-1,428). De certa maneira, esses resultados sugerem uma indexação social da variante -no, ou seja, que a aplicação da regra de lenição do /d/ está inversamente relacionada ao nível de escolaridade: quanto maior o grau de escolarização, menor a probabilidade de ocorrência da regra.

Em contraste, Ferreira, Tenani e Gonçalves (2012), em São José do Rio Preto, verificaram maior lenição entre falantes com menor escolaridade (1º ciclo do Fundamental), enquanto os níveis Médio e Superior apresentaram maior preservação do /d/. Na amostra de Nova Iguaçu e Copacabana, a preservação pelos mais escolarizados também é evidente, mas não se observam diferenças significativas entre os níveis Médio e Fundamental (PR 0.68 e 0.65, respectivamente), conforme ilustrado no gráfico abaixo.

Gráfico 1: Peso relativo da redução de /d/ em gerúndios a partir da variável *escolaridade*.



Fonte: Elaborado pelos autores.

Uma interpretação possível para esses resultados, em consonância com as observações de Ferreira, Tenani e Gonçalves (2012), que identificaram uma relação semelhante entre escolaridade e preservação da norma-padrão na variedade linguística do interior paulista, é que os informantes com níveis mais elevados de escolaridade possuem maior familiaridade com a norma culta/comum/ *standard*[9] e reconhecem os valores sociais a ela associados. Isso ocorre porque as formas linguísticas valorizadas socialmente geralmente derivam da gramática normativa e dos grupos sociais de maior prestígio, que, em sua maioria, são constituídos por indivíduos com maior grau de instrução.

O cruzamento entre as variáveis *sexo* e *escolaridade* foi realizado com o objetivo de investigar possíveis relações entre esses fatores. Contudo, os resultados não apresentaram significância estatística para o programa, indicando que não há evidências suficientes para afirmar uma associação relevante entre essas variáveis na amostra analisada. A variável faixa etária também não foi selecionada pelo programa, assim como no estudo de Almeida e Oliveira (2017) na variedade falada em Maceió/AL.

5.2. Variável contexto fonético-fonológico seguinte

A segunda variável selecionada pelo programa foi *contexto fonético-fonológico seguinte*, variável linguística que, tal como em outros trabalhos (*cf.* Oliveira, 2012; Nascimento *et al*., 2013; Almeida; Oliveira, 2017), reforça a hipótese de que o contexto seguinte ocupado por consoantes que compartilham propriedades fonéticas com a consoante oclusiva /d/ favorece a lenição em gerúndios. Oliveira (2012) afirma que o processo de lenição não tem motivação morfológica quando influenciado pela presença de segmentos que partilham propriedades articulatórias, o que o caracteriza como um processo que não é exclusivo ao gerúndio.

[9] Para Faraco (2008), a norma culta/comum/standard é a variedade utilizada pelos letrados em práticas monitoradas de fala e escrita. O autor difere tal conceito da ideia de norma-padrão, uma vez que esta se trata de uma "codificação abstrata, uma baliza extraída do uso real para servir de referência" (p. 73).

Tabela 4: Resultados da variável *contexto fonético-fonológico seguinte* em relação à lenição de /d/ em gerúndios na região metropolitana do Rio de Janeiro.

Contexto seguinte	Lenição	Total	%	Logodds	Peso relativo
/d/	117	158	74	0.747	0.67
/t/	61	77	79	0.661	0.66
/n/	114	153	74	0.640	0.65
Outras consoantes	542	810	66	-0.102	0.47
Vogal	534	879	53	-0.373	0.40
Pausa	169	315	60	-1.574	0.17

Fonte: Elaborada pelos autores.

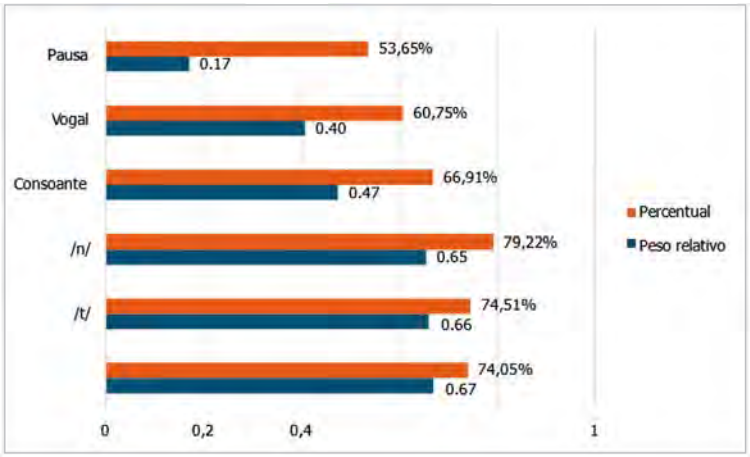
(2) a. "Diz que eu tô *atenden[Ø]o* duas colegas da UFRJ aqui." (NIG-C-3-M)
b. "... isso acaba *prejudican[Ø]o* também a economia." (COP-A-2-M)
c. "... tinha um moleque de rua *dormin[Ø]o* na rua (COP-B-2-F)
d. "... tá *olhan[d]o* pra professora." (NIG-C-1-M)
e. "... ele tá *conhecen[d]o* as coisas, vinte e seis já conhece alguma coisa." (NIG-B-2- M)
f. "... tudo isso tá *fanatizan[Ø]o* um pouco as pessoas de uma forma que não estão *valorizan[d]o_*." (COP-C-1-F)

Os exemplos apresentados em (2) ilustram a variável contexto seguinte, seguindo a mesma ordem da descrição dos resultados: inicialmente, observam-se os contextos com /d/, seguidos por /t/, /n/, outras consoantes, vogais e, por fim, a pausa.

Almeida e Oliveira (2017) observaram que contextos seguidos de /t/, /d/, /n/ apresentam forte favorecimento à lenição da consoante oclusiva /d/ em gerúndios no falar maceioense. Observa-se um comportamento similar nos dados da região metropolitana do Rio de Janeiro: contextos subsequentes preenchidos por consoantes que compartilhem propriedades fonéticas com a consoante /d/ mostram-se fortes favorecedores à redução. Nos resultados, a consoante oclusiva alveolar vozeada /d/ aparece com 74% de lenição (PR 0.67; *logodds* 0.747), seguida da oclusiva alveolar desvozeada /t/, com 79% (PR 0.66; *logodds* 0.661) e da nasal alveolar vozeada /n/, que também apresentou um percentual de lenição de 74% (PR 0.65; *logodds* 0.640).

Ilustramos, com o gráfico abaixo, os percentuais e os pesos relativos de cada elemento considerado em relação à variável contexto fonético-fonológico seguinte.

Gráfico 2: Percentual e peso relativo da redução de /d/ em gerúndios a partir da variável *contexto seguinte*.



Fonte: Elaborado pelos autores.

As consoantes (PR 0.47), vogais (PR 0.40) e pausa (PR 0.17) preenchendo o contexto seguinte favorecem a manutenção de /d/. Destacam-se os resultados encontrados para contextos seguidos por pausa, que, de acordo com Mollica e Mattos (1992), favorecem a manutenção do /d/ porque não há fonema próximo que permita a assimilação[10].

5.3. Variável frequência de item

O programa ainda selecionou a variável *frequência de item*, que, neste estudo, foi definida com base no total de ocorrências (*tokens*) de cada item presente na amostra, bem como no seu respectivo índice de aplicação da regra de lenição.

A análise parece sugerir que a frequência do item lexical exerce alguma influência significativa na aplicação do processo. Isso pode ser verificado pelo fato de os itens mais frequentes – de uma maneira geral – apresentarem tanto os valores de *intercept*[11] quanto os de *peso relativo* mais altos. No entanto, há que se destacar que alguns itens de menor frequência manifestam tais valores também altos, assim como alguns itens de maior frequência revelam esses mesmos valores mais baixos.

Para esta análise, por questões metodológicas, apresenta-se um panorama apenas com as palavras com mais de 20 ocorrências selecionadas, a partir dos resultados do programa.

---

[10] Torna-se importante observar ainda que, no estudo de Mollica e Mattos (1992), foi feita uma oposição entre pausa (silêncio) e qualquer outro ambiente fonológico que não fosse seguido de pausa (não-silêncio). As autoras salientam que o resultado encontrado à época (de PR 0.53 para o segmento seguinte *silêncio*), ainda que tímido, "contraria os postulados neo-gramáticos, uma vez que o fator fonético-fonológico que controla o contexto favorecedor da assimilação só se aplica a uma categoria gramatical" (p. 57), isto é, ao gerúndio.

[11] No modelo de regressão logística, o *intercept* representa a probabilidade de ocorrência da variável dependente quando todas as variáveis previsoras estão em seus níveis de referência.

Quadro 3: Panorama da variável *frequência de item*.

Palavra	Intercept	Frequência	Peso relativo
dando	1.089	43	0.74
trabalhando	1.074	53	0.74
fazendo	1.026	122	0.73
passando	0.885	50	0.70
estudando	0.837	28	0.69
dizendo	0.652	30	0.65
tentando	0.569	23	0.63
falando	0.568	160	0.63
esperando	0.518	21	0.62
olhando	0.454	30	0.61
chegando	0.399	30	0.59
querendo	0.393	47	0.59
andando	0.379	46	0.59
pensando	0.213	30	0.55
correndo	0.14	23	0.53

Fonte: Elaborado pelos autores.

Tendo em vista o panorama da variável *frequência de item*, pode-se dizer que os itens mais frequentes são necessariamente mais suscetíveis à aplicação da regra de lenição de /d/, como mostram os PR da sequência acima, que recorta a extensa lista apresentada pelo programa *R-brul* na seleção da variável. A análise indica que a frequência lexical desempenha um papel significativo no fenômeno, com itens de alta frequência sendo mais suscetíveis à aplicação da regra, o que apoia a hipótese de uma mudança estrutural na língua. Contudo, observa-se que o processo está tão espraiado que também afeta itens de baixa frequência, evidenciado pela diminuição gradual dos pesos relativos.

De acordo com Oliveira (1992), são obscuras as razões que levam um item lexical a ser mais ou menos vulnerável à inovação. Pela perspectiva da Difusão Lexical, o autor aponta certos traços que caracterizam uma palavra afetada pela mudança, tais como [+comum], [-elaborado] e [+contexto fonético natural para a inovação]. A fim de ilustrar o traço [comum], o autor cita como exemplo o processo de monotongação de [ow] para [o], que não tem condicionamento, já que qualquer ditongo [ow] pode ser monotongado [o]: **ou**tro, c**ou**ro, fal**ou**. Por outro lado, a redução não ocorre em palavras com traço [-comum], como Reb**ou**ças, Mosc**ou**, C**ou**to (Oliveira, 1992, p. 37).

Dessa forma, nota-se a importância de integrar diferentes modelos teóricos para uma compreensão mais ampla das mudanças linguísticas. A coexistência das perspectivas da Difusão Lexical e do modelo Neogramático evidencia que as mudanças sonoras não seguem um padrão uniforme, mas são condicionadas por uma complexa interação de fatores fonéticos, lexicais e sociais. Apesar de não destacar as especificidades de cada abordagem, este trabalho reforça a relevância de considerar diferentes perspectivas no estudo linguístico, contribuindo para uma visão mais aprofundada das dinâmicas que regem a língua em uso.

Como etapa futura, pretende-se refinar essa análise, considerando não apenas a frequência bruta dos itens, mas também a frequência específica da lenição em cada um deles, o que possibilitará uma compreensão mais detalhada dos fatores que influenciam o processo variável em estudo.

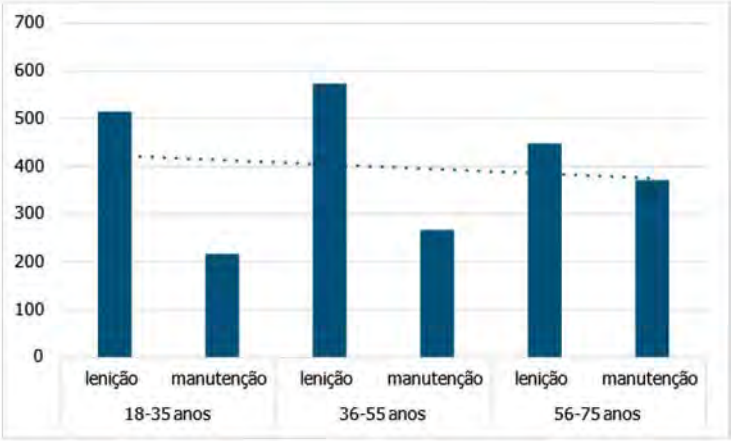
5.4. Variável *informante*

A variável *informante*, incluída como efeito aleatório, apresentou significância limitada, dificultando afirmações robustas a partir dos resultados. Ainda assim, algumas generalizações podem ser feitas. A lenição do /d/ em gerúndios ocorreu com maior frequência entre homens (9/14), indivíduos com baixa escolaridade (10/14) e faixas etárias mais jovens (11/14), sugerindo um possível processo de mudança em progresso (*cf.* gráfico 3). Por ponto de inquérito, a regra foi mais frequente em Nova Iguaçu (8/14) do que em Copacabana (6/14).

Quanto à manutenção da oclusiva /d/, houve predominio entre mulheres (13/22), moradores de Copacabana (12/22) e indivíduos mais velhos (9/22 entre 56-75 anos).

Seguindo os fundamentos da Sociolinguística Variacionista (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972]), a análise da mudança em tempo aparente permite observar o comportamento linguístico de diferentes gerações em uma mesma comunidade durante um período específico. Por sua vez, esse tipo de estudo facilita a identificação de correlações entre a variável social "idade" e a variável linguística em análise, possibilitando evidenciar uma gradação etária, caracterizada por alterações regulares associadas à idade em cada geração ou uma mudança linguística em progresso. No entanto, conforme destacado por Naro (2008), nem toda variação observada na fala reflete necessariamente uma mudança em progresso. Em alguns casos, as diferenças linguísticas entre faixas etárias indicam apenas uma variação estável, sem evidências de transformação contínua no sistema linguístico.

Gráfico 3: Tendência das variantes manutenção de /d/ e lenição de /d/ em relação à *faixa etária*.



Fonte: Elaborado pelos autores.

Nesse mesmo sentido, Tarallo (2001) aponta que considerar a faixa etária dos informantes é importante para identificar se o fenômeno linguístico em análise está estabilizado, quando a idade não interfere em sua ocorrência, ou em processo de mudança, caracterizado por uma maior frequência entre os mais jovens e uma menor aplicação entre os mais velhos.

A tendência apresentada no gráfico foi identificada a partir de nossas próprias conclusões, baseadas nos resultados da variável *informante* selecionada pelo programa. Nesse sentido, ainda que a variável *faixa etária* não tenha sido selecionada, as generalizações buscadas permitiram a utilização do gráfico apenas como suporte explicativo para a tendência destacada anteriormente, a de uma possível mudança em progresso, reforçando a coerência entre os dados e as interpretações aqui propostas.

5.5. Comparando resultados

O fenômeno da lenição da consoante oclusiva /d/ em gerúndios não se restringe à região metropolitana do Rio de Janeiro, sendo, de forma geral, recorrente em diversas variedades do português brasileiro (PB), embora com intensidade variável conforme a localidade e o período.

Estudos realizados ao longo do tempo revelam esse padrão. Mollica e Mattos (1992), no Rio de Janeiro, registraram 62% de manutenção e 38% de lenição em gerúndios. Em 2010, Ferreira constatou forte prevalência da lenição em São José do Rio Preto (72%), resultado semelhante ao observado em Fortaleza por Nascimento *et al.* (2013), com 74,2%. Por outro lado, Almeida e Oliveira (2017), em Maceió, encontraram 42% de lenição e 58% de manutenção, enquanto Freitag, Cardoso e Pinheiro (2018), em Aracaju, registraram 67% de lenição. Recentemente, Ferreira (2023) na Cidade de Goiás apontou 66,9% de lenição.

No Rio de Janeiro, comparações entre diferentes estudos indicam um aumento gradual da lenição ao longo do tempo: Mollica e Mattos (1992) – 38% de lenição; Gonçalves e Melo (2024) – 80,13%.

Este estudo se alinha com os postulados da Sociolinguística Variacionista (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972]), sobretudo os problemas das *restrições* e do *encaixamento*, bem como o conceito de mudança *from below* (abaixo do nível de consciência social), apresentado por Labov (1994).

Ao identificar variáveis sociais e linguísticas que limitam ou favorecem a lenição de /d/ em gerúndios – como o *contexto fonético-fonológico seguinte*, a *frequência lexical*, o *grau de escolaridade* e o *informante* – é possível constatar condições que determinam a variação e, por consequência, as possibilidades de mudança, mostrando quais contextos são mais propícios ao fenômeno e quais oferecem maior resistência.

A análise evidencia, ainda, que a lenição de /d/ não ocorre isoladamente, mas está inserido em um contexto mais amplo de covariação. As variáveis analisadas demonstram como o fenômeno se relaciona com outros aspectos linguísticos (como o contexto fonológico) e sociais (como a escolaridade), reforçando que a mudança está integrada a dinâmicas sociais e linguísticas específicas.

O processo de lenição de /d/ pode ser classificado como uma mudança *from below*, já que o fenômeno ocorre inicialmente no nível vernáculo e parece não estar no radar da consciência social dos falantes. Esta análise permite que, em pesquisa futura, seja realizado complementarmente estudo de percepção, a fim de verificar a hipótese de que a variante [no] é um estereótipo ou um marcador social (Labov, 2008 [1972]).

Considerações finais

Este estudo teve como objetivo investigar os fatores sociais e linguísticos associados à lenição da consoante oclusiva alveolar /d/ no morfema de gerúndio -ndo em dados de fala espontânea coletados em duas áreas distintas da região metropolitana do Rio de Janeiro: Nova Iguaçu e Copacabana. A escolha do fenômeno justifica-se por sua recorrência no português brasileiro (PB) e pelas particularidades estruturais e sociais que o caracterizam, conforme apontado por estudos anteriores.

A análise revelou padrões claros que permitem compreender a atuação de condicionamentos sociais e linguísticos sobre a lenição, bem como sua aplicabilidade no dialeto carioca. Os resultados indicam que o fenômeno ocorre de forma significativa, com 64% de lenição e 36% de manutenção do segmento /d/, evidenciando que se trata de uma mudança em curso no português brasileiro (PB). Além disso, foi possível observar a influência de variáveis sociais e linguísticas selecionadas pelo programa estatístico – escolaridade, contexto fonético-fonológico seguinte, frequência do item e informante – na ocorrência da lenição em formas gerundiais.

A escolaridade mostrou-se um fator decisivo: a lenição do /d/ é mais frequente entre falantes com níveis Médio e Fundamental, enquanto indivíduos com escolaridade superior tendem a preservar o segmento. Os resultados encontrados indicam a indexação social da variante [-no], com redução da lenição à medida que aumenta o grau de escolarização, padrão que se aproxima do observado em

outras variedades do português brasileiro (PB), como a de São José do Rio Preto, ainda que com diferenças mais marcadas entre os níveis Médio e Fundamental no Rio de Janeiro.

Quanto ao contexto fonético-fonológico seguinte, os resultados mostram que consoantes articulatoriamente próximas a /d/ (/t/, /d/, /n/) favorecem significativamente a lenição, enquanto vogais, consoantes distantes e pausas tendem a preservá-la. Esses achados confirmam que o processo não depende exclusivamente de motivação morfológica, mas está relacionado a características articulatórias e de assimilação fonética, alinhando-se a estudos anteriores (Mollica; Mattos, 1992; Oliveira, 2012; Nascimento *et al*., 2013; Almeida; Oliveira, 2017).

As variáveis *frequência do item* e *informante*, selecionadas como efeitos aleatórios, também forneceram *insights* importantes. Itens de alta frequência mostraram maior suscetibilidade à lenição, sugerindo influência estrutural no sistema linguístico. Quanto à variável *informante*, homens mais jovens, de baixa escolaridade e residentes de Nova Iguaçu apresentaram maior propensão à aplicação da regra, enquanto mulheres, indivíduos mais escolarizados, de faixas etárias mais altas e residentes de Copacabana preservaram o /d/, indicando adesão à norma-padrão e resistência à mudança.

Embora a faixa etária não tenha sido selecionada como significativa, a análise permitiu observar tendências intergeracionais: maior ocorrência da lenição entre os jovens sugere um processo de mudança em progresso, alinhado ao conceito de mudança *from below* (Labov, 1994; Weinreich; Labov; Herzog, 2006 [1968]). A não significância da idade pode estar relacionada à sobreposição com a variável informante, que incorpora múltiplas características sociais. Futuras análises podem minimizar essas interferências, realizando rodadas separadas para avaliar individualmente faixa etária, escolaridade, sexo e frequência lexical.

Outros desdobramentos do estudo incluem a análise do impacto do contínuo rural-urbano sobre a realização das variantes, especialmente em Nova Iguaçu, município com características urbanas e rurais, permitindo avaliar influências contextuais adicionais sobre a variação linguística, conforme a perspectiva de Bortoni-Ricardo (2004).

Em síntese, este trabalho contribui para a literatura sociolinguística ao explorar a lenição de /d/ em gerúndios na variedade do Rio de Janeiro, considerando duas áreas socio-geograficamente contrastantes. A abordagem quali-quantitativa adotada reafirma a relevância da Sociolinguística Variacionista como ferramenta para compreender dinâmicas de variação e mudança linguística, evidenciando a relação intrínseca entre língua e sociedade e fornecendo subsídios para futuros estudos sobre padrões e tendências sociolinguísticas na região metropolitana do Rio de Janeiro.

Referências

ALMEIDA, A. N. S.; OLIVEIRA, A. J. Você fala cantano? Uma análise do apagamento de /d/ em gerúndios no falar de Maceió/AL. *Letrônica*. Porto Alegre, v. 10, n. 1, pp. 200-209, jan- jun. 2017.

ARAÚJO, A. A.; ARAGÃO, M. S. S.; ALMEIDA, B. K. M. A redução do gerúndio no atlas linguístico do Pará: uma abordagem variacionista. *Caderno Seminal*, Rio de Janeiro, v. 26, n. 26, 2016. Disponível em: https://www.e-publicacoes.uerj.br/cadernoseminal/article/view/21529. Acesso em: 9 dez. 2024.

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

CALLOU, D.; LEITE, Y. *Iniciação **à** fonética e **à** fonologia*. 5. ed. Rio de Janeiro: Jorge Zahar, 1990.

CÂMARA JR, J. M. *Estrutura da língua portuguesa*. 47. ed. Petrópolis: Editora Vozes, 2015 [1970].

COSTA, G. B. Reflexos pedagógicos da simplificação do gerúndio em estudantes do ensino fundamental. *Letra Magna* (*Online*), v. 11, pp. 1-22, 2009. Disponível em: http://www.uesc.br/eventos/selipeanais/anais/geisaborges.pdf.

CRISTÓFARO-SILVA, T. Fonologia: por uma análise integrada à morfologia e à sintaxe. *In*: SEMANA DE ESTUDOS PORTUGUESES, 2., 1996, Belo Horizonte. *Anais [...].* Belo Horizonte: FALE/UFMG, 1996. v. 2, p. 56-65. Disponível em: Disponível em: http://www.periodicos.letras.ufmg.br/index.php/anais_lingua_portuguesa/article/view/8054. Acesso em: 15 out. 2024.

FERREIRA, J. S. S. R. *O apagamento da oclusiva dental /d/ em gerúndio*: uma análise sociolinguística do falar vilaboense. 2023. 107 f. Dissertação (Mestrado em Língua, Literatura e Interculturalidade) – Câmpus Cora Coralina, Universidade Estadual de Goiás, Goiás, GO, 2023.

FERREIRA, J. S.; TENANI, L.; GONÇALVES, S. C. L. O morfema de gerúndio "ndo" no português brasileiro: análise fonológica e sociolinguística. *Revista Letras e Letras*, v. 28, n. 1, pp. 167-188, jan-jun. 2012.

FERREIRA, J. S. *O apagamento do /d/ em morfema de gerúndio no dialeto de São José do Rio Preto*. 2010. Dissertação (Mestrado) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas, 2010.

FREITAG, R. M. K; CARDOSO, P. B.; PINHEIRO, B. F. M. Saliência na conservação de /d/ no segmento /ndo/: efeitos sociais e estilísticos. *In*: *Gragoatá*, Niterói, v. 23, n. 46, pp. 654-678, maio-ago. 2018.

GONÇALVES, M. B.; MELO, M. A. S. L. A alternância entre [d] e [n] em contexto -ndo na comunidade de fala do Rio de Janeiro. *In*: SEPLA (Seminário de Pesquisas Linguísticas em Andamento) do Programa de Pós-Graduação em Linguística da UFRJ, 29, 2024, Rio de Janeiro. *Anais* [...]. Rio de Janeiro: PPGLIN/UFRJ, 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Censo demográfico 2010*: resultados gerais da amostra. Rio de Janeiro: IBGE, 2012.

LABOV, W. *Principles of linguistic change*: internal factors. Oxford: Blackwell, 1994.

LABOV, W. *Padrões Sociolinguísticos*. São Paulo: Parábola, 2008 [1972].

MARTINS, I. F. M. Apagamento da oclusiva dental /d/: perspectivas variacionista e fonológica. *In*: HORA, Demerval da (org.). *Estudos sociolinguísticos*: Perfil de uma comunidade. João Pessoa: Editora Universitária da UFPB, 2004. pp. 55-82.

MELO, M. A. S. L.; SILVA, I. G. B. [avaliˈãdʊ] ou [avaliˈãnʊ]:o significado social da alternância [d] ~ [n] entre jovens universitários cariocas. *Revista (Con)Textos Linguísticos*. Vitória, v. 16, n. 34, pp. 185-204, 2022. Disponível em: https://periodicos.ufes.br/contextoslinguisticos/article/view/38574. Acesso em: 20 nov. 2024.

MOLLICA, M. C.; MATTOS, P. Pela conjugação das abordagens variacionista e difusicionista. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 1, n. 1, pp. 53-63, 1992.

NASCIMENTO, K. R. S.; ARAÚJO, A. A.; CARVALHO, W. J. A. A redução do gerúndio no falar popular de Fortaleza: um olhar variacionista. *Veredas*. Juiz de Fora, v. 2, pp. 398-413. 2013. Disponível em: https://periodicos.ufjf.br/index.php/veredas/article/view/25011. Acesso em: 5 dez. 2024.

NASCIMENTO, L.; MOTA, J. A ausência do 'd' no gerúndio: com base em inquéritos experimentais do projeto ALIB. *Hyperion Letras*, Salvador, n. 7, [s. p.], 2004. Disponível em: http://www.hyperion.ufba.br/revista_7_04.htm. Acesso em: 15 out. 2024.

R CORE TEAM. *R*: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2022. Disponível em: https://www.R-project.org/. Acesso em: 28 out. 2024.

SOUSA, S. C. T. Interferência da língua falada na escrita de crianças: processos de apagamento da oclusiva dental /d/ e da vibrante final /r/. *D.E.L.T.A*, v. 25, n. 2, pp. 465-495, 2009.

TARALLO, F. *A pesquisa sociolinguística*. 7. ed. São Paulo: Martins Fontes, 2001.

VIEIRA, S. R.; BRANDÃO, S. F. *CORPORAPORT*: Variedades do Português em análise. Rio de Janeiro: Faculdade de Letras-UFRJ, 2018. Disponível em: www.corporaport.letras.ufrj.br. Acesso em: 15 jul. 2024.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola Editorial, 2006 [1968].