Working Papers em Linguística

HISTÓRIA DO PORTUGUÊS BRASILEIRO: UMA HOMENAGEM AO PROFESSOR ATALIBA DE CASTILHO



2025

v. 26

Working Papers em Linguística, v. 26, 2025

Centro de Comunicação e Expressão - CCE
Universidade Federal de Santa Catarina - UFSC
Florianópolis - SC - Brasil

Editor-chefe
Marco Antonio Rocha Martins

Organização
Izete Lehmkuhl Coelho
Marco Antonio Rocha Martins
Paulo Osório

Design e arquivamento
João Paulo Zarelli Rocha

Equipe de revisão
Ana Beatriz Ribeiro
Andriéle Cristina Stasiak
Helena Gouveia
Íris Medeiros da Fonceca
Jonathan Murilo Souza dos Santos
Suzane Cardoso Gonçalves Madruga
Vanessa Grando

Conselho Editorial
Adair Bonini, Universidade Federal de Santa Catarina, Brasil
Adja Balbino de Amorim Barbieri Durão, Universidade Federal de Santa Catarina, Brasil
Adriana Fischer, Centro Universitário de Brusque, Brasil
Aline Cacilda Koteski Emilio, Universidade Estadual de Ponta Grossa, Brasil
Ana Cláudia Souza, Universidade Federal de Santa Catarina, Brasil
Ana Paula Oliveira Santana, Universidade Federal de Santa Catarina, Brasil
André Berri, Universidade Federal de Santa Catarina, Brasil
Clarice Nadir von Borstel, Universidade Estadual do Oeste do Paraná, Brasil
Cláudia Regina Brescancini, Pontifícia Universidade Católica – RS, Brasil
Cristiane Lazzarotto-Volcão, Universidade Federal de Santa Catarina, Brasil
Cristine Gorski Severo, Universidade Federal de Santa Catarina, Brasil
Edair Maria Gorski, Universidade Federal de Santa Catarina, Brasil
Edwiges Maria Morato, Universidade Estadual de Campinas, Brasil
Fabio Luiz Lopes da Silva, Universidade Federal de Santa Catarina, Brasil
Felício Wessling Margotti, Universidade Federal de Santa Catarina, Brasil
Helena Guerra Vicente, Universidade de Brasília, Brasil
Heronides Maurílio de Melo Moura, Universidade Federal de Santa Catarina, Brasil
Izabel Christine Seara, Universidade Federal de Santa Catarina, Brasil
Josias Ricardo Hack, Universidade Federal de Santa Catarina, Brasil
Leandra Cristina de Oliveira, Universidade Federal de Santa Catarina, Brasil
Leonor Scliar Cabral, Universidade Federal de Santa Catarina, Brasil
Lucélio Dantas Aquino, Universidade Federal do Rio Grande do Norte, Brasil
Luizete Guimarães Barros, Universidade Federal de Santa Catarina, Brasil
Magdiel Medeiros Aragão Neto, Universidade Federal da Paraíba, Brasil
Mailce Borges Mota, Universidade Federal de Santa Catarina, Brasil
Maria Inêz Probst Lucena, Universidade Federal de Santa Catarina, Brasil
Maria Izabel de Bortoli Hentz, Universidade Federal de Santa Catarina, Brasil
Maria Teresa Santos Cunha, Universidade do Estado de Santa Catarina
Márluce Coan, Universidade Federal do Ceará, Brasil
Mary Elizabeth Cerutti-Rizzatti, Universidade Federal de Santa Catarina, Brasil
Maurício Eugênio Maliska, Universidade do Sul de Santa Catarina, Brasil
Monica Mano Trindade, Universidade Federal da Paraíba, Brasil
Morgana Fabiola Cambrussi, Universidade Federal da Fronteira Sul, Brasil
Nara Caetano Rodrigues, Universidade Federal de Santa Catarina, Brasil
Nelita Bortolotto, Universidade Federal de Santa Catarina, Brasil
Nívea Rohling, Universidade Tecnológica Federal do Paraná, Brasil
Otávio Goes de Andrade, Universidade Estadual de Londrina, Brasil
Renato Basso, Universidade Federal de São Carlos, Brasil
Roberta Pires de Oliveira, Universidade Federal de Santa Catarina, Brasil
Rodrigo Acosta Pereira, Universidade Federal de Santa Catarina, Brasil
Ronald Taveira da Cruz, Universidade Federal do Parnaíba Piaui, Brasil
Rosângela Hammes Rodrigues, Universidade Federal de Santa Catarina, Brasil
Rosely Xavier, Universidade Federal de Santa Catarina, Brasil
Simone Bueno Borges da Silva, Universidade Federal da Bahia, Brasil
Tarcisio de Arantes Leite, Universidade Federal de Santa Catarina, Brasil
Terezinha da Conceição Costa-Hübes, Universidade Estadual do Oeste do Paraná, Brasil
Vidomar Silva Filho, Instituto Federal de Santa Catarina, Brasil
Werner Heidermann, Universidade Federal de Santa Catarina, Brasil

HISTÓRIA DO PORTUGUÊS BRASILEIRO:
UMA HOMENAGEM AO PROFESSOR
ATALIBA DE CASTILHO

Sumário

HISTÓRIA DO PORTUGUÊS BRASILEIRO E UMA BREVE ENTREVISTA COM ATALIBA TEIXEIRA DE CASTILHO 4
Izete Lehmkuhl Coelho, Marco Antonio Rocha Martins, Paulo Osório

Artigos

AS ORIGENS DA GRAMÁTICA DO PORTUGUÊS: UMA ABORDAGEM HISTORIOGRÁFICO-CRÍTICA 10
Carlos Alberto Faraco

ECONOMIA, SOCIEDADE E FORMAÇÃO DO PORTUGUÊS BRASILEIRO: ELEMENTOS PARA UMA DISCUSSÃO SOBRE AS CONFIGURAÇÕES SOCIOLINGUÍSTICAS DO BRASIL COLONIAL 36
Ana Sartori, Tânia Lobo, Pedro Daniel dos Santos Souza

O PORTUGUÊS NO BRASIL COLONIAL: REFLEXÕES SOBRE EDIÇÕES DE MANUSCRITOS HISTÓRICOS BRASILEIROS 74
Mariana Fagundes de Oliveira Lacerda, Alícia Duhá Lose, Zenaide de Oliveira Novais Carneiro

ANÁLISE DO STATUS FONOLÓGICO DAS CONSOANTES RÓTICAS, LATERAIS E NASAIS DUPLAS DO PORTUGUÊS ARCAICO 106
Débora Aparecida dos Reis Justo Barreto, Gladis Massini-Cagliari

DO PORTUGUÊS CLÁSSICO AO PORTUGUÊS BRASILEIRO: HERANÇA, RAÍZES, DERIVA E CONTATO 132
Charlotte Galves

A CONSTITUIÇÃO DE 1946 E A IDENTIFICAÇÃO DA LÍNGUA NACIONAL 152
Maria Carlota Rosa

HISTÓRIA SOCIAL, CONTATO E A AQUISIÇÃO DO PORTUGUÊS BRASILEIRO COMO L2 171
Maria Eugenia Lammoglia

O SUJEITO NULO EM CARTAS BRASILEIRAS DOS SÉCULOS XIX E XX: NOVOS DADOS PARA UMA VELHA MUDANÇA 204
Silvia Regina de Oliveira Cavalcante

TRAJETÓRIA DE MUDANÇA DO SUJEITO E DO OBJETO DIRETO ANAFÓRICOS EM SANTA CATARINA: ANÁLISE DE CARTAS PESSOAIS DOS SÉCULOS XIX E XX 232
Cecília Augusta Vieira Pinto

A ORDEM DO VERBO NA HISTÓRIA DO PB: QUESTÕES PARA PERIODIZAÇÃO 257
Paulo Ângelo Araújo-Adriano, Williane Corôa

A EXPANSÃO DA PREPOSIÇÃO PARA NAS DITRANSITIVAS DO PORTUGUÊS BRASILEIRO: A PERSPECTIVA DO MODELO DE TRÊS FATORES 302
Ana Calindro, Maria Aparecida Torres Morais

A TOPICALIZAÇÃO DE OBLÍQUOS NUCLEARES EM PEÇAS TEATRAIS BRASILEIRAS DOS SÉCULOS XIX E XX: EVIDÊNCIAS PARA A HIPÓTESE DE COMPETIÇÃO DE GRAMÁTICAS 328
Mônica Tavares Orsini, Marco Antonio Rocha Martins

"REDES SOCIAIS E SOCIOLINGUÍSTICA HISTÓRICA": VARIAÇÃO DOS PRONOMES DE SEGUNDA PESSOA EM CARTAS (1956-1994) 356
Célia Lopes, Janine da Silva Barbosa, João Pedro de C. G. da Silva, Mariana da Silva Soares

A EXPRESSÃO PRONOMINAL DE POSSESSIVOS DE SEGUNDA PESSOA EM CARTAS DE AMOR DO SÉCULO XX 377
Cristiane Namiuti, Carlos Wilson de Jesus Pedreira, Isabel dos Santos Magalhães Gomes, Raiana Cristina Dias da Cruz, Mireia Lêda Silva Santos, Gabriela Santos Barros, Daniela Soares de Oliveira

SOBRE A VARIAÇÃO DOS POSSESSIVOS TEU/SEU NOS SÉCULOS XIX E XX: ANÁLISE EM CARTAS PESSOAIS 390
Dinah Callou, Márcia Cristina de Brito Rumeu

AS FORMAS "NÓS" E "A GENTE" NAS CARTAS PESSOAIS DA AMOSTRA FAMÍLIA ARTHUR REIS 411
Flávia Santos Martins, Edson Galvão Maia, José Fabrício A. F. Júnior, Romário Neves Coelho

MANUTENÇÃO LINGUÍSTICA E A VARIÁVEL 'GRAU DE RURALIDADE' EM FOCO: UM INSTRUMENTO METODOLÓGICO PARA ANÁLISE DE FENÔMENOS LINGUÍSTICOS RURAIS 439
Ivelã Pereira

A (RE)EXISTÊNCIA DOS POVOS INDÍGENAS NAS ENTRANHAS DO SERTÕES BAIANOS: UMA BREVE HISTÓRIA SOCIAL LINGUÍSTICA 466
Paloma Reis Soares, Edson Tosta Matarezio Filho

TRADUÇÃO

CONSTRUÇÕES-WH NO PORTUGUÊS BRASILEIRO: "GÊMEOS" (DOUBLETS) SINTÁTICOS OU VARIAÇÃO ESTILÍSTICA 494
Mary Aizawa Kato

http://doi.org/10.5007/1984-8420.2025.e108131

HISTÓRIA DO PORTUGUÊS BRASILEIRO E UMA BREVE ENTREVISTA COM ATALIBA TEIXEIRA DE CASTILHO

Izete Lehmkuhl Coelho | Lattes | izete.lehmkuhl.coelho@ufsc.br
Universidade Federal de Santa Catarina

Marco Antonio Rocha Martins | Lattes | marcomartins.ufsc@gmail.com
Universidade Federal de Santa Catarina

Paulo Osório | Lattes | Paulo.Osorio@uab.pt
Universidade Aberta

1 Introdução

Este volume 26/nº 1 publicado pela *Working Papers em Linguística* do Programa de Pós-Graduação em Linguística da Universidade Federal de Santa Catarina (UFSC), que reúne 19 artigos voltados a aspectos da história social e gramatical do português no Brasil escritos por especialistas na área, é uma singela homenagem ao professor Ataliba Teixeira de Castilho. É uma singela homenagem porque a relevância e o trabalho do professor Ataliba, especialmente no âmbito do projeto *Para a história do Português Brasileiro (PHPB)* e da área de Linguística Histórica no Brasil não se expressam em palavras, mas é sempre importante tentar! É uma homenagem do projeto *História do Português Brasileiro – da Europa à América* (o Projeto 3!) da Associação de Linguística e Filologia da América Latina (ALFAL), sob a atual coordenação de Izete Lehmkuhl Coelho, Marco Antonio Rocha Martins e Paulo Osório, desde 2012; uma homenagem representativa da própria ALFAL para o qual foi o professor Ataliba um importante representante no Brasil, tendo inclusive presidido a Associação. Muitos dos autores que publicam artigos neste volume ao homenageado foram colegas de trabalho do professor Ataliba no PHPB; muitos foram alunos; todos foram fortemente influenciados por suas ideias, publicações ou participação em projetos coletivos.

Muito recentemente, Martins, Coelho e Osório (2025) publicaram a importante coletânea *História do Português Brasileiro: trajetórias e perspectivas* vinculada ao Projeto 3 da ALFAL em parceira com a editora De Gruyter/Alemanha e a sua coleção Linguística Latino-americana. A coletânea reúne capítulos dos mais renomados especialistas brasileiros que atuam no PHPB, projeto de grande relevância e proporções para a Linguística

Histórica no Brasil. O volume versa sobre a história gramatical e sociocultural do português no Brasil no curso dos séculos, contemplando tópicos de mudanças gramaticais, nos níveis fonético-fonológicos, morfossintáticos, lexicais e textuais; de história social, dialetação e de formação do português brasileiro em solos do continente Americano. Desta forma a obra também documenta a trajetória de pesquisa sobre o português brasileiro refletida nos 11 volumes da coleção *História do Português Brasileiro*, organizada pelo professor Ataliba, que, nas palavras dos autores nos permite "vislumbrar trajetórias traçadas pelos estudos da história social e gramatical do português no e do Brasil, no âmbito desses dois importantes projetos", o projeto 3 da ALFAL e o projeto PHPB."

2 Vida-obra de Ataliba Teixeira de Castilho

Ataliba Teixeira de Castilho foi Professor Titular da Universidade Estadual Paulista, campus de Marília (1961-1975), então CESESP, Professor visitante da Universidade do Texas em Austin (1970), Professor Titular do Departamento de Linguística do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas (1975-1991), cargo em que se aposentou em 1991, Professor Titular de Filologia e Língua Portuguesa na FFLCH da Universidade de São Paulo (1996 a 2006) e Professor Emérito da FFLCH/USP em 2013. Atualmente é professor colaborador voluntário na Universidade Estadual de Campinas. Fez Curso de Graduação em Licenciatura em Letras Clássicas pela Faculdade de Filosofia Ciências e Letras da Universidade de São Paulo, em 1959, com Especialização em Filologia Românica, em 1960, Doutorado em Linguística em 1966, e Livre-docência em Filologia e Linguística Portuguesa em 1993, pela mesma universidade.

No plano institucional, Castilho fundou o Grupo de Estudos Linguísticos do Estado de São Paulo (GEL) em 1969, do qual foi Presidente, propôs a fundação da Associação Brasileira de Linguística (ABRALIN), que viria a presidir, e atuou na Associação de Linguística e Filologia da América Latina (ALFAL), como Vogal, e posteriormente como Presidente, criando ali diversos grupos de pesquisa, objetivando aproximar brasileiros e hispano-americanos no estudo da identidade linguística do continente. Fundou a Revista Alfa, hoje Revista de Linguística da Universidade Estadual Paulista, e colaborou para a fundação da Revista de Estudos Linguísticos da Unicamp e da Revista de Filologia e Linguística da USP. Teve, durante décadas, participação ativa nos encontros do GEL, da ABRALIN e da ALFAL, por meio de conferências e de mesas-redondas. Foi pesquisador do CNPq durante grande parte de sua carreira, participando de importantes projetos de

pesquisa individuais e coletivos. Desde 2006, atua como assessor linguístico do Museu da Língua Portuguesa.

Ataliba de Castilho sempre se empenhou em criar e oferecer condições para formação altamente qualificada de recursos humanos: orientou em torno de 30 estudantes entre Iniciação Científica, Mestrado e Doutorado. Muitos de seus orientandos atuaram (ou ainda atuam) em diferentes universidades brasileiras. Sempre acreditou no trabalho coletivo, coordenando com sua memorável capacidade de gerenciamento de grandes projetos de pesquisa. Foi coordenador geral do Projeto *Norma Urbana Oral Culta de São Paulo* (NURC/SP), do Projeto de *Gramática do Português Falado*, do Projeto de *História do Português de São Paulo* e do Projeto *Para a História do Português Brasileiro*. Esses grandes projetos cumpriram as importantes tarefas de constituição de *corpora* de diferentes variedades brasileiras e de descrição e análise de dados publicados em obras de referência.

Dessa trajetória, Ataliba reúne mais de uma centena de produções bibliográficas, incluindo artigos, livros, organização de livros, capítulos de livros, publicados no Brasil e no exterior. Suas pesquisas têm-se concentrado na descrição e história do Português Brasileiro, no ensino do Português como língua materna e em política linguística, com ênfase nos seguintes temas: descrição da língua falada, sintaxe funcionalista do português brasileiro, história do português brasileiro e análise multissistêmica do português brasileiro. Queremos destacar nessa brilhante trajetória a importante obra *Nova Gramática do Português Brasileiro*, publicada em 2010 pela Editora Contexto, e a edição geral da obra coletiva *História do Português Brasileiro*, com 11 volumes publicados também pela Contexto entre 2018 e 2022. Essa obra coletiva, segundo Castilho (2024), "reúne todos os sistemas que compõem uma língua natural: Gramática, Léxico, Discurso e Semântica. Um empreendimento e tanto!"

A seguir, passamos a palavra ao professor Ataliba de Castilho que muito gentilmente responde algumas questões que colocamos em pauta no sentido de complementar essa breve homenagem pela sua brilhante e marcante colaboração com a linguística (e a Linguística Histórica, mais especificamente) no Brasil. A linguística e o modo de articular pesquisas em grandes projetos não são os mesmos depois de Ataliba!

3 Entrevista

Coelho, Martins e Osório: Queremos iniciar esta nossa conversa falando sobre o Projeto 3: *História do português brasileiro – desde a Europa até a América* que atualmente temos a

honra de coordenar. Gostaríamos de saber como foi a sua experiência como o primeiro coordenador do Projeto 3 da ALFAL.

Castilho: Quando me candidatei a Presidente da ALFAL, argumentei que a associação era muito operante durante os congressos, mas que se apagava após eles, como se fosse um vagalume. Para mantê-la funcionando continuamente, vários projetos foram então lançados, entre eles Projeto 3*: História do português brasileiro – desde a Europa até a América*. Com isto, a ALFAL deixou de ser "una luciérnaga". Minha experiência como primeiro coordenador do Projeto 3 confirmou que as pesquisas coletivas se constituem no melhor modo de fazer ciência entre nós.

Coelho, Martins e Osório: Os estudos em Linguística Histórica no Brasil, como você já mencionou em diversos textos, renasceram na década de 1980 com Fernando Tarallo e Mary Kato na UNICAMP, com um grande projeto sobre a história do português brasileiro, e com Rosa Virgínia Mattos e Silva na UFBA, com seu estudo sobre o português trecentista, a que se seguiu a criação do projeto *Programa de História do Português* (PROHPOR). Em que esses estudos contribuíram para a constituição do projeto coletivo *Para a História do Português Brasileiro* (PHPB), sob sua coordenação geral até 2019?

Castilho: Os projetos mencionados comprovaram que também a Linguística Histórica poderia beneficiar-se desta interessante característica da Linguística brasileira: tratar os grandes temas numa forma coletiva, sem prejuízo do necessário trabalho individual. Assim, ao lançar o PHPB, eu tinha certeza de que tudo daria certo.

Coelho, Martins e Osório: O *Projeto 3: História do português brasileiro – desde a Europa até a América* em seus encontros anuais contou inúmeras vezes com a agenda do PHPB, relacionada à discussão a respeito da constituição do corpus e da descrição de resultados preliminares dos grupos de pesquisa. Como eram esses encontros com o PHPB no âmbito da ALFAL?

Castilho: Os seminários nacionais do PHPB sempre foram muito produtivos, graças ao empenho dos pesquisadores que se associaram a esse projeto. Particularmente no caso do Projeto 3 da ALFAL, ele poderá ampliar seu escopo se desenvolver atividades conjuntas com o Projeto para a História do Espanhol, já existente, e do Francês da América, a ser

criado. Se formos por aqui, novas áreas de pesquisa poderão ser desenvolvidas, tais como (1) Diacronia da România Nova, (2) Teoria da mudança linguística. E a ALFAL precisará desenvolver esforços para recrutar associados na Guiana Francesa.

Coelho, Martins e Osório: Uma das tarefas do PHPB foi a de levantar e organizar, a partir de critérios definidos nos seminários nacionais, um banco de dados linguístico formado por um *corpus mínimo comum* e um *corpus diferencial* dos séculos XVIII, XIX e XX, que subsidiasse análises diversas sobre a história linguística e social do português brasileiro realizadas pelos grupos de pesquisadores nas diferentes áreas de atuação. Como foi conduzida a constituição desse banco de dados?

Castilho: Bancos de dados demandam muitos esforços e, sobretudo, precisam ser organizados com a ajuda de dispositivos computacionais. Charlotte Galves, que hoje está integrada na equipe paulista do PHPB, desenvolveu importantes programas de automatização de corpus, graças ao seu projeto de História do Português Europeu. Sem bancos de dados, os projetos coletivos de pesquisa linguística não podem prosperar.

Coelho, Martins e Osório: Desde o I Seminário do PHPB, que aconteceu na USP em 1997, o projeto foi organizado em equipes regionais e foi se ampliando ao longo do tempo, contando hoje com filiados de várias universidades brasileiras. Além de sua coordenação exemplar, a que fatores se deve o sucesso desse projeto coletivo que já está completando quase 30 anos e continua a percorrer sua tarefa de documentar o português brasileiro?

Castilho: São muitos os fatores que vêm beneficiando o PHPB: a habilidade dos linguistas brasileiros em atuar coletivamente no desenvolvimento de projetos científicos, a facilidade com que eles identificam e desenvolvem projetos de importância para o tratamento de nossa identidade linguística, e sua receptividade às novas orientações científicas, entre outros.

Coelho, Martins e Osório: Desde o início de sua constituição, o projeto PHPB tinha decidido consolidar os resultados obtidos de descrição do português brasileiro nas diferentes áreas de atuação numa coleção própria. Essa atividade teve êxito com a publicação da coleção *História do Português Brasileiro*, composta por 11 volumes publicados pela Contexto entre 2018 e 2022 e um volume em desenvolvimento a sair em edição

eletrônica. Como foi para você a experiência de coordenar essa obra tão importante sobre a história linguística e social do português brasileiro?

Castilho: Foram muitas as boas surpresas que tive ao coordenar o PHPB, e a publicação da coleção *História do Português Brasileiro* foi a melhor dentre elas. Os colegas atuaram com determinação, cumpriram os prazos, e – segundo penso – preparam-se agora para teorizar sobre mudança linguística, com base nos vastos materiais publicados nessa série. A postulação de teorias linguísticas é sem dúvida o próximo passo que a Linguística brasileira vai dar, emparelhando-se ao que de melhor se faz no mundo, atualmente, nesse setor.

Coelho, Martins e Osório: Numa apreciação geral, qual a sua opinião sobre o desenvolvimento desses dois importantes projetos voltados à história social e gramatical do português da Europa à América, e especialmente para o estudo do português no Brasil e a formação intelectual de pesquisadores em universidades brasileiras? Quais foram as principais contribuições desses projetos?

Castilho: Como disse no quesito anterior, para além de descrever e historiar o Português Brasileiro, a produção de teorias é, certamente, o próximo passo da Linguística Brasileira. Em meu discurso de recepção do título de Professor Emérito pela Universidade de São Paulo, lancei a ideia de criarmos um Programa Interinstitucional de Pós-Doutoramento em Teorização Linguística. Enviei o projeto respectivo ao Prof. Cleber Ataíde, atual Presidente da Associação Brasileira de Linguística. Ele integra a equipe pernambucana do PHPB. Logo, vai dar certo.

Coelho, Martins e Osório: Nosso muito obrigado!

Referências

CASTILHO, Ataliba Teixeira de. *Historiando a Língua Portuguesa no Novo Mundo.* Texto enviado para Conferência de encerramento do Projeto 3 do XX Congresso Internacional da ALFAL, Concepción, Chile, 25 de janeiro de 2024.

MARTINS, Marco Antonio Rocha; COELHO, Izete Lehmkuhl; OSORIO, Paulo (orgs.). *História do Português Brasileiro*: trajetórias e perspectivas. 1. ed. Berlin/Boston: Walter de Gruyter GmbH, 2025.

http://doi.org/10.5007/1984-8420.2025.e108134

AS ORIGENS DA GRAMÁTICA DO PORTUGUÊS: UMA ABORDAGEM HISTORIOGRÁFICO-CRÍTICA

THE ORIGINS OF PORTUGUESE GRAMMAR: A HISTORIOGRAPHICAL AND CRITICAL APPROACH

Carlos Alberto Faraco | Lattes | carlosfaraco62@gmail.com
Universidade Federal do Paraná

Resumo: Este texto é uma reflexão crítica sobre as origens da gramática do português, em especial sobre a crença de que a gramática tradicional padroniza os usos, tendo como referência os textos literários. Como se buscará mostrar aqui, é uma crença que, em geral, não corresponde à realidade, na medida em que, nas obras gramaticais tradicionais, predominam (quando não são únicos) os dados criados pelo próprio autor. Não há, em nenhuma delas, um *corpus* literário do qual se parte para as suas asserções e regras. No trato crítico dessa questão, é importante ter claro que há, aí, subjacentes duas concepções distintas de gramática, saídas de vertentes diversas dos estudos das línguas. De um lado, está a gramática vinda da tradição greco-latina (e, por isso, qualificada de *tradicional*), que, em tese, funda seus preceitos nos textos literários. De outro, estão os modelos gramaticais elaborados por diferentes teorias no interior da linguística, cujo escopo empírico alcança todas as manifestações da língua, quaisquer que sejam. Depois de um breve histórico da constituição da gramática tradicional, analisa-se, especificamente, as origens da gramática do português, mostrando que ela se fez, efetivamente, pela transposição da doutrina consolidada na tradição gramatical latina e pela acomodação nela da língua vernácula. A fonte da normatividade não foi propriamente o uso como tal, mas o que já estava dito nas gramáticas latinas.

Palavras-chave: Gramática. Português. Historiografia Crítica.

Abstract: This text is a critical reflection on the origins of Portuguese grammar, in particular on the belief that traditional grammar standardizes usage, using literary texts as a reference. As we will try to show here, it is a belief that, in general, does not correspond to reality, insofar as, in traditional grammatical works, data created by the author himself predominates (if not unique). There is, in none of them, a literary corpus from which to base their assertions and rules. When dealing critically with this issue, it is important to

clarify that there are two distinct conceptions of grammar underlying it, coming from different aspects of language studies. On the one hand, there is grammar coming from the Greco-Latin tradition (and, therefore, qualified as traditional), which, in theory, bases its precepts on literary texts. On the other hand, there are the grammatical models developed by different theories within linguistics, whose empirical scope covers all manifestations of the language, whatever they may be. After a brief history of the constitution of traditional grammar, the origins of Portuguese grammar are specifically analyzed, showing that it was effectively created by the transposition of the doctrine consolidated in the Latin grammatical tradition and by the accommodation of the vernacular language within it. The source of normativity was not actually usage as such, but what was already said in Latin grammars.

Keywords: Grammar. Portuguese. Critical Historiography.

Laudatio

Este texto é uma reflexão crítica sobre as origens da gramática do português. Foi escrito em homenagem a Ataliba Teixeira de Castilho, um dos mais profícuos e importantes linguistas brasileiros. A ideia me veio da leitura da introdução e do primeiro capítulo da sua monumental *Nova Gramática do Português Brasileiro*. Espero que esta homenagem não fique muito aquém das muitas qualidades do homenageado. Ataliba foi da geração que assumiu a tarefa de institucionalizar a linguística no Brasil. Esteve entre os fundadores da Associação Brasileira de Linguística (ABRALIN) e do Grupo de Estudos Linguísticos do Estado de São Paulo (GEL), as duas entidades pioneiras na congregação dos linguistas brasileiros. A ele devemos avanços incalculáveis no estudo do português brasileiro advindos dos dois grandes projetos coletivos que organizou e coordenou com brilhantismo: o projeto "Gramática do Português Falado", do qual resultaram os volumes da *Gramática do Português Culto Falado no Brasil*; e o projeto "História do Português Brasileiro", que motivou tanto a busca da documentação histórica pertinente pelos arquivos do país quanto grande quantidade de estudos da história social e gramatical da língua no Brasil, reunidos em onze expressivos volumes organizados por Ataliba (Castilho, 2018-2022). Além disso, ele tem se dedicado também a provocar os linguistas brasileiros para que se empenhem num esforço teórico criativo de modo a tornar a nossa linguística fonte de proposições teóricas, garantindo uma interação de duas mãos com as vanguardas externas. Por tudo isso, sou grande admirador de Ataliba de Castilho, uma pessoa que

tem me inspirado em várias dimensões da pesquisa linguística. Acima de tudo, me sinto muito honrado por desfrutar de uma longa amizade fraterna com Ataliba.

Introdução

É um lugar comum afirmar que a gramática tradicional padroniza os usos, tendo como referência os textos literários. Como buscarei mostrar neste texto, é uma crença que, em geral, não corresponde à realidade, na medida em que, nas obras gramaticais tradicionais, predominam (quando não são únicos) os dados criados pelo próprio autor. Não há, em nenhuma delas, um *corpus* literário do qual se parte para as suas asserções normativas. No entanto, a crença é tão recorrente que o próprio Ataliba, ao apresentar a sua *Nova Gramática do Português Brasileiro*, teve de registrar a ressalva de que não seguia o rumo da tradição e acrescentou uma frase lapidar: "não acho que os escritores trabalham para nos abastecer de regras gramaticais" (Castilho, 2010, p. 32).

No trato crítico dessa questão, é importante ter claro que há, aí, subjacentes duas concepções distintas de gramática, saídas de vertentes diversas dos estudos das línguas. De um lado, está a gramática vinda da tradição greco-latina (e, por isso, qualificada de *tradicional*), que, em tese, funda seus preceitos nos textos literários. De outro, estão os modelos gramaticais elaborados por diferentes teorias no interior da linguística, cujo escopo empírico alcança todas as manifestações da língua, quaisquer que sejam.

São, portanto, dois projetos epistemológicos bem diferenciados. A gramática tradicional foi concebida pelos gregos alexandrinos, por volta da virada do século II AEC para o século I AEC, como *tékhnē* e não como *episteme*. Nasceu como um saber prático e não como um saber propriamente teórico-abstrato. Em outros termos, foi, primordialmente, vista como arte e não como ciência. Seu eventual estatuto de ciência veio à tona mais tarde e introduziu uma ambivalência (ou um dilema) entre arte e ciência que nunca se resolveu adequadamente.

Já os modelos gramaticais elaborados no interior da linguística são concebidos como atividade de natureza científica e começaram a emergir nas primeiras décadas do século XX com o surgimento da linguística estrutural, a dimensão sistêmico-sincrônica da moderna ciência da linguagem, nascida esta, em sua dimensão histórico-comparativa, na virada do século XVIII para o XIX.

A gramática tradicional agrega dois eixos: um aparato metalinguístico, que funciona como uma matriz descritiva, e um propósito normativo (Vieira, 2020). Já os modelos gramaticais da linguística têm apenas um eixo: um aparato metalinguístico. São, portanto,

modelos formais teórico-analíticos, sem qualquer propósito normativo, como é próprio, aliás, dos modelos teóricos de qualquer das ciências modernas. Nelas, se constroem modelos teóricos para analisar a realidade como tal, sempre sem qualquer intento normativo.

Ataliba, em sua *Nova Gramática do Português Brasileiro*, não está alinhado ao projeto epistemológico da gramática tradicional (o que justifica a ressalva que mencionamos anteriormente), mas ao projeto epistemológico da moderna ciência da linguagem. Seu interesse é apresentar, a partir de dados do projeto NURC (Norma Urbana Culta) e do projeto Gramática do Português Falado, aspectos da organização estrutural do português brasileiro, assumindo, como quadro teórico, uma perspectiva multissistêmica funcionalista-cognitivista, que concebe a língua como um conjunto de processos estruturantes. Não há, por conseguinte, nenhum propósito normativo nessa gramática.

Essa distinção entre os dois projetos epistemológicos nem sempre é bem compreendida pelos letrados, que costumam entender gramática apenas no sentido tradicional, por ser este o formato que lhes foi apresentado em sua experiência escolar. A própria ideia de que a língua é objeto de ciência (no sentido moderno do termo) lhes é, em geral, estranha, quando não desconhecida.

Dessa limitação conceitual e desse desconhecimento científico decorrem incompreensões e equívocos, nos debates públicos, sobre os objetos, objetivos e métodos da ciência da linguagem. O mais corriqueiro deles é o que atribui aos linguistas uma atitude antinormativa e, por consequência, a aceitação de um vale-tudo em matéria de língua. Chega-se ao absurdo de dizer que os linguistas são destruidores da língua, asserção que apenas revela um iletramento científico lastimável.

É óbvio que à ciência da linguagem interessam todas as manifestações da língua, sejam orais ou escritas, sejam prestigiadas ou estigmatizadas, sejam atuais ou antigas. Nesse sentido, para os linguistas vale, sim, tudo. Ou seja, qualquer fenômeno linguístico estará em seu campo de interesse teórico-analítico. Sua meta é produzir um saber sobre a língua como tal em qualquer de suas múltiplas manifestações. Seu foco é empiricamente amplo e não se restringe à língua escrita dita culta conforme se manifesta em textos literários, como, em tese, o faz a gramática tradicional em seu propósito normativo.

A linguística não se restringe a manifestações escritas, mas nada impede que as inclua em suas investigações teórico-analíticas, na medida em que elas são manifestações da língua dentre tantas outras. Nunca o fará, porém, com intuito normativo porque não é tarefa da ciência normatizar. As atividades normatizadoras envolvem outros fatores além do saber científico. Muitas vezes, inclusive, nem sequer se pautaram ou se pautam por esse saber.

Isso tudo não significa que os linguistas são antinormativos. Em primeiro lugar, porque, quando escrevem seus textos acadêmicos, seguem, mesmo que criticamente, as formas da língua de uso costumeiro no mundo acadêmico. Por outro lado, têm sido os linguistas a contribuir para o esclarecimento teórico-analítico dos processos normativos (Faraco; Zilles, 2017). Tal esclarecimento tem lhes dado instrumentos para definir adequadamente o papel sociocultural de um padrão de língua. E tem lhes dado, igualmente, no caso brasileiro (entre outros), instrumentos para uma avaliação crítica da história e de aspectos da norma-padrão tradicional (Vieira; Faraco, 2024, no prelo).

Fazer tais críticas não implica, contudo, em condenar a normatividade como tal. Para usar uma metáfora cansada, mas ainda expressiva, os linguistas não jogam o bebê da norma-padrão com a água do banho de suas pertinentes críticas. Cabe, sim, normatizar. Mas esta não é uma tarefa da linguística como ciência, embora seja racionalmente desejável que os saberes que ela produz sobre a língua sejam tomados em consideração em projetos normatizadores, para evitar decisões anacrônicas e arbitrárias. Esse é, no caso brasileiro, um tema em aberto, considerando que, apesar de pontuado aqui ou ali por filólogos e linguistas desde o século XIX, continua sem um enfrentamento sistemático.

Admitamos que não é fácil, para quem tem apenas a referência do saber gramatical escolarizado (que é, em geral, apresentado de forma categórica e dogmática) e opera apenas com a visão restrita de senso comum a propósito da heterogeneidade da língua, aceitar a historicidade e a relatividade dos preceitos normativos e, ao mesmo tempo, entender a relevância de uma norma-padrão numa sociedade de cultura escrita sedimentada. Só mesmo um saber (socio)linguisticamente bem informado poderá favorecer esse entendimento.

Breve histórico

A crença de que há um vínculo estreito entre a gramática tradicional e os "bons autores" tem, certamente, a ver com suas origens. A tradição gramatical greco-latina foi iniciada pelos filólogos que trabalhavam na Biblioteca de Alexandria. Dedicavam-se ali ao estudo dos textos dos poetas e prosadores prestigiados em sua cultura – Homero em particular. Em sua atividade, comparavam manuscritos diversos do mesmo texto, buscando, por meio do trato crítico das divergências, lacunas, interpolações e trechos ilegíveis, fixar a forma supostamente mais próxima da original. Para o desenvolvimento dessas atividades filológicas, necessitavam de um aparato terminológico e conceitual (uma metalinguagem) para poder falar sobre a língua desses textos.

Sabemos pouco sobre os gestos iniciais da construção desse aparato metalinguísti-

co. Borges Neto (2022), em sua *História da Gramática*, mostra que os gregos, já bem antes dos alexandrinos, refletiam sobre a língua, seja numa perspectiva retórica, seja numa perspectiva lógico-filosófica. Havia, portanto, um conjunto de saberes, um *a priori* categorial e conceitual, do qual os alexandrinos se serviram na constituição de suas *tékhnai grammatikai*, formatando-as por meio de um processo eclético que combinou elementos de lógica, metafísica, prosódia, métrica, retórica, assim como propriedades morfológicas da língua grega clássica, conforme se pode constatar na mais antiga *tékhnē grammatikē* que conhecemos, atribuída ao filólogo alexandrino Dionísio Trácio (c. 160 AEC - 85 AEC) e escrita, provavelmente, na virada do século II AEC para o século I AEC.[1]

Esses manuais articulavam um saber prático utilizado como auxiliar para a execução do trabalho filológico. Não era, nesse sentido, um ramo autônomo do conhecimento, como viria a ser mais tarde. E, em si, eram mais de caráter expositivo do que explicitamente normativo. Na *tékhnē* atribuída a Dionísio Trácio, por exemplo, há um único momento que pode ser entendido como normativo, qual seja, sua caracterização da leitura como "a pronunciação *impecável* dos poemas ou dos escritos em prosa" (*apud* Chapanski, 2003, p. 21 – destaque acrescido), assim detalhada:

> Ao ler é necessário atentar para a entonação, a prosódia, a arte das pausas, o sentido contido <no texto>. É, pois, necessário ler a tragédia em tom heroico, a comédia coloquialmente, as elegias, plangentemente, a épica vigorosamente, a lírica melodiosamente e os lamentos humilde e tristemente. De fato, a <leitura> feita sem a observação desses <princípios> arruína o valor dos poetas e ridiculariza a postura dos leitores. (*apud* Chapanski, 2003, p. 21)

Isso não quer dizer que os filólogos alexandrinos não estivessem pensando a língua também com propósito normativo. O próprio objeto recortado por eles (o grego clássico e não o grego alexandrino de seu tempo) e o material empírico que tomavam como referência para codificar as características do grego clássico (os textos dos autores antigos) os induziam a uma atitude normatizadora.

Desse modo, a descrição como tal era já um guia normatizador: eis aí a língua modelar a ser imitada. Por isso, talvez, a remissão direta aos textos literários fosse, de fato, pequena, quando não rara. Na *tékhnē* atribuída a Dionísio Trácio, o autor recorreu mais a exemplos de sua lavra e só esporadicamente transcreveu trechos de Homero.[2]

[1] Sobre a controversa questão da autoria dessa *tékhnē grammatikē*, pode-se consultar Chapanski (2003) e Borges Neto (2022). A tradução do texto para o português, enriquecida com notas e comentários, se encontra em Chapanski (2003).

[2] Vale notar que Louis Holtz (1981), em seu rigoroso estudo da obra do gramático latino Donato, mostra que os exemplos literários são raros na sua *Ars Minor* e são abundantes apenas na parte terceira da sua *Ars Maior*, dedicada aos vícios, virtudes e figuras de linguagem.

Pode-se tomar esse manual como paradigmático do novo gênero textual/discursivo criado pelos alexandrinos, quer pela matriz de conceitos e termos que assentou, quer pelo modo expositivo que desenvolveu. Seu arcabouço metalinguístico, além de considerações sobre a face sonora da língua (necessárias, como se viu, para orientar a leitura em voz alta dos textos), incluía, como uma espécie de núcleo duro, uma doutrina sobre as classes de palavras. A esta se agregou, duzentos e tantos anos depois, uma primeira sintaxe da oração completa, com o tratado de Apolônio Díscolo (séc. II), que viria a ser retomado e ampliado por Prisciano de Cesareia (sé. V-VI) em sua portentosa obra *Institutiones Grammaticae*, uma espécie de grande suma do labor gramatical greco-latino.

O modo expositivo adotado por Dionísio Trácio, por sua vez, permaneceu como característico do discurso de toda a tradição de raiz greco-latina. Sua marca é a exposição doutrinária. O *dixit* dos gramáticos tradicionais se faz, como diz Ataliba de Castilho (2010, p. 32), no "estilo revelação". Há um saber estabelecido *a priori* que é apresentado *ex-cathedra*, num movimento contínuo de retomada e reenunciação do que já foi dito. Algumas vezes, em obras maiores, esse saber é ampliado ou ressignificado, sem que seu núcleo duro se altere substancialmente. O "estilo revelação" é, certamente, o mais afinado com a apresentação de um saber que se assenta antes num *a priori* metafísico do que num *a posteriori* físico.

Comentando o modo expositivo de Donato (que se pode tomar como representativo do *dixit* dos gramáticos), diz Louis Holtz:

> O conjunto dessa construção tem um aspecto severo e dogmático. A severidade, a austeridade da exposição técnica tributária desse método são características surpreendentes – toda uma estilística está em causa: definições abstratas constituídas com base num único molde; enumerações secas; a frequência do verbo de cópula *est, sunt*; transições artificiais (*nam, autem*) ou a ausência de transição. O dogmatismo se expressa pelo caráter universal e sem réplica das noções consideradas. Não se trata aí de um corpo de doutrinas que se cria a nossos olhos numa meditação livre, de uma pesquisa aberta, pondo em causa seu objeto num movimento dialético, mas de uma série de afirmações de grande peso que têm por efeito constituir uma espécie de inventário. (Holtz, 1981, p. 54)[3]

[3] No original: "L'ensemble de cette construction a un aspect sévère et dogmatique. La sévérité, l'austérité de l'exposé technique tributaire de cette méthode sont des caractéristiques frappantes : toute une stylistique est en cause : des définitions abstraites bâties sur un unique moule ; de sèches énumérations ; la fréquence du verbe copule *est, sunt* ; des transitions factices (*nam, autem*) ou l'absence de toute transition. Le dogmatisme s'exprime par le caractère universel et sans réplique des notions envisagées. Il ne s'agit pas ici d'un corps de doctrines qui se crée sous nos yeux dans une libre méditation, d'une recherche ouverte, remettant en cause son objet dans un mouvement dialectique, mais d'une série de très pesantes affirmations qui ont pour effet de constituer une sorte d'inventaire."

Com o passar do tempo, a gramática da tradição greco-latina deixou de ser apenas um manual auxiliar dos estudos filológicos e se tornou, primordialmente, um objeto pedagógico. A educação linguística da Antiguidade clássica, conforme relatada por Quintiliano (c.35 - 96) em seu livro *Institutio Oratoria*, incluía três estágios. Primeiramente, se fazia o ensino da leitura e da escrita. Na sequência, o estudante passava a frequentar as aulas do *grammaticus*. O currículo dessa etapa se centrava na leitura e comentário dos textos dos autores tomados como modelares. Como auxiliar dessa atividade, entrava, então, o estudo da gramática, da qual, recomendava Quintiliano, se devia ensinar tudo, menos "o que é inútil" (I.VII.34 – Pereira, 2005, p. 172), recomendação que, infelizmente, parece não ter vingado na nossa pedagogia da gramática, tão preocupada com minúcias. O curso do *grammaticus*, esse mergulho nos textos tomados como modelares, era uma etapa propedêutica para o objetivo maior da educação linguística greco-latina, qual seja, o estudo e a prática da retórica, com vistas a formar o bom orador para atuar nos tribunais e nos espaços políticos.

Ao se tornar objeto de ensino, a gramática agregou a seu conceito inicial (como está na *tékhnē* atribuída a Dionísio Trácio) de "conhecimento empírico do comumente dito <nas obras> dos poetas e prosadores" (Chapanski, 2003, p. 21) o de "arte de falar corretamente", como se lê em Quintiliano (I.IV.2 – Pereira, 2005, p. 85). O ensino do *grammaticus* romano abarcava, então, essas duas faces: a leitura e comentário dos textos dos autores consagrados (para isso, a metalinguagem da gramática servia de apoio técnico) e as coordenadas da correção linguística (o propósito normativo da gramática como a arte de falar corretamente), que tinha como meta alcançar o ideal da *Latinitas*, da língua modelar, cultuado pela elite romana letrada.

Os próprios parâmetros de correção se tornaram tema de reflexão. Quintiliano, por exemplo, dedicou todo o capítulo VI do Livro I de seu *Institutio Oratoria* para desenvolver uma longa discussão sobre esses parâmetros. Ele fazia menção a quatro critérios: a razão – que, em termos de hoje, equivaleria a sistema ou organização estrutural da língua, no qual Quintiliano incluía os processos analógicos –, a antiguidade, a autoridade e o uso. Ao mesmo tempo, comentava os limites de cada um deles como efetiva baliza de correção.

O longo arrazoado de Quintiliano sobre esses critérios mostra que, embora a correção linguística fosse um valor cultural, não tinha, de fato, critérios absolutos, nem se esgotava na autoridade de poetas e prosadores. A analogia pode, por exemplo, justificar novidades estruturais; o uso, muitas vezes, supera a autoridade; os poetas costumam, por razões de métrica, escapar do esperado; as expressões antigas podem conferir majestade

ao discurso, mas é preciso moderação para evitar afetação e mesmo prejudicar a clareza, que é a maior virtude do discurso. Por fim, o uso, embora seja ele o "o mais seguro mestre do falar" (I.VI.3 – Pereira, 2005, p. 138), não deve ser entendido liberalmente, ou seja, como qualquer uso, mas apenas o que resulte "[d]o consenso dos eruditos" (I.VI.45 – Pereira, 2005, p. 160). O discurso de Quintiliano sobre os critérios de correção é rigoroso, mas, ao mesmo tempo, ponderado. Não é um dizer categórico e dogmático. Chega a recomendar, em um ponto, ousadia (I.V.72 – Pereira, 2005, p. 136) e, em meio a certas dúvidas, afirma: "não me parece desprovido de graça dizer que uma coisa é falar latim, outra falar gramatiquês" (I.VI.2– Pereira, 2005, p. 150), saudável distinção que raramente recorre em discursos normativos contemporâneos dogmáticos e pouco esclarecidos.

Os dois conceitos de gramática – como o conhecimento empírico do dito nas obras de poetas e prosadores e como a arte de falar corretamente – apareceram ainda em gramáticos da Antiguidade tardia, como em Diomedes (séc. IV), que acrescentou ao segundo conceito a correção do escrever: "(...) mas toda a gramática consiste principalmente na compreensão de poetas e escritores e na pronta exposição de histórias e na maneira correta de falar e escrever"[4] (Diomedes, [séc. IV] 1857, p. 426 – tradução minha).

À medida que o período medieval, pós-queda de Roma, foi avançando, aumentou também a distância entre os falares vulgares (protorromânicos) e o latim do período clássico, tal como aparecia nos textos dos autores romanos consagrados. Em consequência, o primeiro conceito foi desaparecendo e as gramáticas passaram a ser entendidas apenas como a arte de falar e escrever corretamente. E os próprios parâmetros de correção deixaram de ser adequadamente problematizados.

O panorama linguístico da Europa latina, depois das invasões germânicas e da fragmentação do Império do Ocidente, era bastante complexo. De um lado, estavam os eruditos (uma pequena elite) escrevendo em latim e tentando manter os padrões clássicos, o que era cada vez mais difícil, considerando que tais padrões eram, de fato, uma referência linguística congelada no tempo porque perdera sua vitalidade original com a desintegração progressiva do Império. Esses mesmos eruditos procuravam também falar, em situações formais (no ensino, na administração política e religiosa e na diplomacia), um certo latim que tinha como referência só a escrita, na medida em que não era mais a língua viva da população. Na comunicação cotidiana, estavam em franco desenvolvimento as novas línguas vernáculas, herdeiras diretas das diferentes variedades do latim popular falado nas várias regiões do Império, diferenciação que veio a se intensificar pelo contato com as línguas dos povos germânicos.

[4] No original: "tota autem grammatica consistit praecipue intellectu poetarum et scriptorum et historiarum prompta expositione et in recte loquendi scribendique ratione".

Nesse panorama, é perfeitamente compreensível que a gramática se reduzisse a seu conceito de arte de falar e escrever corretamente. O latim dos letrados era cada vez mais artificial. Funcionava, de fato, como uma segunda língua, com o agravante de não haver mais falantes seus como primeira língua. Seu domínio exigia um elaborado trabalho intelectual que tinha de começar por uma apresentação da língua. A gramática deixou de ser, então, suplemento para falantes interessados em aperfeiçoar o domínio de sua língua primeira e passou a ser ponto de partida para se chegar a conhecer a língua que lhes era quase estrangeira. Construiu-se, por conseguinte, um modelo pedagógico que, além de fortemente normativo (era preciso dizer continuamente o que era o latim "correto", já que não havia referência viva de qualquer modelo social), partia sempre do estudo gramatical. Só depois disso é que se passava à leitura e estudo dos textos e à prática da fala e da escrita. A gramática se desvinculou do estudo literário como tal e adquiriu autonomia. Estava entre as sete artes liberais; constituía, com a retórica e a dialética, o *trivium* e era entendida como a base de todo o currículo. Como diz Law (2003, p. 101):

> A gramática era vista como a origem e o fundamento de todo o resto, o [saber] mais básico (nos dois sentidos da palavra). Era frequentemente retratada como uma professora conduzindo crianças pequenas em direção a coisas mais altas, brandindo um canivete para usar não apenas para afiar canetas de pena, mas também para extirpar os erros de seus alunos.[5]

Outro efeito da escolarização da gramática foi a inclusão, nos manuais, de toda uma parte dedicada aos "vícios e virtudes", englobando comentários sobre barbarismos, solecismos e linguagem figurada. Ampliava-se, assim, o eixo normativo da exposição gramatical, conforme se pode ver já em Quintiliano e, depois, em gramáticos da chamada Antiguidade tardia, como na *Ars* de Mário Plócio Sacerdote (século III), na *Ars* de Diomedes (século IV) e na Parte III da *Ars Maior* de Donato (século IV). Esta última foi extensamente usada no ensino do período medieval e será referência para a escrita de gramáticas das línguas vernáculas a partir do século XV. A ela se conjugou a obra monumental de Prisciano (séc. V-VI) retomada pelos intelectuais envolvidos no chamado renascimento carolíngio, no século IX (Law, 2003, cap. 7).

Na Baixa Idade Média, o estudo gramatical tomou um novo rumo, paralelo à tradição greco-latina escolarizada. A filosofia escolástica motivou a busca de uma teoria ca-

[5] No original: "Grammar was viewed as the origin and foundation of all the rest, the most basic (in both senses of the word). It was often depicted as a woman teacher leading young children on to higher things, brandishing a penknife for use not only in sharpening quill pens but also in excising the mistakes of her pupils."

paz de dar suporte lógico-filosófico às afirmações sobre a linguagem, teoria que tomou forma em várias obras, das quais a mais destacada foi o *Tractatus de modis significandi seu grammatica speculativa*, escrito por Tomás de Erfurt (? – c.1310) por volta de 1300.

Sobre essas gramáticas, diz Borges Neto (2017, p. 14):

> Ao avesso da tradição gramatical, iniciada na Grécia do século II a.C. e continuada até o início da Baixa Idade Média, que dava à gramática uma função basicamente pedagógica, as gramáticas dos especulativos procuravam estabelecer uma *teoria* capaz de suportar logicamente as afirmações sobre a linguagem. Enquanto as gramáticas anteriores se dedicavam a determinar os valores fonéticos das letras e das sílabas, estabelecer os paradigmas flexionais das palavras e as possibilidades combinatórias (sintáticas, no fundo) das palavras na composição das orações – sempre com o objetivo de chegar a um ideal de fala/escrita "correta" –, a preocupação dos especulativos era com a significação. Interessava-lhes saber como as expressões linguísticas podiam significar. As gramáticas dos especulativos, de modo geral, só olhavam para as classes de palavras para investigar a contribuição que as palavras de cada classe davam para o significado da oração.

Essa diferença entre as duas abordagens fica bem clara na afirmação de um desses gramáticos, Boécio da Dácia (c.1240 – 1280), a propósito de Prisciano:

> [...] como Prisciano não explica a gramática por meio do conhecimento de seu funcionamento interno, sua teoria é muito limitada. Daí que tire muitas conclusões gratuitas [...] que justifica recorrendo à autoridade dos gramáticos antigos [...] (Boécio da Dácia, Questão 9, *apud* Beccari, 2017, p. 37)

Prisciano era, sim, uma referência respeitada quanto à apresentação da gramática do latim. No entanto, merecia críticas desses pensadores por não ter ido além da mera descrição, deixando de realizar uma investigação sobre os fundamentos lógico-filosóficos da organização da língua. A "correção" estava presente no pensamento desses gramáticos, mas era entendida sob outro ponto de vista. Não se tratava de buscar o "correto" com base, digamos assim, nos critérios de Quintiliano (que incluíam a autoridade dos autores consagrados), mas de estabelecê-lo com base num sistema de regras lógicas manifesto nas propriedades sintáticas e semânticas das expressões.[6] A gramática deixou de ser pensada como *tékhnē* e passou para o patamar da *episteme* – passou a ser pensada por um quadro teórico lógico-metafísico universalista. A teoria era exemplificada com expressões do

[6] É plausível dizer que, dentre os critérios de Quintiliano, os gramáticos especulativos se centravam na *ratio*, ressignificada como o todo organizacional de base lógico-filosófica da língua.

latim (a língua de cultura do mundo medieval), mas, por buscar definir o fundamento lógico-metafísico da linguagem, era entendida como válida para qualquer língua.[7] Por isso mesmo, não havia, na exposição argumentativa dos especulativos, qualquer referência à *auctoritas* de poetas e prosadores. Não lhes interessava nem o "comumente dito <nas obras> dos poetas e prosadores", nem "a arte de falar e escrever corretamente". Seu objeto eram os modos de significar, como se pode ler em Tomás de Erfurt (c. 1300, *apud* Beccari, 2017, p. 99):

> Visto que, em toda ciência, saber e entender relacionam-se com o conhecimento de princípios, como está escrito no primeiro livro e no comentário da *Física* de Aristóteles, é oportuno para nós, desejosos de conhecer a ciência da gramática, determo-nos em todos os seus princípios, cujos modos são os modos de significar.

Como mostra Beccari (2017), a gramática especulativa perdeu fôlego a partir da segunda década do século XIV frente a uma ampla adesão de pensadores escolásticos ao nominalismo. Com isso, um programa teórico que assumia pressupostos realistas, como o dos gramáticos especulativos, deixou de interessar. Ao mesmo tempo, o surgimento do Humanismo, postulando uma retomada das investigações filológicas de obras literárias da Roma clássica e dos estudos da retórica, reavivou uma perspectiva de estudo retórico-estilístico da língua latina, na qual não havia lugar para as teorizações lógico-filosóficas.

Estas retornarão, porém, no século XVI, com o desenvolvimento de uma linhagem gramaticográfica racionalista de caráter logicista e universalizante adotada por autores europeus que escreveram gramáticas do latim, em especial o italiano Júlio César Escalígero (1484-1558) e o espanhol Francisco Sanchez de las Brozas (1523-1600). O primeiro escreveu *De causis linguae latinae* (1540) e o segundo, *Minerva - seu de causis linguae latinae* (1587). Essas obras se opunham à tradição humanista, representada pincipalmente por outro autor italiano, Lorenzo Valla (1407-1457), e sua *Elegantiae linguae latinae*, publicada em 1444. Enquanto Valla – plenamente inserido no Humanismo italiano – produziu uma obra fundamentalmente estilística fundada no uso dos autores latinos clássicos, os gramáticos Escalígero e Sanchez de las Brozas defendiam que a gramática deveria estar fundada não no uso, mas na *ratio* (Apel, 1975). Tratava-se, para eles, não de glorificar e imitar as belezas estilísticas da *latinitas*, mas de investigar as causas e os princípios racio-

[7] Pode-se ler, em Borges Neto (2022), uma detalhada apresentação da gramática especulativa. Por outro lado, Beccari (2017) traz o texto original da gramática de Tomás de Erfurt e sua tradução para o português, acompanhados de um amplo acervo de notas com comentários e informações complementares e antecedidos de um cuidadoso estudo introdutório que situa historicamente o movimento modista.

nais ordenadores das estruturas linguísticas. Nesse sentido, são ilustrativas as asserções de Sanchez de las Brozas apresentadas a seguir.[8] As duas primeiras, extraídas do capítulo 1 do Livro I, justificam o título *Minerva – seu de causis linguae latinae*. A última asserção, agora do capítulo 2 do Livro I, retoma, com outras palavras, a condenação do dito (normativo) "O mestre o disse":

> Por que chamei de *Minerva - ou Causas da Língua Latina* a este produto das minhas insônias? Porque pretendo descobrir a causa e o verdadeiro princípio de cada questão, desaprovando, junto com Cícero, o dito pitagórico: 'O mestre o disse'. (Sanches de las Brozas, 1761 [1587], p. 8)[9]

> Em síntese, pois, há que se dar, em primeiro lugar, a explicação racional das coisas e só depois, se possível, virão os testemunhos, para que as coisas se convertam de boas em claras. (Sanchez de las Brozas, 1761 [1587], p. 7-8)[10]

> Posto que o tema de que tratamos há de ser demonstrado primeiro com a razão, depois com testemunhos e com o uso, ninguém deve estranhar se não sigo os homens ilustres. É que, por mais autoridades em que se apoie o gramático, se não demonstra o que diz com a razão e com exemplos, não será digno de crédito em nada e menos ainda em gramática. (Sanchez de las Brozas, 1761 [1587], p. 8)[11]

Essa perspectiva gramaticográfica desembocou, agora no trato da língua vernácula, na *Grammaire générale e raisonnée de Port-Royal* (em português, *Gramática de Port-Royal ou Gramática Geral e Razoada*), publicada em 1660 pelos franceses Antoine Arnauld (1612-1694) e Claude Lancelot (1615-1695). Esta obra se consolidou como o modelo da gramática dita racionalista ou filosófica. Seus autores defendiam que fossem investigados "os verdadeiros fundamentos da arte de falar" (Arnauld; Lancelot, 2001 [1660], p. 5), tendo como objeto, portanto, não apenas o uso da língua, mas, antes de tudo, o conhecimento de suas razões.

---

[8] Os textos latinos foram retirados da edição publicada em Amsterdam, em 1761. As minhas traduções para o português tomaram como referência a tradução para o espanhol da edição organizada por Sánchez; Caparro Gómez (1995).

[9] No original: "Quae causa fuit, ut has nostras vigilias, *Minerva, seu de causis linguae Latinae*, inscriberemus; quia cujus libet quaestionis causas & vera principia eruere conamur; improbantes cum Cicerone, Pythagoricum illud, Ipse dixit."

[10] No original: "Relinquum est igitur, ut omnium rerum ratio primum adhibeatur, tum deinde, si fieri poterit, accedant testimonia, ut res ex optima fiat illustrior."

[11] No original: "Quoniam res, de qua agimus, ratione primum, deinde testimoniis & usu, est comprobanda, nemo mirari debet, si magnos interdum viros non sequamur; nam quantacunque auctoritate mihi Grammaticus polleat, nisi ratione propositisque exemplis, quod dixerit, confirmaverit, nullam in re praesertim grammatica fidem faciet."

A gramática de Port-Royal viria a ser, positiva ou criticamente, referência de muitas das discussões sobre a linguagem que atravessaram o século XVIII, principalmente na França, onde se desenvolveram estudos gramaticais de monta, em especial aqueles vinculados à *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers* (Diderot; D'Alembert, 1751-1772). Os seus verbetes de temas gramaticais adotaram tanto perspectivas racionalistas, alinhadas a Port-Royal, quanto empiristas, alinhadas ao projeto antilatinizante iniciado, na primeira década do século XVIII, pelo gramático francês Claude Buffier (1661-1737) e expandido por Gabriel Girard (1677-1748) e Étienne de Condillac (1714-1780).[12]

Ambas as perspectivas repercutiram na gramaticografia do português. Contador de Argote (1676-1749) e Reis Lobato (1721?-1803?)[13], por exemplo, embora seguissem uma linhagem latinizante (pouco afeita, portanto, ao que propunham Arnauld e Lancelot), mencionaram Port-Royal nos preâmbulos de suas gramáticas, sem que se note qualquer repercussão estruturante daquele pensamento nas análises apresentadas nelas. A menção a Port-Royal nos preâmbulos foi, talvez, apenas uma concessão à moda teórica.

Mais à frente, autores como Bacelar (1783), Melo (1818) e Barbosa (1822), em Portugal, e Loureiro (1828), Duarte (1829) e Bezerra (1861), no Brasil, apresentaram suas obras como "gramáticas filosóficas", sendo a de Barbosa (1822) a mais substancial delas. Gonçalves (2006) traz uma lista de algumas dessas gramáticas e analisa o movimento gramatical racionalista em sua conjuntura e em suas características, destacando, em particular, seu caráter mais eclético do que dogmático. Diz ela em suas considerações finais (2006, s.p.):

> Fruto de condições sociais, políticas e religiosas que contribuíram para manter o ensino sob o controlo de agentes de mentalidade não "iluminada" – conjuntura conducente à expulsão dos Jesuítas e ao encerramento dos seus colégios –, o electismo ideológico manifestou-se em todas as formas do saber devido à recepção em diferido das ideias europeias, mormente as francesas. Essa característica denota-se em todos os domínios, não sendo a gramática excepção. Na verdade, tal como a língua, poderoso instrumento de comunicação e de circulação das ideias, a gramática participou da agenda do legislador e, devido a isso, foram franqueadas as portas a certas correntes e autores de pensamento pedagógico e linguístico. Para lá das

[12] Um estudo sobre as diferentes linhagens seguidas na gramaticografia das línguas vernáculas europeias pode ser lido em Faraco; Vieira (2021).

[13] Nada se sabe da vida de Antônio José dos Reis Lobato. Assunção (2000) levanta a hipótese de que ele nunca tenha, de fato, existido e o nome seja apenas um pseudônimo usado por alguém a serviço do Marquês de Pombal no processo de reforma do ensino, produzindo a gramática pedagógica necessária para a introdução do ensino de português, acoplado ao ensino de latim, nas escolas de Portugal, uma das diretrizes da reforma pombalina.

> influências implícitas, a variedade das correntes representadas pelas fontes citadas (racionalista, "sensista" ou "sensualista" e "ideologia"), bem como a integração desses contributos numa mesma obra, regra geral sem problematização nem dos conceitos nem de certa terminologia, faz da gramática filosófica portuguesa um produto ideologicamente ecléctico, de filiação variada, que exige um laborioso e paciente trabalho de cotejo textual, trabalho ainda por fazer. Talvez o interesse desta gramaticografia resida precisamente na tessitura híbrida que lhe advém da conciliação de pontos de vista nem sempre convergentes na origem.

A linhagem empirista, por sua vez, sob influência da Gramática do Castelhano, publicada pela Real Academia Española (RAE) em 1771, apareceu, pela primeira vez em Portugal, na obra *Rudimentos da gramática portugueza*, de Pedro José da Fonseca (1737-1816), publicada em 1799. No fim do século XIX, essa linhagem voltará a inspirar a produção de gramáticas portuguesas e brasileiras, que se declaravam "científicas" por criticarem a tradição das gramáticas filosóficas e por se dizerem seguidoras da linguística histórico-comparativa, a nova onda teórica que vinha da Alemanha.

Essa declaração de cientificidade por alinhamento com a linguística histórico-comparativa em nada alterou a organização das gramáticas. Primeiro, porque não estava no foco de interesse desse movimento teórico-analítico a produção de gramáticas expositivo-normativas. Por ser um projeto essencialmente histórico, motivou, sim, a criação de um novo gênero textual/discursivo – as gramáticas históricas. Assim, as informações históricas que apareciam nas gramáticas expositivo-normativas do fim do XIX e começo do XX eram, geralmente, apenas agregadas ao texto como uma espécie de ilustração erudita. As gramáticas em si não fugiam da matriz greco-latina. Eram, na verdade, projetos ecléticos em que, ao núcleo duro da tradição, agregavam comentários de história da língua (sob o impacto da grande novidade europeia da época, a linguística histórico-comparativa), formulações das gramáticas filosóficas (mesmo quando no plano retórico as criticassem) e apropriações de gramáticos de outras línguas, principalmente do inglês (como se pode ler, por exemplo, no Prefácio da segunda edição da gramática de Júlio Ribeiro[14]). Não se nega que eram manuais volumosos e de grande fôlego e contribuíram para sumarizar e sistematizar o saber gramatical tradicional. Essa consolidação vai ser repetida, sem significativas alterações, nas gramáticas brasileiras posteriores e, mais tarde, apenas ajustada à Nomenclatura Gramatical Brasileira de 1959.

[14] Prete (2020), além de fazer uma ótima revisão crítica das avaliações da gramática de Júlio Ribeiro correntes em estudos historiográficos contemporâneos da gramaticografia brasileira, mostra, com bastante propriedade, como se deu essa sistemática integradora na gramática desse autor, a qual, pode-se afirmar sem erro, caracterizou todas as gramáticas do período.

É preciso reconhecer, neste ponto, que a matriz greco-latina, tendo se constituído dentro das coordenadas epistemológicas especulativo-metafísicas do mundo grego, não tem em si a dinâmica da ciência moderna. Mantém-se, por conseguinte, como um saber acabado, fechado, estático. Não se deve, portanto, estranhar que as gramáticas tradicionais contemporâneas, no fundo, apenas se repitam. São as mesmas classes de palavras desde Dionísio Trácio (com alguns pequenos ajustes ao longo da história). Quanto à sintaxe, seu objeto nunca foi além da oração completa, como já estava posto, em linhas gerais, em Dionísio Trácio e ficou consolidado em Prisciano.

Cabe pontuar que, nessa tradição, os fenômenos sintáticos da coordenação e da subordinação foram sempre tratados a partir do estudo da classe das conjunções e dos pronomes relativos (quase como um epifenômeno dessa classe e subclasse) até se consolidarem no atual modelo, em que a subordinação das orações é tratada num capítulo próprio da sintaxe e as orações são classificadas pelo seu funcionamento, no interior de uma oração principal como substantivos, adjetivos ou advérbios. Essa classificação, que aproxima um tipo de oração subordinada e uma determinada classe de palavras, foi proposta, no início do século XIX, pelo médico e gramático alemão Karl Ferdinand Becker (1775-1849).

Becker foi autor de várias obras sobre linguagem, entre as quais uma gramática do alemão, publicada em 1829. Foi nesta que ele propôs a classificação das orações subordinadas em substantivas, adjetivas e adverbiais. A sua gramática foi traduzida para o inglês em 1830 e, a partir daí, sua classificação foi incorporada nas gramáticas do inglês e, progressivamente, adotada pelas gramáticas de outras línguas.[15] A primeira gramática do português a adotá-la, ao que tudo indica, foi a de Júlio Ribeiro (1881), leitor que foi das gramáticas do inglês da época, conforme mencionou no Prefácio da segunda edição, de 1884.

Pode-se dizer que este foi o último ajuste mais significativo na matriz greco-latina. Embora interessante, não passou, porém, de um rearranjo interno, considerando que a matriz como tal não tem mais vitalidade e, por isso mesmo, é passível apenas de repetição. como testemunham as gramáticas tradicionais contemporâneas. No caso do Brasil, as balizas dessa matriz foram congeladas, em 1959, por portaria do ministro da Educação, na Nomenclatura Gramatical Brasileira (NGB).[16]

Já foi comum dizer que a NGB embaraça os estudos da língua. Tal juízo, contudo, não procede. Ela foi fixada com o único objetivo de uniformizar a terminologia adotada

[15] Um detalhado estudo do longo processo que culminou no trabalho de Becker pode ser lido em Andrieu (2007).

[16] Sobre a história e a configuração da NGB, pode-se consultar Henriques (2009).

nos conteúdos gramaticais dos programas escolares. Por consequência, ela uniformiza a escrita de gramáticas escolares, que são, em essência, gramáticas tradicionais. Como esse modelo gramatical é acabado, fechado e estático (independentemente de haver ou não uma nomenclatura congelada), a NGB não limita ou embaraça nenhuma investigação no interior da linguística. Como apontei, no início, são dois projetos epistemológicos distintos que caminham, de fato, em linhas paralelas. É, por isso, um equívoco achar que a linguística veio reconstruir a gramática da tradição greco-latina ou que a ciência moderna da linguagem seja mera extensão da gramática tradicional.

É bem verdade que a terminologia tradicional e mesmo certos pressupostos sobre a natureza da língua vindos da tradição foram apropriados por teorias linguísticas e estão, conforme diz Borges Neto (2012), "naturalizados" e tem "uso protocolar" nessas teorias. Afora o complexo problema epistemológico aí embutido e discutido por Borges Neto (2012), essa apropriação dá visibilidade ao fato de que dois milênios e meio de uma tradição conceitual-analítica deixaram marcas profundas no pensamento ocidental sobre a linguagem e repercutem no modo como são construídos os modelos teórico-analíticos no interior da jovem ciência da linguagem, que nasceu, em sua dimensão histórico-comparativa, há apenas dois séculos e desenvolveu sua dimensão estrutural-sincrônica apenas a partir do início do século XX.

É curioso, por fim, observar que algumas gramáticas tradicionais contemporâneas, embora mantenham o núcleo duro da matriz greco-latina, agregam informações e teorizações retiradas de teorias linguísticas e apresentadas como parte de comentários de fatos gramaticais (como faz, por exemplo, Bechara (2000), no capítulo sobre o verbo, a propósito de tempo e aspecto). Trata-se, de certa forma, de uma naturalização reversa. Ou seja, as teorizações resultantes de análises linguísticas são tomadas como atemporais, definitivas e desvinculadas de seus modelos teóricos. Parece que as eventuais incongruências advindas dessas apropriações deixam de ser devidamente avaliadas.[17]

Há, portanto, relevantes e complexas questões epistemológicas envolvidas nas aproximações e apropriações, de parte a parte, dos saberes tradicionais e dos saberes oriundos da moderna ciência da linguagem. São questões que raramente têm sido enfrentadas. Debatê-las aqui fugiria ao escopo deste texto. Fica, porém, o registro de sua relevância.[18]

[17] Sobre essa questão, há uma interessante análise crítica da gramática de Bechara (2000) em Mulinacci (2016).

[18] O século XXI viu nascer, no Brasil, um outro tipo de gramática – as chamadas gramáticas de linguistas. Não há espaço aqui para discuti-las. Sobre elas, há análises interessantes em Vieira (2015), Faraco; Vieira (2016) e Borges Neto (2022).

Origens da gramática do português

Dá-se, como marco inicial da gramaticografia das línguas europeias modernas, a *Gramática Castellana*, escrita por Antônio de Nebrija (1441?-1522) e publicada em 1492.[19] Nebrija era professor de latim em Salamanca e tinha escrito uma celebrada gramática do latim (*Introductiones latinae*), publicada em 1481 e que conheceu sucessivas edições, sendo usada ainda no século XVIII. Conhecia bem, portanto, a tradição gramatical greco-latina, da qual se apropriou e com a qual dialogou na feitura de sua gramática do castelhano.[20]

A esse propósito, é curioso notar que, embora se apoiasse extensa e claramente em Prisciano, nunca mencionou seu nome no corpo da gramática do castelhano. São, porém, frequentes nela as remissões explícitas a Quintiliano. Talvez essa diferença se justifique pelo fato de Prisciano ser do conhecimento comum dos letrados desde, pelo menos, o renascimento carolíngio. Já o texto completo da obra de Quintiliano tinha sido redescoberto apenas em 1416, num monastério de St. Gall (Suíça), por Poggio Bracciolini (1380-1459), um dos mais importantes humanistas do Renascimento italiano. A partir daí, a obra passou a ser uma referência explícita obrigatória nos estudos de gramática e de retórica até, pelo menos, o século XVIII (Faraco, 2018).

Nebrija construiu sua gramática em dois movimentos expositivos complementares. Primeiro, tomou das gramáticas latinas tudo o que pudesse ser também aplicado ao castelhano. Desse modo, enobrecia a língua vernácula, mostrando-a semelhante ao latim. Num segundo movimento, foi completando esse seu quadro expositivo com formulações que davam conta daquilo que era específico do castelhano.

Desse modo, a gramática de Nebrija não foi da língua para a gramática, mas, sim, da gramática para a língua. Ou seja, o arcabouço da gramática greco-latina foi tomado como uma matriz pré-dada na qual o castelhano foi enquadrado. Daí decorre o fato de que não há um *corpus* de referência. O próprio gramático e só ele é quem criou os exemplos. Foi de si mesmo, de seu modo de falar a língua, que Nebrija tirou seu material empírico. Só muito esporadicamente trouxe exemplos literários, que ficaram restritos a comentários sobre métrica e sobre algumas das figuras de linguagem e foram retirados da obra de um único poeta – Juan de Mena (1411-1456). O trabalho de reduzir a língua castelhana a *artifício* (como declarou, no Prólogo, ser sua intenção) se constituiu, no fundo, numa ativi-

[19] Um panorama do pensamento linguístico do Renascimento e do processo inicial da gramatização das línguas vernáculas modernas pode ser lido em Bagno (2023, t. I, cap. 4).

[20] Entre as várias edições de sua gramática latina, Nebrija publicou, possivelmente em 1488, uma em que comparava o latim e o castelhano (Nebrija, 1996), uma espécie de exercício preparatório para a escrita de sua gramática do castelhano.

dade de acomodá-la a um quadro doutrinal já praticamente pronto. O *dixit* do gramático pode, então, ser assim resumido: Eis – como espelho da gramática latina – a gramática da língua da qual sou falante. Sigam-na! Foi por esse diapasão enunciativo que autores de gramáticas de outras línguas vernáculas escreveram seus manuais no século XVI.

Diante disso, parece que não há como escapar da conclusão de que não passa de um mito de origem a crença de que a gramática tradicional das línguas modernas toma os "bons autores" como ponto de partida de suas asserções. O seu início se fez, efetivamente, pela transposição da doutrina consolidada na tradição gramatical latina, acomodando nela a língua vernácula. A fonte da normatividade não foi propriamente o uso como tal, mas o que já estava dito nas gramáticas latinas.

Na gramaticografia inicial do português, Fernão de Oliveira (1507-1581), em sua *Gramática da Linguagem Portuguesa*, publicada em 1536, disse que iria notar o bom costume dos que sabem a língua, "para que outros muitos aprendam" (Oliveira, 1975, p. 38). No entanto, todos os exemplos são de sua lavra. O autor se assumiu, portanto, como representante qualificado e suficiente dos que "mais sabem a língua".

João de Barros (1496-1570), em sua *Grammatica da lingua portuguesa*, publicada quatro anos depois, não fez diferente. Definiu a gramática como "hũ modo certo e iusto de falar, & escrever, colheito do uso, e autoridade dos barões doutos" (Barros, 1540, p. 2). Assim dito, parece que o movimento seria da língua para a gramática. No entanto, na prática, ele se assumiu também como representante qualificado e suficiente dos "barões doutos". Todos seus exemplos são de sua criação. Nesse sentido, o modo como construiu sua gramática em nada diferiu de Nebrija ou de Oliveira.

E, por essa trilha, seguiram os gramáticos portugueses nos dois séculos e meio seguintes. Foi só com Pedro José da Fonseca, em 1799, que se introduziu, na gramaticografia portuguesa, a prática de abonar as asserções e regras com exemplos literários. Isso, porém, pouco ou nada alterou o *modus faciendi* dos gramáticos. A exposição continuou sendo da gramática para a língua e não ao contrário.[21] Em geral, se assume que a doutrina está consolidada e o gesto primordial é repeti-la, o que não impede que se façam alguns ajustes terminológicos e conceituais, críticas ao que disseram gramáticos anteriores e acréscimos que tomam por base o que escrevem gramáticos de outras línguas. E a exemplificação, por sua vez, oscila. Há gramáticas em que não há, praticamente, exemplos literários, só os forjados pelo autor (como a de Sotero dos Reis, de 1866); em outras,

[21] Tendo isso em conta, afirmar, como se lê ocasionalmente, que as gramáticas tradicionais do português registram os usos linguísticos assumidos como exemplares, adotando como referência os falantes mais cultos da sociedade, é revelar um profundo desconhecimento da história da gramaticografia da língua.

ocorre a exemplificação literária, mas predominam exemplos criados pelo próprio autor (como a de Júlio Ribeiro, de 1881); e, em outras ainda, é relativamente farta a presença de exemplos literários (como a de Grivet, de 1881, e a de Maciel, de 1894), em especial na parte dedicada à sintaxe. Por outro lado, é bem sabido que a exemplificação é sempre seletiva: só são recolhidos exemplos que abonem o que diz a doutrina. Se um "bom autor" tem usos que não correspondem a ela, estes ou nunca são tomados em conta ou são eventualmente citados com a observação de que devem ser evitados.[22] Essa seletividade reforça nossa tese de que o gesto enunciativo vai sempre da gramática para a língua e nunca ao contrário.

Considerações finais

O sociólogo Norbert Elias (1897-1990) dizia, a propósito de seu trabalho científico, que seu objetivo era "levantar o véu das mitologias que mascara nossa visão da sociedade" (Elias, 2001, p. 45). Ao lê-lo, alguns anos atrás, me ocorreu que talvez valesse a pena tentar levantar também o véu de algumas mitologias que mascara nossa visão da língua e, especialmente, da gramática.

Uma dessas mitologias é o lugar comum de que as gramáticas tradicionais formulam seus ditos e regras com base nos textos dos "bons autores". Procurei, neste texto que escrevi em homenagem a Ataliba de Castilho, fazer uma leitura crítica dessa mitologia. Penso que "desmitificar as ocultações" (Elias, 2001, p. 48) pode contribuir para levar à frente uma necessária discussão dos nossos processos normatizadores, de modo a delinear, para o português brasileiro, uma norma-padrão livre de artificialismos e arbitrariedades, com a qual se possa superar, positivamente, um ensino anacrônico e desfuncional, contribuindo, assim, para o desenvolvimento das práticas de escrita, hoje tão prejudicadas pela falta de referências realistas e pela consequente insegurança dos falantes. A *Gramática do Português Brasileiro Escrito* (Vieira; Faraco, 2023), tendo, como pano de fundo, os dados historiográficos aqui discutidos, busca dar contribuições para esse necessário debate.

[22] A esse propósito, vale mencionar dois dentre vários exemplos desse tipo de juízo. Almeida (2005, p. 180), por exemplo, dizia, sobre construções do tipo *seu pai dele*: "Tais modos de dizer, ainda que aforados em textos portugueses de bom cunho, devem evitar-se". E Laudelino Freire (s.d., p. 18), sobre a regra que estipula a obrigatoriedade da próclise do pronome oblíquo ao verbo quando precedido de *onde*, *quando*, *como*, dizia: "Os exemplos colhidos nos clássicos mais corretos, que porventura contrariem esta regra, não lhes tiram o caráter imperativo".

Referências

ALMEIDA, N. M. de. *Gramática metódica da língua portuguesa*. 45. ed. São Paulo: Saraiva, 2005.

ANDRIEU, W. *Le concept de subordination et les grammaires de l'anglais*. Thèse de Doctorat. Paris: Université de Paris X – Nanterre, École Doctorale Lettres, Langues et Spectacle, 2007.

APEL, K. O. *L'idea di lingua nella tradizione dell'umanesimo da Dante a Vico.* Bologna: Il Mulino, 1975.

APOLÔNIO DÍSCOLO. *Sintaxis.* Trad. e ed.: V. B. Botas. Madrid: Gredos, 1987.

ARGOTE, J. C. de. *Regras da Lingua Portugueza, espelho da lingua Latina*. Lisboa: Officina da Musica, 1725.

ARNAULD, A.; LANCELOT, C. *Gramática de Port-Royal*. [1660] Trad. B. F. Bassetto; H. G. Murachco. 2. ed. São Paulo: Martins Fontes, 2001.

ASSUNÇÃO, C. *A arte da Grammatica da lingua portugueza de António José dos Reis Lobato*: estudo, edição crítica, manuscritos e textos subsidiários. Lisboa: Academia das Ciências de Lisboa, 2000.

AZEREDO, J. C. de. *Gramática Houaiss da Língua Portuguesa.* 4. ed. São Paulo: Parábola, 2018.

BACELAR, B. de L. M. *Grammatica Philosophica, e Orthografia Racional da Lingua Portugueza, para se pronunciarem, e escreverem com acerto os vocabulos d'este idioma.* Lisboa: S. T. Ferreira, 1783.

BAGNO, M. *Uma história da linguística.* Tomo 1. São Paulo: Parábola, 2023.

BARBOSA, J. S. *Grammatica Philosophica da Lingua Portugueza ou principios da grammatica geral applicados á nossa linguagem. Lisboa: Academia das Sciencias, 1822.*

BARROS, J. *Grammatica da lingua portuguesa*. Lisboa: apud Lodovicum Rotorigiũ Typographum, 1540.

BECCARI, A. J. *Tratado sobre os modos de significar ou Gramática especulativa, de Tomás de Erfurt.* Curitiba: Editora da UFPR, 2017.

BECHARA, E. *Moderna gramática portuguesa*. 37. ed. Rio de Janeiro: Lucerna, 2000.

BECKER, K. F. *Grammar of the German Language*. London: John Murray, 1830.

BEZERRA, M. S. da S. *Compendio de grammatica philosophica.* Ceará: Typ. Social, 1861.

BOÉCIO DA DÁCIA. Modi significandi sive quaestiones super priscianum maiorem. *In*: McDERMOTT, A. C. S. (ed.) *Goffrey of Fontaine's abridgement of Boetius of Dacia's modi significandi sive questiones super priscianum maiorem.* Amsterdam: John Benjamin B. V., 1980.

BORGES NETO, J. Gramática tradicional e linguística contemporâneas: continuidade ou ruptura? *Todas as Letras*, São Paulo, v. 14, n. 1, p. 87-98, 2012.

BORGES NETO, J. Prefácio. In: BECCARI, A. J. *Tratado sobre os modos de significar ou Gramática especulativa, de Tomás de Erfurt.* Curitiba: Editora da UFPR, 2017, p. 13-18.

BORGES NETO, J. *História da gramática.* Curitiba: Editora da UFPR, 2022.

BUFFIER, C. *Grammaire Françoise sur un plan nouveau pour en rendre les principes plus clairs & la pratique plus aisée*. Paris: Nicolas Le Clerc, Michel Brunet, Leconte et Montalant, 1709.

CASTILHO, A. T. de. *Nova gramática do português brasileiro.* São Paulo: Contexto, 2010.

CASTILHO, A. T. de. (coord.) *Gramática do português culto falado no Brasil.* 7 volumes. São Paulo: Contexto, 2013-2016.

CASTILHO, A. T. de. (coord.) *História do português brasileiro.* 11 volumes. São Paulo: Contexto, 2018-2022.

CONDILLAC, E. de. *Grammaire (Cours d'études pour l'instruction du Prince de Parme). In*: *Oeuvres completes de Condillac, revue, corrigées par l'auteur et imprimées sur ses manuscrits autographes.* Tome huitième. Paris: Chez Dufart, 1803, p. 3-384.

DIDEROT, D.; D'ALEMBERT, J. le R. (Orgs.). *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers*. 35 vol. Paris: Briasson, David, Le Breton, Durand, 1751-1772. Disponível em: https://enccre.academie-sciences.fr. Acesso em: 14 maio 2024. Vários verbetes foram traduzidos para o português e publicados pela Editora da UNESP em 6 volumes (2015-2017).

DIOMEDES. Artis Grammaticae. Libri III. In: *Grammatici Latine – ex recensione Heinrich Keil.* V. 1. Lipsiae (Leipzig): B. G. Teubner, 1857, p. 297-529.

DIONÍSIO TRÁCIO. *Tékhne Grammatiké.* In: CHAPANSKI, G. *Uma tradução da* Tékhne *Grammatike, de Dionísio Trácio, para o português*. Dissertação (Mestrado em Letras). Curitiba: Universidade Federal do Paraná – Programa de Pós-Graduação em Letras, 2003. Disponível em: https://acervodigital.ufpr.br/handle/1884/24818. Acesso em 14 maio 2024.

DONATO. *Ars Minor* e *Ars Maior.* Ed. crítica. *In*: HOLTZ, L. *Donat et la tradition de l'enseignement grammatical.* Paris: Centre National de la Recherche Scientifique, 1981, p. 585-674. Trad. para o português: DEZOTTI, L. C. Arte menor *e* Arte maior *de Donatos: tradução, anotação e estudo introdutório.* Dissertação (Mestrado em Letras). São Paulo: Universidade de São Paulo – Programa de Pós-Graduação em Letras Clássicas, 2011.

DUARTE, A. da C. *Compendio de grammatica philosophica da lingua portugueza.* Maranhão: Typ. Nacional, 1829.

ELIAS, N. *Norbert Elias por ele mesmo.* Rio de Janeiro: Jorge Zahar Editora, 2001.

ESCALÍGERO (SCALIGER), J. C. *De causis linguae latinae.* [1540] Ed. et Trad.: Pierre Larder; Geneviève Clerico; Bernard Colombat. Genève: Droz, 2018.

FARACO, C. A. Quintiliano: Ars Oratoria e educação linguística. *In*: BECCARI, A. J.; BINATO, C. V. P.; FERREIRA, E. A. G. R. (orgs.). *Retórica - perspectiva histórica e atualidade*. Campinas: Mercado de Letras, 2018, p. 15-40.

FARACO, C. A.; VIEIRA, F. E. (eds.) *Gramáticas brasileiras: com a palavra os leitores.* São Paulo: Parábola, 2016.

FARACO, C. A.; VIEIRA, F. E. A linhagem empirista na gramática do século 18. *Revista da ABRALIN*, v. 20, n. 3, p. 464-492, 2021. DOI: 10.25189/rabralin.v20i3.1963. Acesso em: 20 abr. 2023.

FARACO, C. A.; ZILLES, A. M. *Para conhecer norma linguística.* São Paulo: Contexto, 2017.

FONSECA, P. J. da. *Rudimentos da grammatica portugueza*. Lisboa: Off. de Simão Thaddeo Ferreira, 1799.

FREIRE, L. *Regras práticas para bem escrever.* Rio de Janeiro: Editora A Noite, s.d.

GIRARD, G. *Vrais Principes de la Langue Française, ou la parole réduite en méthode conformement aux lois de l'usage*. Tome I. Paris: Chez Le Breton, 1747.

GONÇALVES, M. F. Iluminismo e pensamento linguístico em Portugal: o exemplo

das gramáticas filosóficas. *Actes VII Congrés de Lingüística General.* CDRom. Barcelona: Universitat Barcelona - Departament de Lingüística General, 2006, p. 146.

GRIVET, A. *Nova Grammatica Analytica da Lingua Portugueza.* Rio de Janeiro: Typ. De G. Leuzinger & Filhos, 1881.

HENRIQUES, C. C. *Nomenclatura gramatical brasileira: 50 anos depois.* São Paulo: Parábola, 2009.

HOLTZ, L. *Donat et la tradition de l'enseignement grammatical.* Paris: Centre National de la Recherche Scientifique, 1981.

LAW, V. *The History of Linguistics in Europe: from Plato to 1600.* Cambridge: Cambridge University Press, 2003.

LOBATO, A. J. dos R. *Arte da grammatica da lingua portugueza.* Lisboa: Regia Officina Typografica, 1770.

LOUREIRO, L. T. de. *Grammatica razoavel da lingua portugueza.* Rio de Janeiro: Typ. Imp. Nac., 1828.

MACIEL, Maximino. *Lingua Portugueza – Grammatica Descriptiva.* [1894] 9. ed. aum. e refund. Rio de Janeiro: Livraria Francisco Alves, 1925.

MELO, J. C. do C. e. *Gramática Filosofica da Linguagem Portuguêza*. Lisbôa: Impressão Régia, 1818.

MULINACCI, R. *Moderna Gramática Portuguesa*: habemus grammaticam? *In*: FARACO, C. A.; VIEIRA, F. E. (eds.) *Gramáticas brasileiras: com a palavra os leitores.* São Paulo: Parábola, 2016, p. 113-147.

NEBRIJA, E. A. de. *Gramática castellana*. [1492] Introducción y notas de Miguel Ángel Esparza y Ramón Sarmiento. Madrid: Fundación Antonio Nebrija/SGEL, 1992. Disponível em: https://bdh-rd.bne.es/viewer.vm?id=0000174208&page=1. Acesso em: 14 maio 2024.

NEBRIJA, E. A. de. *Instructiones latinae-Recognitio*. Introdución, edición crítica, tradución y notas de Eustaquio Sánchez Salor, Santiago Lopez Moreda, Maria Luisa Harto Moreda y Joaquin Villalba Álvarez. Vol. 1 y 2. Cáceres: PublicaUEx Editorial, 2022.

NEBRIJA, E. A. de. *Introducciones latinas contrapuesto el romance al latin*. [c.1488] Edição crítica de Miguel Angel Esparza e Vicente Calvo. Munster: Nodus Publikationen, 1996.

OLIVEIRA, F. de. *A Gramática da linguagem portuguesa.* Introdução, leitura actualizada e notas por Maria Leonor Carvalhão Buescu. Lisboa: Imprensa Nacional-Casa da Moeda, 1975.

PEREIRA, M. A. *Quintiliano gramático:o papel do mestre de Gramática na* Institutio Oratoria. 2. ed. São Paulo: Humanitas, 2005.

PRETE, M. *O perfil teórico-metodológico da* Grammatica Portugueza *de Júlio Ribeiro*. Tese (Doutorado em Filologia e Língua Portuguesa). São Paulo: Universidade de São Paulo – Faculdade de Filosofia, Letras e Ciências Humanas, 2020.

PRISCIANO. *Institutionum Grammaticarum Libri I-XII* – ex recensione Martin Hertz. *In*: *Grammatici Latini* – ex recensione Heinrich Keil. Vol. II. Lipsias (Leipzig): B. G. Teubner, 1855. PRISCIEN. *Grammaire.* Texto latino e trad. para o francês dos livros XI e XII: Groupe *Ars Grammatica.* Paris: J. Vrin, 2020.

PRISCIANO. *Institutionum Grammaticarum Libri XIII-XVIII* – ex recensione Martin Hertz. *In*: *Grammatici Latini* ex recensione Heinrich Keil. Vol. III. Lipsias (Leipzig): B. G. Teubner, 1858. PRISCIEN. *Grammaire.* Texto latino e trad. para o francês dos livros XIII a XVIII: Groupe *Ars Grammatica.* Paris: J. Vrin, 2010-2020.

QUINTILIANO, M. F. *Instituição oratória.* Tomos I-IV. Ed. em latim e português. Trad. e notas Bruno F. Bassetto. Campinas: Editora da UNICAMP, 2015-2016.

REAL ACADEMIA ESPAÑOLA. *Gramática de la Lengua Castellana*. Madrid, 1771.

REIS, F. S. *Grammatica Portugueza.* [1866] 2. ed. revista, corrigida e annotada. Maranhão: Typ. De R. d'Almeida & C. Editores, 1871.

RIBEIRO, J. *Grammatica portugueza*. São Paulo: Typ. de Jorge Seckler, 1881.

RIBEIRO, J. *Grammatica portugueza.* 2. ed. São Paulo: Typ. De Jorge Seckler, 1884.

SACERDOTE, M. P. *Artium Grammaticarum. In*: *Scriptores Artis Metricae* – ex recensione Heinrich Keil. Lipsiae (Leipzig): B. G. Teubner, 1874, p. 427-546.

SANCHEZ DE LAS BROZAS, F. *Minerva, seu de causis linguae latinae.* [1587] Commentarius Gasp. Scioppius & Jac. Perizoni. Amstelaedami: Fratrum de Tournes, 1761.

SANCHEZ DE LAS BROZAS, F. *Minerva, seu de causis linguae latinae.* Trad. para o espanhol: SÁNCHEZ SALOR, E.; CHAPARRO GÓMEZ, C. (eds.) Cáceres: Institución Cultural El Brocense, 1995.

TOMÁS DE ERFURT. *Tractatus de modis significandi seu grammatica speculativa.* (c. 1300) Tradução para o português: BECCARI, A. J. *Tratado sobre os modos de significar ou Gramática especulativa, de Tomás de Erfurt.* Curitiba: Editora da UFPR, 2017.

VALLA, L. *Elegantiae linguae latinae.* [1441] S. LÓPEZ MOREDA (ed.) Cáceres: Univerisdad de Extremadura, 1999.

VIEIRA, F. E. *Gramáticas brasileiras contemporâneas do português: movimentos de ruptura e linhas de continuidade com o paradigma tradicional de gramatização.* Tese (Doutorado em Linguística). Recife: Universidade Federal de Pernambuco – Programa de Pós-Graduação em Letras, 2015.

VIEIRA, F. E. Tradição gramatical: história, epistemologia e ensino. *In*: VIEIRA, F. E.; BAGNO, M. (Orgs.). *História das línguas, histórias da Linguística*: homenagem a Carlos Alberto Faraco. São Paulo: Parábola Editorial, 2020. p. 85-125.

VIEIRA, F. E.; FARACO, C. A. *Gramática do português brasileiro escrito.* São Paulo: Parábola, 2023.

VIEIRA, F. E.; FARACO, C. A. A construção de um padrão exonormativo para o português do Brasil. *In*: SOUSA, S. C. T. de; BRAGA, A. (Orgs.). *A dimensão política da língua(gem)*: perspectivas da Linguística Aplicada e das Teorias do Discurso. Campinas, SP: Pontes, 2024 (no prelo).

http://doi.org/10.5007/1984-8420.2025.e108135

ECONOMIA, SOCIEDADE E FORMAÇÃO DO PORTUGUÊS BRASILEIRO: ELEMENTOS PARA UMA DISCUSSÃO SOBRE AS CONFIGURAÇÕES SOCIOLINGUÍSTICAS DO BRASIL COLONIAL

ECONOMY, SOCIETY AND THE FORMATION OF BRAZILIAN PORTUGUESE: ELEMENTS FOR A DISCUSSION ON THE SOCIOLINGUISTIC CONFIGURATIONS OF COLONIAL BRAZIL

Ana Sartori | Lattes | anasartorii@hotmail.com
Universidade Federal da Bahia

Tânia Lobo | Lattes | tanlobo@gmail.com
Universidade Federal da Bahia

Pedro Daniel dos Santos Souza | Lattes | pdan.uneb@gmail.com
Universidade do Estado da Bahia

Resumo: A análise das fontes inquisitoriais legadas pelas duas primeiras visitações do Santo Ofício à América Portuguesa (Lobo; Sartori, 2020; Lobo; Sartori; Mota, 2016; Siqueira, 1978) não sustenta a interpretação da estratificação socioeconômica do Brasil colonial em termos dicotômicos – "casa-grande" *versus* "senzala" (Freyre, [1933] 2002) –, o que tem implicações na interpretação da formação do PB e na compreensão da atual configuração sociolinguística do país. Ressalte-se ainda outro aspecto pouco problematizado: a necessidade de compreender as interações linguísticas no seio das diversas configurações da família colonial. Perspectivada diacronicamente, a dita polarização sociolinguística do PB (Lucchesi, 1994, 2015), tão clara quando contrastadas a elite socioeconômica e a população escravizada, torna-se uma emaranhada teia, se postas em análise nuances que a realidade linguística e cultural no âmbito social e doméstico abrigava. A partir de Pagotto (2018), propõe-se aprofundar a reflexão sobre o que interpretações engendradas a partir da década de 1970 sobre a economia e a sociedade coloniais desvelam sobre a história do PB, destacando-se os seguintes pontos: o fato de a economia colonial não poder ser apreendida exclusivamente pela visão reducionista dos "ciclos", não se ter limitado ao modelo agro-minério-exportador, não guardar muitas semelhanças com a economia de regiões como o Caribe e o Sul dos Estados Unidos da América e, finalmente, a relevância do papel exercido pelo mercado interno, mais visível quando se consideram a conquista e a colonização dos sertões (Santos, 2010).

Palavras-chave: *Continuum* de normas X polarização sociolinguística; Português colonial; Configuração da estrutura econômica, social e familiar do Brasil Colônia.

http://doi.org/10.5007/1984-8420.2025.e108135

Abstract: The analysis of the inquisitorial sources bequeathed by the first two visitations of the Holy Office to Portuguese America (Lobo; Sartori, 2020; Lobo; Sartori; Mota, 2016; Siqueira, 1978) does not support the perspective of interpretation of the socioeconomic stratification of colonial Brazil in dichotomous terms –"casa-grande" *versus* "senzala" (Freyre, [1933] 2002) –, which has implications for the interpretation of the process of formation of BP and for understanding the current sociolinguistic configuration of the country. In addition to this point, another aspect that is rarely discussed is highlighted: the need to understand linguistic interactions within the different configuration patterns of the colonial family. The so-called sociolinguistic polarization (Lucchesi, 1994, 2015) of BP, so clear when contrasting the socioeconomic elite with the enslaved population, becomes a tangled web, if nuances that linguistic and cultural reality are put into analysis in the social and domestic sphere it sheltered. Based on Pagotto (2018), we propose to deepen the reflection on what interpretations engendered from the 1970s on the colonial economy and society reveal about the history of BP, highlighting the following points: the fact that the colonial economy cannot be understood through the reductionist vision of "cycles"; the fact that it was not limited to the agro-mineral-export model; the fact that it does not bear many similarities to the economy of regions such as the Caribbean and the South of the United States of America and, finally, the relevance of the role played by the internal market, more visible when considering the conquest and the colonization of the backlands (Santos, 2010).

Keywords: *Continuum* of norms X sociolinguistic polarization; Colonial Portuguese; Configuration of the economic, social and family structure of Colonial Brazil.

1. Introdução

A coleção *História do português brasileiro* – integrada por 12 volumes publicados entre 2018 e 2022[1] e cuja coordenação geral foi levada a cabo por Ataliba Teixeira de Castilho, especialmente homenageado neste número temático da *Working Papers em*

[1] Vol. 1: Ataliba T. de Castilho (2018). *O português brasileiro em seu contexto histórico*; Vol. 2: Ataliba T. de Castilho (2019). *Corpus diacrônico do português brasileiro*; Vol. 3: Dermeval da Hora, Elisa Battisti e Valéria de Oliveira Monaretto (2019). *Mudança fônica do português brasileiro*; Vol. 4: Célia Regina Lopes (2018). *Mudança sintática das classes de palavra*: perspectiva funcionalista; Vol. 5: Ataliba T. de Castilho (2019). *Mudança sintática das construções*: perspectiva funcionalista; Vol. 6 (2018): Sônia Cyrino e Maria Aparecida Torres Morais. *Mudança sintática do português brasileiro*: perspectiva gerativista; Vol. 7: Maria Lúcia C.V.O. Andrade e Valéria Severina Gomes (2018). *Tradições discursivas do português brasileiro*: constituição e mudança dos gêneros discursivos; Vol. 8: Rodolfo Ilari e Renato Basso (2020). *História semântica do português brasileiro*; Vol. 9: Dinah Callou e Tânia Lobo (2020). *Da história social à história linguística*; Vol. 10: Jânia M. Ramos e Marilza de Oliveira (2021). *Dialetação e povoamento*: da história linguística à história social; Vol. 11: Eduardo Penhavel e Marcos Rogério Cintra (2022). *Diacronia dos processos de construção de textos*. Vol. 12: Vanderci Aguilera e Fabiane Cristina Altino. *Léxico histórico do português brasileiro.* Enquanto os onze primeiros volumes foram publicados em formato impresso, o de número 12 virá à luz no formato eletrônico.

Linguística – consolida os resultados de pesquisa daquela que se pode considerar a primeira e longa fase do projeto nacional *Para a história do português brasileiro* (PHPB), fundado, também por iniciativa de Ataliba Teixeira de Castilho, no ano de 1997.

Aproximando-se já de quase três décadas de existência, é o PHPB um projeto de importância fundamental na Linguística brasileira, como também o foi outro projeto de âmbito nacional proposto dez anos antes, em 1987, também por Ataliba Teixeira de Castilho, o da *Gramática do português falado*, o qual, tendo procedido a uma análise de caráter global do *corpus* de língua falada constituído no âmbito do projeto de estudo da *Norma linguística urbana culta* (NURC), nos legou a coleção *Gramática do português culto falado no Brasil*, conferindo, assim, ao português brasileiro o estatuto de ser a primeira língua românica a ter a sua variedade linguística socialmente mais prestigiada amplamente descrita.

Uma análise de caráter preliminar da extensa produção bibliográfica gerada pelo PHPB ressalta, ao menos, três aspectos que demandam a nossa atenção, tendo em vista o planejamento da sua história futura:

1º) A notável dissociação entre a pesquisa produzida no âmbito do PHPB e o ensino de história do português brasileiro nos cursos de graduação do país. Nem em seus textos fundadores, nem nos textos apresentados e discutidos nos Seminários já realizados, nem ainda na grande coleção *História do português brasileiro*, a questão da relação entre pesquisa e ensino de Graduação foi problematizada. Analisando as grades curriculares, as ementas e os componentes curriculares dos cursos de Letras de 17 universidades públicas (das esferas federal e estadual), com equipes de pesquisadores vinculados ao PHPB, Lobo, Sartori e Souza (2021) observaram que apenas 04 delas possuíam um componente curricular específico para tratar da história do português brasileiro, com diferentes graus de aprofundamento das questões relativas à história social da língua e à mudança lexical ou gramatical. Na esmagadora maioria dos casos, o que se prioriza é a história da formação da língua no espaço europeu, do latim ao português arcaico, apresentando-se apenas alguns tópicos sobre o português do ou no Brasil.

2º) A concentração das pesquisas fundamentalmente no século XIX, avançando, a partir daí, para o século XX, sem ainda ter havido o necessário e aprofundado recuo à história linguística do Brasil e à história da língua portuguesa no e do Brasil nos três séculos de conquista e colonização da chamada América portuguesa. A abordagem do período colonial nos remete à consideração de uma separação necessária, mas geralmente pouco feita, entre o Estado do Brasil e o Estado do Grão-Pará e Maranhão, as duas colônias

portuguesas na América, que se distinguiam tanto do ponto de vista político-administrativo e socioeconômico, como também tiveram histórias linguísticas diferenciadas (cf. Bessa Freire, 2004). Considerando não só as dimensões continentais do atual território brasileiro, mas até mesmo de cada uma das duas colônias separadamente, reafirmamos a necessidade de que, para a produção de sínteses em perspectiva macroscópica, haja investimento em investigação de histórias linguísticas regionais e de micro-histórias linguísticas. Foi, assim, com este objetivo, correlacionado à projeção de uma segunda fase de pesquisas do PHPB[2], que se elaborou o projeto *História do português brasileiro colonial* (2024), o qual tratará,

> [...] sobretudo, mas não exclusivamente, do Estado do Brasil e, de forma mais verticalizada, das Capitanias da Bahia, de Ilhéus e de Porto Seguro, espaços reestruturados na segunda metade do século XVIII e cuja área corresponde aproximadamente ao que passou a ser a Província da Bahia e, por fim, o Estado da Bahia (p. 01)

3º) Por fim, chama também a atenção a quase ausência de debate quer sobre as hipóteses explicativas sobre a formação do português brasileiro – *grosso modo* referidas como a *hipótese da deriva* (Naro; Scherre, 1993, 2007) e a(s) *hipótese(s) do contato*[3]-, quer sobre as propostas de caracterização sociolinguística do português brasileiro contemporâneo – a proposta dos *continua de normas* (Bortoni-Ricardo, [1998] 2002) e a da *polarização sociolinguística* (Lucchesi, 1994, 2015) – e sobre como cada uma das duas referidas propostas dialogaria com o passado sócio-histórico-linguístico do Brasil.

Neste texto, com o objetivo de contribuir para uma história social do português colonial e partindo do presente em direção ao passado, discutiremos, inicialmente, na seção 2, as duas propostas de caracterização sociolinguística do português brasileiro contemporâneo – o *continuum* de normas *versus* a polarização sociolinguística –, demonstrando, com a reinterpretação de conclusões referentes à análise da avaliação social sobre a variação na concordância verbal junto à 3ª pessoa do plural, ser a proposta dos *continua* de normas a mais adequada. Nas seções seguintes, buscaremos compreender o encaixamento

[2] O projeto *História do português brasileiro colonial*, cuja vigência se estende de 01.01.2024 a 31.12.2029, está sendo desenvolvido por 17 docentes pesquisadores vinculados à UFBA, UEFS, UNEB, UnB e UNIT.

[3] A questão do contato na história linguística brasileira tem sido objeto de diversas interpretações, dentre as quais se podem mencionar a hipótese da crioulização prévia (Guy, 1989) e a da transmissão linguística irregular (Baxter; Lucchesi, 1997 e Lucchesi, 2008). Assumido o contato como central, investiga-se se as mudanças mais significativas na gramática do PB teriam sido desencadeadas pela ação de "princípios gerais" que regeriam os processos de aquisição de L2 e produziriam inovações gramaticais que são passadas para variedades L1 ou se teriam decorrido da transferência de propriedades gramaticais de línguas africanas para as novas variedades do português (Álvarez-López; Gonçalves; Avelar, 2018).

histórico desse *continuum* de normas: na terceira, confrontaremos modelos explicativos distintos sobre a formação da economia colonial, destacando que o processo de conquista e colonização do sertão da Bahia, diferentemente do que ocorreu no Recôncavo, ativou principalmente um mercado interno, ancorado na atividade pecuária e na economia de subsistência; na quarta, refletiremos sobre a configuração social do Brasil colonial, com foco no Nordeste, relacionando estratificação social, mestiçagem e estrutura familiar à complexa configuração linguística do período.

2. *Continuum* de normas *versus* polarização sociolinguística

Três macro-parâmetros sócio-históricos devem ser tidos em conta em quaisquer proposições que visem à caracterização sociolinguística do português brasileiro contemporâneo. Partindo do fato incontornável de que o português brasileiro emergiu em contexto multilíngue, sendo, portanto, o contato linguístico elemento *constitutivo* da sua formação, o primeiro parâmetro é o que diz respeito às "sucessivas distribuições demográfico-linguísticas" (Houaiss, [1985] 1992), que levaram a que o multilinguismo generalizado do passado fosse cedendo lugar ao multilinguismo localizado do presente (Mattos e Silva, 1993), processo em cujo âmbito o português brasileiro não só se vai constituindo, como também se tornando língua hegemônica. O segundo parâmetro diz respeito ao crescimento populacional associado às migrações internas, à urbanização e à industrialização, opondo o Brasil colonial e imperial eminentemente rural-agrário ao Brasil eminentemente urbano-industrial da segunda metade do século XX. Por fim, o processo de difusão social da leitura e da escrita e o processo de escolarização associados ambos ao processo de estandardização da língua, opondo um Brasil colonial e imperial marcado por altos índices de analfabetismo a um Brasil crescentemente industrializado e com índices também crescentes de escolarização-letramento.

Ao estudar o nascimento do dialeto de Brazlândia, cidade-satélite de Brasília, tornada, em 1960, a nova capital do Brasil, Bortoni-Ricardo ([1985] 2011) pôs em evidência o segundo dos três parâmetros referidos, afirmando que o aspecto fundamental para a compreensão do português brasileiro contemporâneo seria o estudo da transformação das variedades linguísticas rurais, segundo ela, caracterizadas por surpreendente alto grau de uniformidade, em variedades urbanas não padrão, por ela designadas de variedades *rurbanas*, as quais, no *continuum* dialetal, estariam adjacentes às variedades rurais isoladas e altamente estigmatizadas, e seriam faladas não só nas cidades, por indivíduos não alfabetizados ou semialfabetizados de classes mais baixas e, em sua maioria, com antecedentes rurais, mas também nas áreas rurais cujos falantes se encontrassem sob a influência tecnológica e modernizadora emanada das cidades. Propôs Bortoni-Ricardo ([1998] 2002) um modelo constituído por três *continua*, em que, para além do já mencionado *continuum*

rural-urbano, destinado à análise dos atributos sociológicos do falante (variedades rurais isoladas – variedades rurbanas – variedades urbanas padronizadas), haveria o *continuum* oralidade-letramento, referente à análise das práticas sociais letradas em que o indivíduo toma parte (eventos de oralidade – eventos de letramento), e o *continuum* de monitoração estilística, atinente à análise dos processos psicológicos de planejamento e atenção no momento da enunciação (- monitoração – + monitoração).

É Lucchesi (2015) quem vai aprofundar a análise do *continuum* rural-urbano. Além de aprofundá-la, refina-a, ao considerar, em toda a sua reflexão, o primeiro macro-parâmetro sócio-histórico acima referido – o contato linguístico em contexto colonial e imperial escravagista – como um aspecto fundante da história social linguística brasileira. Contudo, a despeito de manter a posição de defesa da tese da polarização sociolinguística do Brasil, cuja primeira formulação data de 1994, a sua proposta de estratificação sociolinguística do Brasil contemporâneo, calcada na análise da variação na concordância verbal junto à 3ª pessoa do plural (fenômeno considerado o "cerne" da polarização), vem, paradoxalmente, corroborar – quer no plano empírico, quer no plano das hipóteses que formula – a tese dos *continua* de normas.

Passamos, a seguir, à apresentação dos resultados da referida análise de Lucchesi (2015), feita a partir da observação controlada de três aspectos dos padrões coletivos de comportamento linguístico: (i) a frequência de uso das variantes linguísticas; (ii) a avaliação social das variantes linguísticas e (iii) as tendências dos processos de variação e mudança[4].

Relativamente à frequência de uso das variantes linguísticas, os estudos são abundantes, não havendo carência de dados empíricos. Assim, na Tabela 1, abaixo, podemos ver a compilação feita por Lucchesi (2015) dos resultados de estudos[5] sobre a variação na concordância verbal junto à 3ª pessoa do plural:

[4] Neste texto, apenas serão abordados os dois primeiros aspectos.

[5] **Português urbano culto**: Diva Graciosa (1991). *Concordância verbal na fala culta carioca*: 18 entrevistas do tipo DID do acervo do Projeto NURC-RJ; **português urbano médio**: Marta Scherre e Anthony Naro (1997). A concordância de número no português do Brasil: um caso típico de variação inerente: amostra com base no *corpus* do PEUL, Isabel Monguilhott (2001). *Variação na concordância verbal de terceira pessoa do plural na fala dos florianopolitanos:* amostra com base no *corpus* do VARSUL de Florianópolis; **português popular urbano**: Anthony Naro (1981). The social and structural dimensions of a syntactic change; Gregory Guy (1981). *Linguistic variation in Brazilian Portuguese: aspects of phonology, syntax and language history*; amostra com base no *corpus* do Projeto Competências, com 20 falantes adultos, alunos do MOBRAL; ***português rurbano***: Stella-Maris Bortoni-Ricardo ([1985] 2011). *Do campo para a cidade*: estudo sociolinguístico de migração e redes sociais: amostra de 33 entrevistas com migrantes rurais de uma cidade-satélite de Brasília, com pouca ou nenhuma escolaridade; **português popular rurbano**: Sílvia Vieira (1995). A não concordância em dialetos populares: uma regra variável: amostra com base no *corpus* do acervo do Projeto APERJ (*Atlas Etnolinguístico do Estado do Rio de Janeiro*), com 72 inquéritos com informantes homens, pescadores, com pouca ou nenhuma escolaridade; **português popular rural**: amostra do português popular no município de Santo Antônio de Jesus, na zona do Recôncavo Baiano, com 24 falantes dos dois sexos, de pouca ou nenhuma escolaridade, sendo 12 moradores da sede do município e 12 da zona rural, distribuídos em três faixas etárias. Remete-se a Lucchesi (2015) para as referências completas.

Tabela 1: Frequência geral de aplicação da regra de concordância verbal em diversas variedades do português brasileiro

Variedade do português do Brasil	**Frequência geral de aplicação da regra de concordância verbal**
Português urbano culto (RJ)	94%
Português urbano médio (RJ e SC)	73% (RJ) e 79% (SC)
Português popular urbano (RJ)	48%
Português popular rurbano (RJ)	38%
Português popular rurbano (Brasília)	35%
Português popular rurbano (BA)	25%
Português popular rural (BA)	17%

Fonte: Lucchesi (2015, p. 251).

Observando, quanto aos resultados referentes ao Rio de Janeiro, uma diferença de aproximadamente 20 pontos percentuais entre os falantes com nível de escolaridade superior completo, os com escolaridade de 1 a 11 anos e os analfabetos ou semialfabetizados, aponta Lucchesi para o estabelecimento de três normas distintas – a *norma urbana culta*, a *norma urbana média* e a *norma popular urbana*. Tal proposta não só é corroborada, mas avança, quando desmembra os falantes que representariam o português urbano médio do Rio de Janeiro, com escolaridade de 1 a 11 anos, em três grupos distintos e observa 63% de frequência de aplicação da regra para os que têm entre 1 e 4 anos de escolaridade, 78% para os que têm entre 5 e 9 anos de escolaridade e, finalmente, 81% para falantes com 9 a 11 anos de escolaridade. Propõe, então, que os falantes com 63% de aplicação da regra representariam uma *norma urbana média baixa*, enquanto os falantes dos outros dois grupos – com percentuais muito próximos (78% e 81%) – representariam a *norma urbana média*.

Conforme suas próprias palavras, "prosseguindo no *continuum* sociolinguístico" (p. 252), divisa ainda o autor a *norma popular rurbana* e a *norma popular rural*. No que tange à *norma popular rurbana* – com frequências de aplicação da regra que opõem, por um lado, os 25% relativos à Bahia e, por outro, os 35% e 38%, correspondentes, respectivamente, a uma cidade satélite de Brasília e ao Rio de Janeiro, propõe considerar a frequência de 35% como valor de referência. Por fim, para a *norma popular rural*, o valor de referência assumido é a frequência de 20% de aplicação da regra de concordância, fundamentando-se nos 17% atestados em área rural da Bahia. Na conclusão da análise sobre a frequência de uso, faz um leve ajuste nos percentuais entre as normas linguísticas, fixando um intervalo constante de 15% entre elas, o que torna ainda mais evidente a gradualidade do contínuo (cf. Tabela 2, abaixo).

Se dados empíricos sistemáticos e representativos não constituíram nenhum óbice à análise da frequência de uso das variantes linguísticas, o mesmo já não se pode afirmar quanto à análise da avaliação subjetiva, restando, portanto, ao autor a formulação de uma hipótese, que, todavia, carece ainda de comprovação. Na Tabela 2, apresentamos os dados referentes à frequência geral de aplicação da regra de concordância verbal, conjugados a uma síntese que fizemos das conclusões a que chega Lucchesi (p. 253-254) ao cabo da discussão sobre a sua hipótese:

Tabela 2: Frequência geral de aplicação da regra de concordância verbal nas normas sociolinguísticas X Hipótese sobre avaliação social das variantes

Norma sociolinguística	**Frequência geral de aplicação da regra**	**Hipótese sobre a avaliação social das variantes**
Norma urbana culta (falantes com nível superior de escolaridade)	95%	**Estereótipo** Avaliação explicitamente negativa da não aplicação da regra de concordância
Norma urbana média (falantes com 5 a 11 anos de escolaridade)	80%	**Norma de transição** Oscilação entre a situação de marcador e estereótipo
Norma urbana média baixa (falantes com 1 a 4 anos de escolaridade)	65%	**Norma de transição** Oscilação entre a situação de marcador e estereótipo
Norma urbana popular (falantes analfabetos)	50%	**Marcador** Ausência de avaliação explicitamente negativa da não aplicação da regra de concordância, apesar da consciência do prestígio social decorrente da sua aplicação
Norma rurbana popular	35%	**Norma de transição** Oscilação entre a situação de indicador e marcador
Norma rural popular	20%	**Indicador** Ausência de consciência sobre o prestígio social decorrente da aplicação da regra

Elaborada pelos autores.

Apesar de longo, reproduzimos, abaixo, o trecho que sintetiza a posição do autor após a análise dos dois parâmetros (p. 254-255):

> Confrontando as frequências de uso da regra de concordância com os sistemas de avaliação social do seu emprego, pode-se constatar que, enquanto o primeiro parâmetro aponta no sentido de um *continuum*, o segundo revela a polarização sociolinguística. No que concerne à frequência de uso, as normas mantêm entre si um intervalo regular de 15 pontos percentuais, ou seja, exibem uma gradação sem solução de continuidade. Porém, no plano da avaliação subjetiva, postula-se uma diferença qualitativa. Na norma culta, a ausência de concordância recebe uma avaliação negativa explícita; o mesmo não se pode dizer da norma

> popular, na qual não se observa essa avaliação negativa, podendo ocorrer no máximo uma avaliação positiva da regra. Entre esses contrários, figurariam as normas médias, cuja definição ainda carece de novas investigações empíricas. Já no polo da norma popular, uma nova distinção qualitativa poderia ser proposta, em face da ausência de avaliação subjetiva na norma popular rural.
>
> Conjugando os dois parâmetros até aqui observados, pode-se postular uma relação dialética entre uso e avaliação. A avaliação positiva da regra de concordância se basearia na alta frequência de uso dessa regra por parte de um grupo social, que é minoritário, mas dominante no plano político e social [...], em contraste com os demais grupos sociais, particularmente aqueles mais marginalizados na estrutura socioeconômica, nos quais a regra praticamente não é usada. Fica, assim, desnudado o caráter ideológico da avaliação positiva da regra de concordância e, *a fortiori*, do preconceito contra o padrão de não concordância, predominante na fala popular, pois tanto a avaliação positiva quanto o preconceito nada mais são do que o reflexo da distribuição sociolinguística da variável (o que revela sua motivação socioeconômica), e não decorrente de algum valor linguístico intrínseco da regra, o que efetivamente não existe.

A hipótese formulada é plausível, porém o que **inequivocamente** observamos, a partir da síntese feita, é que também em relação à avaliação social das variantes há "uma gradação sem solução de continuidade" no que diz respeito à "consciência" dos falantes de cada uma das normas linguísticas sobre o valor social da aplicação da regra de concordância verbal (passa-se "gradualmente" do polo de ausência de consciência para o polo em que a consciência seria plenamente atestada); observemos, ainda, que a própria proposição da existência de "normas de transição" vem corroborar essa interpretação.

Por fim, resta observar que Lucchesi (2015) postula um *continuum* de variedades potenciais da norma popular brasileira no eixo rural-urbano, que seria assim constituído: *português popular urbano das grandes cidades – português popular rurbano das grandes cidades – popular urbano do interior – português popular rurbano do interior – português popular rural – português afro-brasileiro*[6]. Tal postulação – também plausível – não encontraria, conforme Lucchesi (2015), paralelo na norma culta, fazendo-nos supor ser ela bastante mais homogênea do que talvez de fato o seja e fazendo-nos supor ainda que os seus falantes apresentariam o mesmo nível de consciência quanto à não aplicação da re-

[6] Observe-se que o contínuo proposto não contempla variedades linguísticas resultantes do contato entre português e línguas indígenas. A Linguística Histórica, no Brasil, tem dado protagonismo ao contato do português com línguas africanas – o que é inquestionável – e negligenciado o contato com as línguas indígenas. Isto talvez revele o quanto ainda estamos distantes de uma renovada historiografia da América indígena, que, há aproximadamente 30 anos, reage a tratar, conforme Monteiro (2001, p. 4), "a história dos índios no Brasil [como se resumindo] basicamente à crônica de sua extinção"

gra de concordância. Contudo, tal ideia deve, pelo menos, ser matizada, como podemos comprovar com a leitura do texto abaixo, publicado em rede social por uma mulher de 37 anos, graduada em Letras e estudante de Mestrado[7]:

> "TÁ ME ESCUTANDO VÓ?" Eu sempre te perguntava e mesmo com toda dificuldade de ouvir a senhora se esforçava e sempre me respondia "ESTOU MINHA FILHA!" Espero que esteja me escutando também agora. EU TE AMO VÓ! Feliz dia das avós. A saudade da senhora me sufoca as vezes, mas sigo bem como a senhora me deixou. "TÁ ME ESCUTANDO VÓ?" Obrigado por ter estado comigo em todos os momentos. Ainda consigo sentir suas mãos que passava [**sic**] brilhantina em meu cabelo fazendo as tranças nagô, ainda consigo sentir suas mãos que encostava [**sic**] em meus lábios me ensinando a pedir a sua benção. Ainda consigo sentir as suas mãos que testava [**sic**] minha temperatura em um dia de febre, acarinhava [**sic**] minha barriga durante minhas gestações, e que entregou [**sic**] o canudo da minha graduação. Sinto essas mãos tão enrugadas que me confortaram na minha separação, que alisou [**sic**] minhas costas ao vê [**sic**] minha primeira casa pronta, que apertou [**sic**] com emoção a minha mão passando a sua força e apoio depois de vê [**sic**] minha mãe oxum nascer em mim. Mãos que nos últimos momentos alisou [**sic**] meu ori pq minha carequinha era gostosa de pegar. Você esteve em todos os momentos. Todos! E agora como minha ancestral peço que essas mãos sigam presente, lindas como só elas eram, com suas unhas sempre pintadas em variados tons de rosa e seus anéis, no mínimo dois ou três, não mais me tocando, mas como setas, mostrando o melhor caminho a seguir. TÁ ME ESCUTANDO VÓ? EU TE AMO!

Podemos situar nas reflexões de Silva Neto ([1950] 1986) sobre o que denominou de "dualidade linguística" os antecedentes históricos da proposta de polarização sociolinguística desenvolvida por Lucchesi: "(...) dos princípios da colonização até 1808, e daí por diante com intensidade cada vez maior, se notava a dualidade linguística entre a nata social, viveiro de brancos e mestiços que ascenderam, e a plebe, descendente dos índios, negros e mestiços da colônia." (p. 80) Além disso, é também relevante observarmos que Mattos e Silva (2000a) assumiu a ideia da polarização e buscou, assim como também o fez Lucchesi, identificar as "raízes" das duas supostas grandes vertentes da formação histórica da realidade sociolinguística brasileira: a norma culta – supostamente mais unitária e proveniente do português falado pelos colonos portugueses e seus descendentes nasci-

[7] A citação deste texto seguiu as normas estabelecidas pelas *Diretrizes para a ética na pesquisa e a integridade científica* (2024), elaboradas pelo "Fórum de Ciências Humanas, Sociais, Sociais Aplicadas, Linguística, Letras e Artes" (FCHSSALLA), tendo a autora dado autorização expressa para a sua divulgação e análise. Disponível em: https://www.cgee.org.br/documents/10195/11009772/CGEE_FCHSSALLA_diret_etica_pesq_integ_cient.pdf. Acesso em 30 de agosto de 2024.

dos no Brasil, além de moldada sob o influxo do padrão normativo de escrita difundido via escolarização – e a norma popular – supostamente mais diversificada e supostamente descendente direta do português falado como segunda língua por índios e negros, os quais, juntamente com seus descendentes, do século XVI a meados do XIX, sempre constituíram o segmento maioritário da população brasileira e até aproximadamente meados do século XX estiveram, em sua maioria, à margem do sistema formal de escolarização. Ao fazê-lo, Mattos e Silva e Lucchesi embasaram-se em um modelo que não dá conta da interpretação da sociedade e da economia colonial em sua complexidade, o mesmo em que, aliás, *mutatis mutandis*, também se apoiou Silva Neto. Se a análise dicotômica teve o mérito de nos fazer começar a enfrentar a complexidade histórica da questão, paradoxalmente, foi também responsável por também nos desviar dessa mesma complexidade.

Nas duas seções seguintes deste texto, com base na premissa do princípio do uniformitarismo da Sociolinguística ([1972] 2008), o objetivo será o de "começar" a compreender o encaixamento histórico da questão, buscando desvendar quais seriam as raízes econômicas e sociais não das "duas" grandes vertentes da formação histórica da realidade sociolinguística brasileira, mas, sim, do complexo contínuo de normas sociolinguísticas observadas no presente.

3. Economia e ocupação territorial

Nesta seção, trataremos de modelos explicativos sobre a formação da economia colonial e da sua relevância para a compreensão da história social linguística brasileira. A seguir, o foco da análise recairá sobre a região Nordeste, abordando, mais verticalmente, o processo de conquista e colonização do sertão da Bahia, alicerçado também na mineração, mas, principalmente, na expansão da atividade pecuária, na preação de indígenas, na economia de subsistência e na constituição de um mercado interno. Adicionalmente, destacaremos as guerras e alianças travadas com as populações indígenas predominantemente nômades e seminômades que habitavam essa área, com o objetivo de destacar a relevância do contato entre o português e suas línguas, aspecto ainda pouco estudado e geralmente considerado de somenos importância, comparativamente ao destaque dado ao papel das línguas africanas na formação do português brasileiro[8].

[8] Uma exceção notável é a já longa tradição de pesquisa da UEFS sobre a história linguística do sertão, conforme se pode comprovar, dentre muitos outros, pelos seguintes trabalhos: CARNEIRO, Zenaide; LACERDA, Mariana; ALMEIDA, Norma. Indícios sobre a participação dos povos tapuias no contato com o português nos sertões baianos seiscentistas. *In:* CALLOU, Dinah; LOBO, Tânia. *História social do português brasileiro*: da história social à história linguística. Coleção *História do português brasileiro*, v. 9, Coordenação Geral de Ataliba T. de Castilho. São Paulo: Contexto, 2020. p. 98-125. CARNEIRO, Zenaide; LACERDA, Mariana. Projeto *Corpus* eletrônico de documentos históricos do sertão: a documentação epistolar. *In*: PIMENTA, Ricardo M.; ALVES, Daniel. *Humanidades digitais e o mundo lusófono*. Rio de Janeiro: Editora FGV, 2021. p. 215-233.

3.1 Breve olhar sobre modelos explicativos da economia colonial

Abordar a história social linguística do Brasil e, em especial, o processo de constituição histórica do português brasileiro analisando o modo como se equaciona a relação entre economia e funcionamento sociolinguístico – em outras palavras, abordar os efeitos sociolinguísticos dos processos econômicos – tem sido algo relativamente incomum entre nós, linguistas historiadores do português brasileiro, e, exatamente por isso, chama-nos a atenção a publicação, no ano de 2018, de um breve, porém instigante texto de Emilio Pagotto, intitulado "Ebulição e sedentarização linguística: o lugar da economia de subsistência na formação do português brasileiro".

Inicialmente, adverte-nos o autor para o fato de poder tal relação ser considerada em duas distintas perspectivas. A primeira, de caráter mais amplo, analisaria as relações sociais e linguísticas em sociedades comunitárias, feudais, capitalistas pré-industriais, industriais ou pós-industriais etc., tendo como foco verificar se mudanças no modo de produção seriam correlacionáveis a mudanças nas relações sociolinguísticas. A segunda, de caráter mais específico, estaria dedicada a desvelar se e como, no âmbito de um mesmo modo de produção, rupturas econômicas engendradas no seio de sociedades particulares ecoariam nos processos de variação e mudança linguística.

É a partir da segunda perspectiva que Pagotto (2018) vai tratar da formação do português do Brasil, trazendo à luz, inicialmente, um esclarecimento indispensável, mas, até então, não observado em nosso campo de estudos: o de que a história econômica da sociedade colonial brasileira é alvo de duas interpretações que se chocam[9], sendo a primeira tributária ao pensamento de Caio Prado Jr. em *Formação do Brasil contemporâneo*, obra de 1946, pensamento este que se tornou hegemônico ao longo de toda segunda metade do século XX, e a segunda, derivada das pesquisas conduzidas por João Fragoso e Manolo Florentino, que se apresentam na obra *O arcaísmo como projeto: mercado atlântico, sociedade agrária e elite mercantil em uma economia colonial tardia*, cuja primeira edição data de 1993. Ainda tributária da primeira interpretação temos *Formação econômica do Brasil*, obra de Celso Furtado, publicada em 1958.

Caio Prado Jr., em *Formação do Brasil contemporâneo*, e Celso Furtado, em *Formação econômica do Brasil*, defendiam que "a sociedade da América lusa dos séculos XVII e XVIII fora construída com o propósito de fomentar a transição do feudalismo para o capitalismo na Europa, ou ainda com o intuito de viabilizar a revolução industrial inglesa

[9] No âmbito da historiografia brasileira e brasilianista, a questão dos modelos interpretativos da economia colonial não se restringe a "duas" interpretações. Para uma visão da complexidade do tema, remetemos a ZOREK (2021).

do século XIX" (Fragoso, 2012, p. 107). Nesses termos, como resultado do capitalismo comercial europeu, o que teríamos, na passagem do século XVI para o século XVII, na América portuguesa, seria um "grande canavial" gerenciado por senhores de engenho, ficando, a economia colonial à mercê do mercado europeu e, portanto, sem uma dinâmica própria. A inexistência de um mercado interno ou ainda de produções mercantis para abastecimento da própria Colônia seria uma consequência desse modelo interpretativo da economia colonial.

O embate historiográfico sobre os antigos modelos explicativos da chamada economia colonial dominante, que se assentavam, sobretudo, nos "sucessivos ciclos" do mercado agro-minério-exportador e na relação senhores *versus* escravizados, emergiu nas décadas de 1970 e 1980[10], revelando "a fragilidades de várias certezas – em verdade, hipóteses – da tradicional historiografia colonial" (Fragoso, 2012, p. 210). Abriram-se, portanto, caminhos para outras interpretações, na medida em que se passou a demonstrar que

> [...] a economia colonial era mais do que uma *plantation* exportadora, existia um circuito de mercados internos disseminados pela América. Mesmo nas regiões até então vistas como açucareiras, como o Recôncavo Baiano, observou-se a existência de áreas dedicadas à lavoura mercantil de alimentos (Fragoso, 2012, p. 110).

> [...] no Recôncavo Baiano, ao lado da atividade açucareira, havia os pequenos e médios produtores de alimentos que, utilizando o trabalho familiar e/ou escravo, abasteciam os engenhos e as cidades. Algo semelhante ocorria no Rio de Janeiro do século XVII. Além da produção de mantimentos nas próprias áreas açucareiras, ao longo do século XVII e no seguinte, respondendo à demanda dessas áreas, **os agricultores** de São Paulo, **do sul da Bahia**, do Espírito Santo e, posteriormente, os do Maranhão **passaram a orientar as suas atividades para um comércio intercapitanias, esboço de um incipiente mercado interno colonial**[11] (Fragoso; Florentino; Faria, 1998, p. 49).

Na segunda citação acima, destacamos o trecho que faz referência ao sul da Bahia, *grosso modo* correspondente às antigas Capitanias de Ilhéus e Porto Seguro. Tendo analisado especificamente a importância da Capitania de Porto Seguro no amplo mercado interno de abastecimento de víveres, Cancela (2012) destacou o objetivo da Coroa portuguesa de destinar o grosso da sua produção econômica ao mercado de Salvador e do Rio

[10] Os debates e as revisões historiográficas resultantes foram consequência da multiplicação dos programas de Pós-Graduação em História no Brasil, em cujo âmbito passaram a ser escritas dissertações e teses fundamentadas em sólidas pesquisas empíricas.

[11] Grifo nosso.

de Janeiro, o que exigiu a criação de uma infraestrutura de transporte e comunicação para garantir o escoamento dos produtos. Observou ainda que, estando o sistema produtivo da referida capitania conjugado às demandas dos grandes centros urbanos da colônia, não perdeu Porto Seguro "a oportunidade de produzir e comercializar mais de uma dezena de gêneros agrícolas, extrativistas e manufatureiros" (Cancela, 2012, p. 259), tendo desenvolvido, por exemplo, o cultivo de mandioca, milho, algodão, cana, arroz e feijão, a pesca da garoupa, a extração de madeira de lei e tabuado, assim como a manufatura e/ou semimanufatura de farinha, aguardente, açúcar, embira, cordoaria e embarcações.

Cabe, neste ponto, um parêntese para destacar a proposta de periodização da história linguística do Sul da Bahia formulada por Argolo (2015), segundo a qual é só no período compreendido entre 1820 e 1872, quando se observa a dizimação mais acentuada da população indígena e quando se consolida o processo de migração sertaneja, deflagrado pela atração da lavoura do cacau, que o multilinguismo indígena da região "perde a batalha" para o unilinguismo em português brasileiro.

Ainda segundo o mesmo autor (2015, p. 293):

> [...] em 1760, já é possível encontrar registros documentais de invasões de grandes extensões territoriais na Capitania de Ilhéus, possivelmente para plantar cacau. Desse período em diante, o movimento migratório ganhou força, culminando na disputa pela posse da terra entre os índios, mamelucos e brancos pobres do Sul da Bahia e os sertanejos que migraram para lá, no intuito de obter melhores condições de vida com o plantio do cacau. Sendo os índios, mamelucos e brancos pobres do Sul da Bahia os depositários do multilinguismo da região, e os sertanejos, neste caso específico, os depositários do unilinguismo em português brasileiro, iniciou-se, em termos metafóricos, uma batalha entre o multilinguismo e o unilinguismo na costa sul baiana.

O caso da Capitania de Porto Seguro, bem como o do sertão da Bahia, que será abordado no item seguinte, evidencia, conforme sintetizou Pagotto (2018, p. 56), que, no núcleo do debate historiográfico sobre modelos interpretativos da economia colonial, está o papel do mercado interno na história econômica da Colônia:

> [...] o quanto a economia colonial encontra sua lógica de crescimento apenas voltada para a exportação, com a expropriação dos lucros pela metrópole *versus* o quanto a constituição de um mercado interno propiciou a acumulação de capitais na própria colônia, o que inverteria a lógica do mercantilismo de colonização.

O modelo dos sucessivos ciclos econômicos, conjugado à exclusividade do mercado externo, "cria a falsa ideia de que o artigo anterior deixou de ser produzido" (Fragoso; Florentino; Faria, 1998, p. 53). O caso do "açúcar" é emblemático: apesar de ter sofrido drástica queda de exportação pela queda do preço no mercado internacional – o que teria encerrado o seu "ciclo" –, não sofreu queda de produção, já que foi progressivamente atendendo também ao sempre crescente mercado interno.

Ademais, não podemos perder de vista que a compreensão da importância do mercado interno se relaciona diretamente com o processo de ocupação territorial dos chamados sertões e ampliação das fronteiras. Assim, passaremos, a seguir, a refletir sobre a questão, tomando como referência, sobretudo, a Capitania da Bahia.

3.2 Caminhos da ocupação dos sertões da Bahia e da difusão do português

Ao propor uma reconstrução da história social linguística do Brasil como um dos campos de pesquisa para enfrentar a questão histórica do português brasileiro, Mattos e Silva (1998) destaca que dois percursos investigativos devem ser trilhados: (i) a recuperação da articulação entre os fatos de ocupação territorial, as sucessivas distribuições demográfico-linguísticas e as prevalências e desaparecimento das línguas; e (ii) a reconstrução da história da escolarização no Brasil, mais amplamente, a difusão social da leitura e da escrita. Para uma melhor compreensão sobre o percurso investigativo (i), podemos considerar que Santos (2010), em sua tese intitulada *Fronteiras do sertão baiano: 1640-1750*, embora centrando sua atenção de meados do século XVII até a primeira metade do século XVIII, fornece-nos, com base em ampla pesquisa arquivística, um significativo panorama sobre o processo de ocupação do sertão baiano.

No processo de ocupação da América portuguesa, o autor destaca dois movimentos distintos, nos quais se podem perceber diferentes modos e ritmos de ocupação: na faixa litorânea, encontrava-se "a economia do açúcar no cerne das relações intercontinentais", "presa à dinâmica do mercado externo"; já no interior, "as opções econômicas desprendiam-se do vínculo quase exclusivo com o mercado externo, o que possibilitava a ação de pequenos e médios empreendedores" (Santos, 2010, p. 55). No caso da ocupação do sertão nordestino, os terrenos da caatinga e do cerrado eram propícios às pastagens que formariam as 'zonas de criação', "em razão da topografia relativamente plana, da vegetação arbustiva e da ausência de matas fechadas" (p. 52).

A partir de meados do século XVII, o interior do Nordeste é percorrido e ocupado por sertanistas paulistas, baianos e portugueses, "que estabelecem núcleos avançados de

povoamento" (Santos, 2010, p.56). Vemos, progressivamente, uma ampliação das fronteiras por meio da atividade pecuária. Puntoni (2002) destaca o crescimento dessa economia, bem como, paradoxalmente, as suas dificuldades, como fatores que teriam permitido a expansão do povoamento para o sertão. Para o autor, as necessidades da criação do gado forçavam a pecuária a ocupar regiões mais interioranas, sobretudo quando se valorizou o preço das terras próximas dos portos de embarque e dos cursos fluviais. Por outro lado, a busca por minerais preciosos ajudaria a promover o desbravamento do interior.

Esse processo de alargamento das fronteiras, por meio de duas correntes de povoamento, uma proveniente da Bahia, que acompanhou o curso do Rio São Francisco e do Rio Itapicuru e colonizou o chamado "sertão de dentro", e outra saindo de Pernambuco, que seguiu as regiões mais próximas do litoral até o Ceará e colonizou os "sertões de fora", não poderia deixar de resultar em conflitos com os povos indígenas que ocupavam essas vastas paragens. Para avaliarmos a dimensão do território, é preciso informar que

> [...] o sertão baiano, ou "interior", compreendia toda a região que ocupa o atual território do estado, incluída a margem ocidental do São Francisco, mais o interior do Piauí e o "território dos Pastos Bons", região do alto ltapicuru e rio das Balsas até Tocantins. O sertão "exterior", ou pernambucano, era mais próximo ao litoral. A corrente de povoamento, inicialmente acompanhando a linha costeira num território semi-árido que impossibilitava o plantio mesmo da cana e se estendia da Paraíba ao Ceará, passando pelo Rio Grande, acabou se encontrando com as correntes baianas, à medida que se interiorizava, particularmente pela bacia do rio Jaguaribe (Puntoni, 2002, p. 26).

Revendo a perspectiva historiográfica que se assenta na explicação de ocupação contínua do território, Santos (2010) propõe um processo de expansão geográfica e de dilatação das fronteiras na Bahia marcado por reversões, lacunas e descontinuidades. A bem da verdade, tal processo de expansão, por meio da atividade agropecuária, fundamentou-se na conquista e ocupação de um território de massiva população indígena, genericamente tratada como "tapuia", em oposição aos tupis da costa. Daí serem espaços de instabilidades, de constantes conflitos, caracterizados pela ameaça de retomada indígena ou ainda de esvaziamento, em função das pressões do meio, como longas distâncias, estiagens, enchentes, doenças provocadas por águas malsãs, animais selvagens.

A ampliação dessas fronteiras internas, portanto, não teria sido possível sem os conflitos com os chamados "tapuias", "povos ameríndios majoritariamente não sedentários, caracterizados pela heterogeneidade, diversidade linguística, migrações sazonais e popu-

lações menos densas e concentradas do que na faixa litorânea" (Santos, 2010, p. 118). As concentrações populacionais mais expressivas, que podiam incluir povos semissedentários ou mesmo sedentários, devem ter ocorrido nas áreas mais úmidas, como as margens dos rios, as ilhas fluviais e os brejos, justamente os espaços em que iriam se concentrar as frentes de ocupação do sertão, devido às diversas vantagens que possuíam os territórios ribeirinhos: provimento de água para homens e animais, abastecimento de peixes e crustáceos, fertilização de terrenos, navegação e orientação geográfica (Santos, 2010, p. 341).

Inevitavelmente, nesse processo, os próprios indígenas passaram também a integrar as frentes de ocupação. Como nos informa Santos (2010, p. 95), "seja como escravos, seja, o que foi certamente menos comum, como aliados ocasionais dos invasores, grandes grupos de índios, de centenas ou milhares de indivíduos, foram quase sempre incorporados aos exércitos da conquista". Além da participação direta nos combates, indígenas escravizados ou "aliados" eram utilizados nas expedições como guias, remadores, caçadores, operários na construção de casas-fortes e carregadores (Santos, 2010). Um recurso utilizado nesse contexto, como estratégia de enfraquecimento, foi a desterritorialização dos escravizados: indígenas da cidade da Bahia e de São Paulo eram transferidos para o sertão e "tornavam-se a mão de obra principal nos currais e lavouras do novo território", enquanto os indígenas "hostis" escravizados nos sertões "eram submetidos, desalojados e enviados para a cidade da Bahia" e para a capitania de São Vicente (Santos, 2010, p. 226-227).

Ainda segundo o autor, a "parceria" entre índios e colonizadores foi sempre assimétrica, em detrimento do indígena, sendo que o que houve, via de regra, foi uma "**trágica e quase rotineira sucessão de massacres de homens e escravização de mulheres e crianças**" (Santos, 2010, p. 118, grifo nosso). Um dos muitos exemplos desse fato se encontra no relato do padre Martinho de Nantes sobre uma expedição repressiva contra indígenas comandada pelo coronel Francisco Dias d'Ávila, o qual informa que 500 índios já rendidos do médio São Francisco foram mortos a sangue frio, tendo sido **escravizadas as mulheres e crianças** (Santos, 2010, p. 105, grifo nosso). Nesse cenário brutal, em que se encontra o extermínio de homens indígenas, ao lado da escravização de mulheres e crianças, veem-se criadas as condições para a emergência de uma substancial população mameluca, população cuja história linguística tem passado muitas vezes despercebida nos estudos da sócio-história linguística do Brasil[12].

Outro aspecto da mestiçagem ocorrida em território baiano, vastamente descrito

[12] Remetemos à nota 8 deste texto, em que destacamos a importância da tradição de pesquisa da UEFS, voltada para a história linguística do sertão da Bahia.

na documentação analisada por Santos (2010), é o que diz respeito à relevante presença dos sertanistas paulistas na ocupação da área. Assim como se deu em São Paulo, também no sertão da Bahia, os agentes da colonização portuguesa eram, não raro, homens mamelucos, frutos da união entre homens brancos e mulheres indígenas, ou descendentes de indivíduos mamelucos.

Por fim, quanto à permanência indígena na ocupação dos sertões, destacamos ainda a sua participação nas atividades econômicas que ali se desenvolveram. Segundo Santos (2014, p. 45),

> há notícias bastante seguras da contratação de índios pelos fazendeiros para o trabalho de condução das boiadas até a cidade. De acordo com Antonil, por exemplo, essa atividade era feita por 'brancos, mulatos e pretos, e também índios que com este trabalho procuram ter algum lucro'.

O autor ainda observa que, no ano de 1759, o cronista José Antônio Caldas também fez referência ao trabalho indígena na condução de boiadas pelo sertão. Esta seria a principal causa por que várias aldeias, antes existentes, se encontravam extintas. Nas palavras do próprio cronista,

> [...] servindose os viandantes dos Indios por estipendios, que lhes contribuiaõ para conduzirem as boyadas de gado de hum, e outro genero, foraõ ficando diminutas de tal sorte, que chegaraõ a dezertar Aldeas inteiras, principalmente aquelas, que ficavaõ vizinhas as estradas, que vaõ para os Sertoens e Minas (Caldas, 1759, p. 61).

Para Santos (2014, p. 46), tais testemunhos revelariam que os índios, aldeados ou não, buscavam meios de garantir ou ampliar por conta própria suas condições de sobrevivência, não dependendo inteiramente dos religiosos ou das autoridades coloniais. Também os missionários possuíam fazendas no sertão e recorriam à mão de obra indígena para a condução do gado, como se observa em episódio, ocorrido em 1732, com os índios das aldeias de Natuba, Canabrava e Saco dos Morcegos, os quais foram detidos, a mando do coronel Garcia de Ávila Pereira, e liberados pelo governador geral Conde de Sabugosa, alegando que o motivo da viagem seria a condução de boiadas dos reverendos padres da Companhia e de outros particulares (Santos, 2014, p.46).

A presença indígena no vasto território do sertão e sua relativa dizimação, marcada por uma profunda miscigenação, constituem-se como peças fundamentais na difusão e configuração do português no interior do território baiano. No âmbito das discussões sobre o processo de colonização e difusão da língua portuguesa no território brasileiro, não

podemos deixar de considerar esse quadro e a importância que esses espaços têm para uma história social linguística do Brasil, embora a historiografia linguística tenha dado pouca atenção a esse aspecto e priorizado quase exclusivamente a história do contato do português com línguas africanas.

Ampliando o olhar para o processo de ocupação de todo o território da Bahia, devemos considerar o povoamento do sul do Estado, caracterizado por uma demografia histórica com significativa presença indígena, "espaço periférico ao latifúndio canavieiro", o qual "redundou numa forma diferente de conquista, ocupação do solo e contato com os indígenas", como apontado por Mott (2010, p. 195). Destacando como característica geral a relevante presença indígena, Mott (2010) descreve, no período que vai de meados do século XVIII a meados do século XIX, as vilas e aldeias do sul da Bahia, delimitando as diferenciações entre a Comarca de Porto Seguro, em que os indígenas se encontravam na fase dos primeiros contatos com a população "civilizada", e a Comarca de Ilhéus, em que a interação com a sociedade colonial já seria, para algumas etnias, bicentenária, encontrando-se os indígenas "acaboclados", "residentes em vilas e nalguns aldeamentos remanescentes, fortemente deculturados" (Mott, 2010, p. 195).

O cenário aqui analisado aponta para a complexidade da economia colonial e para a participação de diversos atores na conquista e ocupação do espaço sertanejo, marcado pela relativa dizimação e também pela etnogênese dos povos indígenas (Almeida, 2001, 2010), ao se apropriarem dos instrumentos coloniais como estratégias de negociações e resistências. A questão da língua não poderia ficar à margem desses processos em que se relacionam economia e sociedade colonial, sobretudo no âmbito do mercado interno e da economia de subsistência, espaços onde não se encaixa a estrutura social dos engenhos, marcada pela relação entre senhores e escravizados.

A citação abaixo, extraída de Prado Jr. ([1946] 2011, p. 298-299), é lapidar para ilustrar o que pretendemos trazer à luz na seção seguinte:

> Abre-se assim um **vácuo imenso entre os extremos da escala social: os senhores e os escravos**; a pequena minoria dos primeiros e a multidão dos últimos. Aqueles dois grupos são os dos bem classificados da hierarquia e da estrutura social da colônia: os primeiros serão os dirigentes da colonização em seus vários setores; os outros, a massa trabalhadora. **Entre estas duas categorias nitidamente definidas e entrosadas na obra da colonização, comprime-se o número, que vai avultando com o tempo, dos desclassificados, dos inúteis e inadaptados**; indivíduos de ocupações mais ou menos incertas e aleatórias, ou sem ocupação alguma.

4. Estratificação social e estrutura familiar

Nesta seção, faremos uma análise da configuração social do Brasil no período colonial, trazendo o foco para a região Nordeste, tal como fizemos na seção anterior, e buscando relacionar as estruturas sociais com a configuração linguística. O objetivo é refletir não apenas a respeito da presença e difusão da língua portuguesa em concorrência/coexistência com as outras línguas da nossa história linguística, mas ainda melhor compreender historicamente a variação no português brasileiro, tratando da questão das normas linguísticas e da polarização do português brasileiro. Para tanto, traremos à baila três aspectos da configuração social aqui considerados de suma importância: estratificação social, mestiçagem e estrutura familiar.

4.1 Estratificação social e mestiçagem

Como estava constituída a sociedade colonial brasileira? De que maneira essa estrutura social se refletia na realidade linguística, naquele momento? A ideia de uma sociedade bipartida, em que se encontravam, de um lado, os senhores de terras e engenhos, e de outro, a população escravizada, complexifica-se em um olhar mais detido para a estratificação social, na qual podemos observar uma considerável diversificação de papéis sociais, já desde o primeiro século de colonização do Brasil.

As ocupações existentes na sociedade quinhentista foram divididas por Siqueira (1978) em três grandes grupos, nomeados pela autora de grande burguesia, pequena e média burguesia e povo, em uma análise que se debruçou sobre a população brasileira do século XVI, tendo como *corpus* as fontes do Tribunal da Inquisição em sua primeira visitação ao Brasil[13]. No primeiro grupo, da grande burguesia, encontravam-se os senhores de engenhos e grandes comerciantes, "partícipes da mesma concepção de vida, donos das tendências autoritárias, que transmitiam por herança seus bens e suas profissões" (Siqueira, 1978, p. 76), e ainda o alto clero, que não raro possuía terras ou engenhos. O segundo grupo, da pequena e média burguesia, englobava artesãos, assalariados, burocratas, homens do mar, pequenos lavradores, liberais, mercadores de loja e clero médio e baixo, estando presentes, em cada um desses grupos, uma grande lista de ofícios. No terceiro

[13] O Tribunal do Santo Ofício atuou na América Portuguesa através de uma justiça itinerante, as chamadas "visitações". Foram quatro as visitações realizadas, sendo a primeira, liderada pelo visitador Heitor Furtado de Mendonça, instalada nas capitanias da Bahia, Pernambuco, Itamaracá e Paraíba, entre os anos de 1591-1595. Encontram-se incluídos no *corpus* dessa primeira visitação os Livros das Confissões e Denunciações da Capitania da Bahia e os Livros das Confissões e das Denunciações da Capitania de Pernambuco, os quais contêm registros referentes às Capitanias de Pernambuco, de Itamaracá e da Paraíba (Cf. a lista de documentos da "PRIMEIRA visitação do Santo Officio às partes do Brasil pelo licenciado Heitor Furtado de Mendonça" nas Referências).

grupo, estava o povo, "a massa indefinida e informe da plebe, e os escravos responsáveis por todo o trabalho das lavouras e das fábricas" (Siqueira, 1978, p. 77); incluíam-se aí os indivíduos responsáveis por pequenos ofícios, como boticários, canoeiros, confeiteiros, carreiros, cortadores de carne, barbeiros, vendeiros, ferradores, hortelãos, línguas, pedreiros, pescadores, pasteleiros, parteiras, padeiros, camareiros, criados, pajens, "mulheres do mundo", taverneiros, vinhateiros, trabalhadores de soldada, e somavam-se a estes os mendicantes e os escravos.

Em uma análise dos níveis de letramento nos distintos estratos sociais, tendo como *corpus* as mesmas fontes inquisitoriais[14] abordadas por Siqueira, Lobo e Sartori (2020) observaram percentuais de assinantes que decresciam, da classe mais alta à classe mais baixa: os homens da grande burguesia apresentaram, em média, 98% de assinaturas, mostrando-se praticamente todos capazes de assinar, os homens da pequena e média burguesia exibiram também percentuais bastante altos de assinantes, assinando em uma média de 92% e, entre os homens do povo, havia 77% de assinantes nos livros da Inquisição. Tais percentuais, que podem ser considerados muito altos, sobretudo quando avaliado que se trata do primeiro século da colonização, devem ainda ser melhor entendidos, ao observar-se que os assinantes apresentaram diferentes níveis de execução gráfica, indicando que estariam por trás gradações no seu nível de letramento, estando aí incluídos desde aqueles que pouco ou nada dominariam da tecnologia da escrita, além de assinar – e mal – seu próprio nome, até aqueles que exibiam um alto grau de letramento, com grande fluência nas práticas de leitura e escrita.

A análise dos níveis de letramento corrobora o que apontou o estudo de Siqueira sobre a estratificação da sociedade colonial brasileira: trata-se de uma sociedade com uma estratificação social muito mais complexa do que a da clássica divisão entre "casa-grande" e "senzala". Se, em termos culturais, podemos considerá-la assim também diversificada, englobando indivíduos com diferentes níveis de acesso à cultura escrita, podemos considerar que algo semelhante ocorreria a respeito das variantes linguísticas empregadas por esses indivíduos: por trás da já comentada "polarização linguística" do português brasileiro, haveria, ao que tudo indica, um quadro bem mais complexo, em que falantes de variantes mais cultas e mais populares estariam em interação, e as gradações no nível de acesso à cultura escrita e a um padrão normativo europeizado seriam inúmeras. Tal aspecto da

[14] Trata-se igualmente dos documentos da Primeira Visitação da Inquisição ao Brasil: os livros de Confissões e Denunciações da Capitania da Bahia e os Livros de Confissões e Denunciações da Capitania de Pernambuco. As demais referências que se farão neste artigo aos registros do Tribunal da Inquisição têm essa mesma fonte, os livros da Primeira Visitação. Serão ainda citadas as fontes da Segunda Visitação da Inquisição ao Brasil, ocorrida em 1618: o Livro de Confissões e o Livro de denunciações da Bahia, analisados em Lobo, Sartori e Soares (2016).

história linguística brasileira ganha bastante evidência quando analisamos a questão da mestiçagem entre indivíduos falantes de línguas diferentes ou indivíduos descendentes de falantes de línguas diferentes, e como este aspecto interage com a estratificação social.

O primeiro aspecto da mestiçagem a ser destacado na história colonial brasileira diz respeito à mestiçagem que gerou uma população mameluca, ocorrida, desde o primeiro século de colonização, entre homens brancos e mulheres indígenas. Rodrigues (1996, 2010) trouxe importantes contribuições a respeito da presença da população mameluca na história social linguística do Brasil, ressaltando a sua relevância para a difusão das chamadas "línguas gerais", de base indígena tupiniquim e tupinambá e descrevendo sua presença nas regiões Sudeste, Centro-Oeste e Norte do Brasil. Mais tarde, Lobo, Machado Filho e Mattos e Silva (2006) discutiram indícios de sua presença no sul da Bahia, e Argolo (2015) apresentou evidências da sua existência mais difundida entre vilas do sul da Bahia, nas Capitanias de Ilhéus e Porto Seguro, que, em meados do século XVIII, passaram a integrar a Capitania da Bahia.

A respeito da formação das línguas gerais no Brasil, que teria suas origens na utilização de línguas indígenas por uma população mameluca, destacamos aqui o seguinte trecho de Rodrigues:

> A expressão *língua geral* tomou um sentido bem definido no Brasil nos séculos XVII e XVIII, quando, tanto em São Paulo como no Maranhão e Pará, passou a designar as línguas de origem indígena faladas, nas respectivas províncias, por toda a população originada no cruzamento de europeus e índios tupi-guaranis (especificamente os tupis em São Paulo e os tupinambás no Maranhão e Pará), à qual foi-se agregando um contingente de origem africana e contingentes de vários outros povos indígenas, incorporados ao regime colonial, em geral na qualidade de escravos ou de índios de missão (Rodrigues, 1996, p. 5).

Como podemos observar na citação acima, Rodrigues situa a sua descrição das línguas gerais nos territórios de São Paulo, Maranhão e Pará. Quando tentamos ampliar a análise para todo o território brasileiro, questionando se seria possível fazer generalizações a partir da descrição citada, algumas questões vão se colocando, as quais se podem concentrar aqui em duas perguntas principais: 1) Onde se encontrava a população mameluca no Brasil?; 2) Que língua(s) era(m) falada(s) por essa população?

A respeito da primeira questão, é preciso de início pontuar que não haveria propriamente "uma" população mameluca no Brasil, mas uma mestiçagem que, se apresentava notáveis características culturais em determinadas regiões, como seria o caso dos

bandeirantes paulistas, em outras regiões, como provavelmente se daria em quase todo o Nordeste, parecia se diluir na população branca colonial, de maneira que nunca foi possível verificar, nas estimativas demográficas desse período, onde estavam e quantos eram os mamelucos. A maior parte das estimativas calculava o número de brancos a partir da quantidade de famílias de portugueses presentes na colônia, e o número de índios integrados ao sistema colonial, que se encontravam nas aldeias missionadas. A população que era fruto da miscigenação entre portugueses e índias fica, assim, invisibilizada, sendo muitas vezes inserida na quantificação dos membros das famílias de portugueses, se de fato residiam na casa paterna. Exemplos de tal inserção dos mamelucos nas famílias de portugueses são notados em dados do Tribunal da Inquisição em visitações ao Brasil nos séculos XVI e XVII. Nos depoimentos registrados nos livros da Inquisição, vislumbramos uma população nas capitanias do Nordeste que possuía um significativo número de mamelucos, filhos de portugueses, os quais demonstravam herdar o *status* social e o estado religioso de "cristão velho" ou "cristão novo" de seus pais portugueses (Cf. Lobo; Sartori, 2020; Lobo; Sartori; Soares, 2016).

Para respondermos à segunda questão sobre os mamelucos, em que se indaga que língua(s) seria(m) falada(s) por esses sujeitos, faz-se necessário rompermos com a ideia de que haveria uma correspondência direta entre mamelucos e línguas gerais. Se nem todo falante de língua geral seria mameluco, como admite o próprio Rodrigues (1996) no trecho supracitado, podemos ainda admitir o contrário: nem todo mameluco seria falante de língua geral. Se é fiel o relato de Vieira sobre a sociedade paulista em fins do século XVII, ao dizer que "as famílias dos Portugueses e Indios em São Paulo, estão tão ligadas hoje umas com as outras, que as mulheres, e os filhos, se crião mistiça e domesticamente, e a lingua, que nas ditas famílias se fala, he a dos Indios, e a Portuguesa a vão os meninos aprender a escola" (Vieira, 1736, p. 161-162) – e não há elementos para duvidar de tal afirmação – , não devemos, por outro lado, tomar esse relato de forma generalizante. A respeito disso, não deixam dúvida os depoimentos presentes nos já citados livros do Tribunal da Inquisição. Mamelucos e mamelucas de primeira e segunda geração que depuseram perante o Tribunal português instalado nas Capitanias do Nordeste demonstraram ser muito mais influenciados social e culturalmente pelos seus pais brancos, sendo falantes de língua portuguesa, professando a fé cristã, herdando cargos políticos de seus pais, do que por suas mães indígenas, das quais muito pouco se fala; por vezes, nos depoimentos, sequer o nome de suas mães é mencionado.

No que tange ao letramento dos mamelucos que depuseram durante a Primeira

Visitação da Inquisição, alocada na Capitania da Bahia e nas Capitanias de Pernambuco, Itamaracá e Paraíba, em uma análise das assinaturas deixadas nos livros de registro dos depoimentos prestados ao Tribunal, observamos que os mamelucos foram capazes de redigir sua assinatura não apenas tão bem quanto os homens brancos, mas mesmo com percentuais maiores: 93% de homens brancos assinaram seu nome, ao lado de 100% de mamelucos assinantes. Entre as mulheres, embora fosse sempre baixo o percentual de assinantes, as mamelucas assinaram também mais que as brancas: encontramos 17% de assinantes entre as mamelucas e 14% entre as brancas (Lobo; Sartori, 2020).

Vemos, assim, claramente, que os mamelucos das capitanias do Nordeste diferem social, cultural e linguisticamente dos mamelucos paulistas da descrição de Rodrigues (1996). Ainda dentro do território do Nordeste, não parece apropriado fazer generalizações. Como já vimos, nas capitanias do sul da Bahia, há evidências da presença de língua geral sendo falada pela população.

É preciso ainda retomarmos a questão de saber que língua(s) seria(m) falada(s) pela população mameluca e dizermos que não foi encontrada, nas fontes inquisitoriais, nenhuma informação além da mais simples menção de que se trataria, nas Capitanias do Nordeste que foram alvo do Tribunal nos séculos XVI e XVII, da língua portuguesa. Tais falantes seriam proficientes em língua(s) indígena(s)? Não sabemos. A variante da língua portuguesa que era utilizada seria uma variante mais próxima do português europeu? Não é certo. Não há elementos suficientes para abordarmos, no momento, qual seria o grau de interferência de possíveis línguas maternas de base indígena sobre o português falado por esses mamelucos.

Do que já vimos até aqui, é possível pontuarmos, repetimos, a evidente falta de unidade – social, cultural, linguística – entre os integrantes da chamada "população mameluca". Se a sua aproximação com uma cultura indígena saltava aos olhos em algumas regiões da Colônia, como seria o caso do Sudeste, em outras regiões, ficam apenas as dúvidas de qual seria o nível de tal aproximação.

O segundo aspecto da mestiçagem, também fundamental para a história linguística do Brasil, é aquele que deu origem a uma população mulata. Já no primeiro século de colonização, inicia-se, juntamente à vinda dos colonizadores portugueses, a vinda de uma população africana escravizada. Também desde o primeiro século de colonização vislumbramos o surgimento de uma população mulata, fruto de uniões, geralmente não matrimoniais, que se davam, no início, entre homens brancos e mulheres negras, e que em seguida vão ocorrendo entre brancos e mulatas, mulatos e negras, mulatos e mulatas, cafuzos...

Ao lado das crianças mulatas, cresciam crianças brancas que, podemos considerar, tinham "um pé" na mestiçagem – se não tinham propriamente sangue mestiço, a sua criação era ao lado de crianças negras e mulatas, para além das mamelucas, que viviam dentro de suas casas. Todas recebiam cuidados, e por vezes também o mesmo leite, de uma mulher negra.

Qual(is) seria(m) a(s) línguas(s) dos mulatos? De forma generalizante, podemos responder que línguas africanas e a língua portuguesa seriam faladas por essa população, além de haver a possibilidade do uso das línguas gerais de base indígena em determinadas regiões, caso o contexto social o favorecesse, isto é, caso houvesse uma interação social com falantes de língua geral a ponto de gerar a aquisição dessa língua.

No que tange ao uso das línguas africanas por negros e mulatos, é válido salientarmos a amplitude que algumas delas alcançaram no período colonial. Esse seria o caso do quimbundo, descrito na gramática do jesuíta Pedro Dias, a *Arte da lingua de Angola,* publicada ao final do século XVII (Dias, 1697), testemunho do emprego corrente de uma língua africana no Brasil (Petter, 2006) sendo utilizada como língua veicular entre africanos, portugueses e brasileiros: a gramática teve a finalidade de auxiliar na comunicação entre jesuítas e negros escravizados, facilitando o processo de catequese. Igualmente relevante parece ter sido o uso de uma língua veicular, conhecida como "língua geral de Mina", que teria se desenvolvido no quadrilátero mineiro de Vila Rica - Vila do Carmo - Sabará - Rio dos Montes, originada do contato entre línguas africanas próximas entre si, vindas da Costa do Benin, da qual ficou o registro no documento manuscrito *Obra nova da língua geral de Minna,* redigido em 1731/1741, por Antônio da Costa Peixoto, documento que tinha a intenção de facilitar aos senhores de escravizados o aprendizado da língua utilizada nas minas (Petter; Cunha, 2016).

A respeito do uso da língua portuguesa por essa população negra e mulata, cabe retomar aqui a ideia de Mattos e Silva (2000b), de que um "português geral brasileiro", antecedente histórico do nosso português popular brasileiro, teria sido falado por africanos e afrodescendentes, e ainda, em menor número, por uma população indígena e seus descendentes integrados à sociedade colonial. Tendo em conta que africanos e afrodescendentes constituíam maioria da população durante todo o período colonial, esses teriam sido os principais difusores do português no território brasileiro. Mattos e Silva pontua também, apoiada nos estudos de Mattoso (1990), que haveria uma gradação no que diz respeito ao aprendizado da língua portuguesa pelos escravizados: enquanto, para muitos, seria suficiente um conhecimento precário da língua, que permitisse a compreensão das

ordens – como seria o caso daqueles que trabalhavam nas lavouras de cana de açúcar –, para outros, como os escravizados domésticos, teria sido comum se tornarem bilíngues, e seus filhos, criados na primeira infância junto aos filhos do senhor, aprenderiam melhor o português e menos a(s) língua(s) africana(s) de seus pais.

Observando as relações sociais que se davam nesse período e essa gradação no acesso à língua portuguesa, podemos afirmar que mulatos e negros brasileiros seriam os principais atores no processo de difusão do português no Brasil, enquanto os africanos, em sua maioria, não teriam tido tanto acesso ao português. Isso porque, quanto mais "integrados" à sociedade colonial, melhor seria a aquisição da língua portuguesa por esses indivíduos. Entre africanos era mais difícil, de maneira geral, essa integração à sociedade. Aqueles que viviam nas senzalas e de lá saíam para as lavouras, onde passavam os dias e de lá retornavam, exaustos, para as senzalas, teriam não só pouco acesso à língua portuguesa, como pouca oportunidade de diálogo em sua língua materna.

Quanto aos negros e mulatos já nascidos no Brasil, esses tinham acesso mais prolongado à língua portuguesa, podendo tê-la inclusive como língua materna e, em geral, encontravam-se também mais integrados à sociedade colonial. Mattoso (1990) diferencia o "escravo boçal", negro nascido na África, que "falava mal a língua dos brancos", dos "crioulos", nascidos no Brasil, "que, logicamente, falam realmente o português", os quais "em geral, foram criados na família do senhor e são fortemente marcados pela sociedade dos brancos" (Mattoso, 1990, p.105-106).

Se, no século XVI, a integração desses indivíduos era ainda incipiente, nos séculos seguintes esta já se fazia bastante visível, nas diversas atividades desempenhadas por esses indivíduos na sociedade, na quantidade de negros e mulatos que conquistavam alforria, cuja ascensão social resultava em que, por vezes, comprassem "seus próprios escravos"; na sua inserção em irmandades de negros, espaços que poderiam proporcionar inclusive o aprendizado da leitura e da escrita e um auxílio para comprar sua alforria; na atuação como oficiais nos exércitos brasileiros; ou ainda como religiosos, professando a fé católica.

Nas cidades, negros e mulatos se inseriam nas profissões humildes. Trabalhavam como serralheiros e como ajudantes nas tarefas menos nobres de cada ofício mecânico, sobretudo as que exigiam maior esforço físico, como o transporte de madeira, o serramento e o desbastamento (Flexor, 1974). Mestiços artesãos trabalhavam nos diversos ofícios da arquitetura, como oleiros, ladrilheiros, telheiros, pedreiros, canteiros, rebocadores; carpinteiros, carapinas, entalhadores, ferreiros, serralheiros, latoeiros e até mes-

mo pintores e escultores (Holanda, 1985). Os escravizados conhecidos como "negros de ganho" exerciam diversas atividades, proporcionando lucros a seus senhores, trabalhando como carregadores, doceiros, cozinheiros, alfaiates, sapateiros (Vianna Filho, 2008); entre as mulheres, algumas praticavam a prostituição, imposta pelos seus senhores como fonte de renda complementar ao comércio ambulante em que elas atuavam (Figueiredo, 2004). Os negros libertos trabalhavam de barbeiros, dentistas, fabricantes de vassouras de piaçava, importadores de sabão-da-costa, alguns lavando chapéus-do-chile, enquanto as negras trabalhavam de doceiras, com seus tabuleiros, ou lavadeiras (Freyre, [1933] 2002). A arte de sangrar era também exercida por negros. Havia ainda parteiras negras, que, além de partejarem, curavam doenças ginecológicas por meio de bruxedos, rezas, benzeduras (Flexor, 1974; Freyre, [1933] 2002).

A respeito da inserção de negros nas chamadas irmandades, encontram-se análises em Oliveira (2006), Lobo e Oliveira (2007), Lobo e Oliveira (2009) e Galves e Lobo (2019), nas quais se veem os documentos guardados pelas irmandades negras no Brasil como importantes testemunhos do letramento de africanos e afrodescendentes, dos quais são exemplo os documentos da Sociedade Protetora dos Desvalidos, fundada em Salvador, em 1832, por africanos, e integrada, a partir de 1835, exclusivamente por sócios negros nascidos no Brasil.

No que tange à presença de homens negros entre os religiosos, é sintomática a consulta do Conselho Ultramarino ao rei D. Pedro II, de 23 de fevereiro de 1701, em que se aborda o recrutamento de negros forros, ou, em sua falta, escravizados, que pudessem ser instruídos nos mistérios da fé para atuarem na catequese de africanos, melhor se comunicando com estes por conhecerem a sua língua (Consulta, 1701). Como afirma Marcussi (2013), a formação de catequistas negros, falantes nativos de quimbundo e outras línguas africanas, e o envio de padres angolanos ao Brasil fizeram parte de uma mesma iniciativa jesuítica para a catequese dos escravizados. Ligada a esta iniciativa, estaria a construção de um anexo do colégio de Luanda, na África, onde seria possível estudarem negros africanos, pedido feito pela Companhia de Jesus e aprovado em 1684 pelo rei de Portugal (Marcussi, 2013). O clero pode ser avaliado como uma classe representativa de uma elite cultural, bastante letrada; encontrar, entre estes, negros forros e até mesmo escravizados indicia a diluição das fronteiras sociais que punham de um lado uma elite culta e de outro, o povo.

Os chamados terços de homens de cor, milícias de homens negros e mulatos, surgiram no país no século XVII e se disseminaram pelas capitanias ao longo do século XVIII.

Dentre os diversos documentos redigidos por oficiais dos terços guardados no Arquivo Histórico Ultramarino, encontra-se uma petição escrita pelo Capitão-mor do Terço de Henriques da Bahia Luís Tavares de Meneses, e assinada por outros sete oficiais do terço, em que se nota certa ascensão social que conquistavam os negros do exército, em detrimento do que viviam os negros africanos, sobretudo os da Costa da Mina, os quais, caso entrassem como soldados nos terços, não poderiam ascender a outros cargos, por serem considerados "indignos de subirem aos postos de sargento para cima". Na petição citada, os oficiais Henriques solicitavam que fosse dada uma rigorosa punição a quem promovesse os soldados da nação da Costa da Mina às posições superiores, por serem estes considerados "de gênio revoltoso, e faltos totalmente de fé e verdade, e capitais inimigos de todas as pessoas brancas, crioulos e pretos da nação Angola" (Silva, 2012, p. 03-04).

Do já exposto, vislumbramos não apenas o alargamento dos papeis sociais exercidos pela população negra, mas ainda a notável inserção social dos mestiços, filhos de homens brancos ou de seus descendentes já mestiços, quer fossem filhos de mulheres indígenas ou mamelucas, quer fossem filhos de negras ou mulatas. A posição econômica e social, as profissões exercidas, os cargos políticos desempenhados, o nível de letramento desses indivíduos foram extremamente variáveis ao longo de nossa história.

Em termos de aquisição linguística, observamos, ao tratar da língua portuguesa falada em nosso passado colonial, que tal diversificação na estrutura social se coaduna mais com diversas gradações de normas linguísticas em uso do que com normas em polos apartados. Nesse âmbito, faz-se necessária a reflexão sobre o modelo de língua materna ao qual esses indivíduos tiveram acesso. Aqui, não há como desconsiderar o papel das mulheres na aquisição linguística, as quais não raro são silenciadas nos estudos históricos. Sobre essa relevante questão, trataremos a seguir.

4.2 Estrutura familiar e o papel das mulheres

Para além da questão da estratificação social, a configuração das famílias brasileiras guarda também íntima relação com a situação cultural e linguística no período colonial. No seio familiar, marido e mulher tinham papeis muito diferenciados. Enquanto os homens participavam de maneira mais ativa na sociedade, atuando na vida pública, exercendo cargos políticos, frequentando, quando jovens, os colégios, às mulheres cabia a reclusão do lar.

Araújo (2004, p.40) informa que se repetia "como algo ideal, nos tempos coloniais, que havia apenas três ocasiões em que a mulher poderia sair do lar durante toda sua vida:

para se batizar, para se casar e para ser enterrada", e considera que a afirmação, embora exagerada, encontra respaldo em impressões como a do viajante Froger, de passagem por Salvador em 1696, que dizia que ali as mulheres eram de dar pena, jamais viam ninguém, saindo de suas casas apenas aos domingos, para ir à igreja. Também em Salvador, o arcebispo da cidade, em 1751, queixava-se de que os pais proibiam as moças até de assistir às devotas lições no Colégio das Mercês (Araújo, 2004).

Tal situação de reclusão era sempre mais acentuada entre as camadas mais altas da sociedade. Como afirma Figueiredo (2004), entre as camadas baixas da população, a divisão dos papéis obedecia mais às necessidades econômicas de cada família do que a um preconceito sexual na distribuição de tarefas, podendo-se considerar que "a transferência da chefia dos domicílios para a mulher nos núcleos familiares simples tornou a atuação feminina tão mais importante quanto mais íntima era a associação entre vida doméstica e trabalho produtivo" (p. 149). Assim sendo, nas camadas sociais menos privilegiadas, via-se uma divisão de papéis no domicílio, caracterizada por uma maior atuação feminina do que a prevista no casamento cristão, sendo o maior estímulo para a definição de papéis não mais o discurso teológico, que determinava a submissão feminina no casamento, e muito mais "as exigências de um cotidiano em que era vital a repartição de tarefas ou a transferência de papéis para a sobrevivência do grupo doméstico" (Figueiredo, 2004, p. 147-148).

Nas Visitações do Tribunal da Inquisição ao Brasil de finais do século XVI, as mulheres que iam prestar depoimento mencionavam, geralmente, a ocupação do marido, quando eram casadas ou viúvas, e a do pai, quando solteiras. Pouquíssimas declararam possuir um ofício, figurando essas geralmente entre as mulheres do povo e exercendo ofícios como o de vendeira, estalajadeira, padeira, parteira, costureira, tecedeira, criada. Notamos ainda, nos registros do Tribunal da Inquisição, uma sub-representação feminina, também indicativa da situação de reclusão feminina: de um total de 685 testemunhas que se apresentaram para depor nas capitanias da Bahia, Pernambuco, Itamaracá e Paraíba, 508 (74%) eram homens e 177 (26%), mulheres.

Nos séculos seguintes, tal realidade parecia se manter de maneira semelhante, também no Sudeste, como se vê no seguinte relato a respeito das atividades desempenhadas por mulheres em Minas Gerais, no século XVIII:

> Em que lugar encontraríamos a mulher mineira? Comecemos pela negação, que parece ter sido a característica central na vida dessas mulheres. Estiveram nas Minas excluídas de qualquer exercício de função política nas câmaras municipais, na administração eclesiástica, proibidas

> de ocupar cargos da administração colonial que lhes garantissem reconhecimento social. Os papéis sexuais na colônia reproduziam o que se conhecia na metrópole. Tudo parecia confirmar isso. O exame da atuação feminina nos ofícios mecânicos é desalentador. Entre os ofícios que se multiplicam pelas Gerais, por multidões de ferreiros, latoeiros, sapateiros, pedreiros, carpinteiros, ourives, pouco se vislumbra da presença feminina. Apareciam, sim, ocupadas na panificação, tecelagem e alfaiataria, dividindo com os homens essas funções, cabendo-lhes alguma exclusividade quando eram costureiras, doceiras, fiandeiras e rendeiras. Ainda como cozinheiras, lavadeiras ou criadas reproduziam no Brasil os papéis que tradicionalmente lhes eram reservados. Algumas, através de uma prova prática, assistida por médicos e sangradores, promovida pelas câmaras municipais, receberam "cartas de exame", uma espécie de diploma que as tornava aptas ao exercício legal da função de parteira.
>
> Nos pesados trabalhos de extração mineral parecia confirmada a imagem legada pela pena do viajante Rugendas, *Lavagem de ouro, perto de Itacolomi*, em que mulheres em segundo plano aparecem carregando gamelas com pedras que seriam lavadas. Não havia impedimentos formais a seu trabalho na mineração, mas, diante de exigências de resistência e força física, restringiam-se à função de carregadoras de gamelas. [...] Tampouco seria aí o lugar em que as mulheres se destacavam: proprietários de minas de ouro no século XVIII possuíam muito mais escravos homens (Figueiredo, 2004, p. 120).

Figueiredo (2004, p. 120) ainda completa sua descrição sobre a atuação das mulheres na economia colonial informando sobre uma "participação respeitável de mulheres que, como roceiras em pequenas propriedades arrendadas, aparecem nas listagens de algumas freguesias que pagavam o dízimo à Coroa", em que as mulheres representavam em torno de 10%, chegando até 23% do total de proprietários. Seria um número significativo de mulheres criando gado, aves, plantando gêneros alimentícios para abastecimento local, produzindo queijos, aguardente e pão, número que ainda teria crescido no século XIX, segundo os dados do censo de Vila Rica desse período, que registravam 51 mulheres para 27 homens entre roceiros, lavradores e hortelãos. A presença feminina teria sido ainda sempre destacada no exercício do pequeno comércio em vilas e cidades do Brasil colonial, sendo o comércio ambulante uma ocupação marcadamente feminina (Figueiredo, 2004)

No que tange ao acesso negado às moças à escolarização e letramento, podemos observar relatos quinhentistas, como o de José de Anchieta ([1554-1594] 1933), em que as referências feitas aos alunos dos colégios jesuíticos eram sempre sobre "moços" ou "rapazes", não havendo qualquer menção a garotas. Também nos depoimentos prestados

por mestres leigos na primeira Visitação do Santo Ofício ao Brasil, estes se declaravam perante o Tribunal "mestres de ensinar **moços**" (*Quarto Livro das Denunciações*, fol. 88r; fol. 247r, grifo nosso). As referências a aulas oferecidas às moças são apenas de aulas de costura que tomavam em casas de senhoras como *Branca Dias*, cristã-nova residente na Vila de Olinda, citada nos livros da Inquisição por conservar hábitos judeus proibidos pela Inquisição.

Ainda vemos, nos dados do Tribunal da Inquisição do século XVI, que era extremamente baixo o percentual de mulheres capazes de assinar o próprio nome, dado que salta aos olhos sobretudo quando contrastado ao percentual masculino de assinantes: se, entre os homens, se encontram 92% de assinantes; entre as mulheres, o percentual de assinaturas cai para 12%. Além do índice extremamente baixo de assinantes entre as mulheres, vemos ainda que as assinaturas por elas executadas eram, não raro, mal executadas, apresentando grande dificuldade no traçado das letras e revelando, assim, que a sua habilidade na escrita provavelmente não passaria de assinar, e mal, o seu nome (Lobo; Sartori, 2020).

Dois séculos depois da situação retratada nos livros da Inquisição de finais do século XVI, em que se viam mulheres quase totalmente alheias às práticas da escrita e ao letramento, o francês Le Gentil de la Barbinais, em meados do século XVIII, em visita em uma noite de Natal às clarissas do Convento de Santa Clara do Desterro, um convento que abrigava mulheres da elite, em Salvador, afirmou que elas falavam como os escravos (Nascimento, 1994).

O enorme contraste entre os níveis de letramento de homens e mulheres no Brasil colonial impõe a seguinte e relevante indagação: Qual(is) seria(m) a(s) variante(s) linguística(s) do português utilizada(s) em uma família em que coabitavam, muitas vezes, um homem branco, mais ou menos letrado, uma mulher que, quer fosse branca, índia, mameluca ou mulata, era, na maioria das vezes, analfabeta ou semialfabetizada, e ainda os seus filhos, mais ou menos letrados, em contato com variantes distintas, ou até mesmo línguas diferentes?

A desigualdade das normas linguísticas utilizadas por homens e mulheres prolongou-se durante séculos na história do Brasil. No século XIX, vemos ainda claras diferenças no nível de letramento de homens e mulheres em cartas pessoais, como aquelas escritas por Christiano e Bárbara Ottoni aos seus netos. Christiano, professor e senador do Império e da República, demonstrava grande desenvoltura na escrita, com um uso de um português culto e bastante formal, enquanto sua esposa, Bárbara, que era dona

de casa, escrevia cartas curtas e bastante simples, que fugiam à norma-padrão (Lopes, 2005). Também nas cartas de amor escritas por Carlos e Iracema, já do século XX, observamos diferenças na escrita, ainda que os dois tivessem o nível primário de escolarização. Carlos, que foi comerciante, pecuarista e político, redigia, ele próprio, seus discursos e outros documentos, e revelava maior domínio da prática de escrita, escrevendo com maior desenvoltura e eloquência e demonstrando ter alcançado um alto nível de letramento, enquanto Iracema, que se dedicou inteiramente ao lar, não exercendo uma profissão, escrevia bem, mas sem muita desenvoltura, demonstrando pouca prática de escrita (Sartori, 2010).

5. Considerações finais

A despeito do grande desenvolvimento atingido pela Linguística Histórico-Diacrônica no Brasil nos últimos 35 anos, aprofundando o nosso conhecimento sobre a constituição histórica da língua que se tornou hegemônica – o português brasileiro –, estamos ainda dando os primeiros passos no que tange à história linguística dos três séculos de conquista e colonização da América Portuguesa, o que, indubitavelmente, pode resultar em pretensas sínteses generalizantes, que, em realidade, são, ainda, sínteses parciais.

Se é verdade, conforme afirma Petter (2017, p. 13), com base em Alencastro (2000), "que o Brasil se formou fora do Brasil, não como um prolongamento da Europa, mas como participante de um sistema de exploração colonial singular em ação no Atlântico Sul", e que "As evidências da história permitem extrapolar para o terreno da língua a tese de Alencastro [e afirmar que] o PB também se formou no convívio com as línguas do Centro-Sul da África, sem negar, contudo, a matriz portuguesa, nem os contatos posteriores com línguas da África Ocidental", é também igualmente verdade que "o Brasil se formou dentro do Brasil", no litoral e no sertão.

Neste texto, apresentamos, inicialmente, argumentos a favor da caracterização sociolinguística do português brasileiro contemporâneo não como uma realidade polarizada, mas como um *continuum* de normas. Isto posto, os passos seguintes – iniciais e ainda claudicantes – foram no sentido de desvelar o encaixamento histórico desse *continuum* de normas. Para tal, elegemos um espaço – o Nordeste e, mais particularmente, o sertão da Bahia – e um tempo – o período colonial –, analisando aspectos das estruturas econômicas e sociais, bem como da estrutura familiar, e indagando sobre os seus prováveis correlatos linguísticos.

Referências

ALENCASTRO, L. F. de. *O trato dos viventes*: formação do Brasil no Atlântico-Sul. São Paulo: Companhia das Letras, 2000.

ALMEIDA, M. R. C. de. Os índios aldeados: histórias e identidades em construção. *Tempo*, Rio de Janeiro, v. 6, n. 12, p. 51-71, jul. 2001.

ALMEIDA, M. R. C. de. *Os índios na história do Brasil*. Rio de Janeiro: Editora FGV, 2010.

ÁLVAREZ-LÓPEZ, L. Á.; GONÇALVES, P.; AVELAR, J. O. de. *The Portuguese language continuum in Africa and Brazil*. John Benjamins, 2018.

ANCHIETA, J. de. *Cartas*: informações, fragmentos históricos e sermões do padre José de Anchieta (1554-1594). Rio de Janeiro: Officina Industrial Graphica, 1933. Disponível em: https://bdlb.bn.gov.br/acervo/handle/20.500.12156.3/441059. Acesso em: 20 abr. 2013.

ARAÚJO, E. A arte da sedução: sexualidade feminina na Colônia. *In:* DEL PRIORE, Mary; BASSANEZI, Carla. *História das mulheres no Brasil*. 7. ed. São Paulo: Contexto, 2004. p. 45-77.

ARGOLO, W. 2015. *História linguística do Sul da Bahia (1534-1940)*. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia. Salvador, 2015.

BAXTER, A.; LUCCHESI, D. A relevância dos processos de pidginização e crioulização na formação da língua portuguesa no Brasil. *Estudos linguísticos e literários*. Salvador, n. 19, p. 65-84, 1997.

BESSA FREIRE, J. R. *Rio Babel:* a história das línguas na Amazônia. Rio de Janeiro: Atlântica, 2004.

BORTONI-RICARDO, S. M. *Do campo para a cidade*: estudo sociolinguístico de migração e redes sociais. São Paulo: Parábola, [1985] 2011.

BORTONI-RICARDO, S. M. Um modelo para a análise sociolingu***ística do português do Brasil***. *In:* BAGNO, Marcos. *A linguística da norma*. São Paulo: Edições Loyola, [1998] 2002. p. 333-346.

CALDAS, J. A. *Noticia geral de toda esta Capitania da Bahia desde o seu descobrimento até o prezente anno de 1759*. Editado por Alícia Duhá Lose, Vanilda Salignac de Souza Mazzoni e Perla Peñailillo. 2. ed. Salvador: Memória e Arte; EDUFBA, [1759] 2017.

CONSULTA do Conselho Ultramarino ao rei D. Pedro II acerca do recrutamento de catequistas entre os negros forros da Bahia para doutrinar os negros da Costa da Mina. Lisboa, 23 de fevereiro de 1701. AHU-Baía, cx. 3, doc. 112 / AHU_ACL_CU_005, Cx. 3, D. 314.

DIAS, P. *Arte da lingua de Angola,* oferecida a Virgem Senhora N. do Rosario, Mãy & Senhora dos mesmos Pretos. Lisboa: Oficina de Miguel Deslandes, 1697.

FIGUEIREDO, L. Mulheres nas Minas Gerais. *In:* DEL PRIORE, M.; BASSANEZI, C. *História das mulheres no Brasil*. 7. ed. São Paulo: Contexto, 2004. p. 141-188.

FLEXOR, M. H. *Oficiais mecânicos na cidade do Salvador*. Salvador: Prefeitura Municipal de Salvador, 1974.

FLORENTINO, M. *Em costas negras*. São Paulo: Companhia das Letras, 1997.

FRAGOSO, J. Modelos explicativos da chamada economia colonial e a ideia de Monarquia pluricontinental: notas de um ensaio. *História*, São Paulo, v. 31, n. 2, p. 106-145, jul./dez. 2012.

FRAGOSO, J.; FLORENTINO, M.; FARIA, S. de C. *A economia colonial brasileira (séculos XVI-XIX)*. 3. ed. São Paulo: Editora Atual, 1998.

FREYRE, G. *Casa-grande & senzala*: formação da família brasileira sob o regime da economia patriarcal. Rio de Janeiro: Nova Aguilar, [1933] 2002.

FURTADO, C. *Formação econômica do Brasil*. São Paulo: Companhia das Letras, [1958], 2007.

GALVES, C.; LOBO, T. *O português escrito por afro-brasileiros no século XIX*: as atas da Sociedade Protetora dos Desvalidos. Salvador: EDUFBA, 2019.

GUY, G. R. On the nature and origins of popular Brazilian Portuguese. *In*: *Estudios sobre español de América y Lingüistica Afroamericana.* Bogotá: Instituto Caro y Cuervo, 1989. p. 227-245.

HOLANDA, S. B. de. *História geral da civilização brasileira*. 7. ed. São Paulo: DIFEL, 1985. t. 1, v. 1 e v. 2.

HOUAISS, A. *O português no Brasil*. 3. ed. Rio de Janeiro: Revan, [1985] 1992.

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

LIVRO das Denunciações da Bahia que se fizerão na Visitação do Santo Officio à Cidade Salvador da Bahia de Todos os Santos do Estado do Brasil, no anno de 1618. Introdução de Rodolfo Garcia. Annaes da Bibliotheca Nacional do Rio de Janeiro, v. XLIX, 1936.

LOBO, T.; MACHADO FILHO, A. V. L.; MATTOS E SILVA, R. V. Indícios de língua geral no sul da Bahia na segunda metade do século XVIII. *In:* LOBO, T.; RIBEIRO, I.; CARNEIRO, Z.; ALMEIDA, N. *Para a história do português brasileiro*: novos dados, novas análises. Coleção *Para a história do português brasileiro*, v. 6. Salvador: EDUFBA, 2006. p. 609-630.

LOBO, T.; OLIVEIRA, K. África à vista: dez estudos sobre o português escrito por africanos no Brasil do século XIX. Salvador: EDUFBA, 2009

LOBO, T.; OLIVEIRA, K. Escrita liberta: letramento de negros na Bahia do século XIX. *In:* CASTILHO, A. T. de; MORAIS, Maria Aparecida Torres *et al. Descrição, história e aquisição do português brasileiro*. São Paulo: FAPESP; Campinas: Pontes/FAPESP, 2007. p. 437-460.

LOBO, T.; SARTORI, A. Aos olhos do Santo Ofício: difusão social da escrita nas capitanias da Bahia, Pernambuco, Itamaracá e Paraíba em finais de Quinhentos. *In:* CALLOU, D.; LOBO, T. *História social do português brasileiro*: da história social à história linguística. Coleção *História do português brasileiro*, v. 9, Coordenação Geral de Ataliba T. de Castilho. São Paulo: Contexto, 2020. p.182-227.

LOBO, T.; SARTORI, A.; MOTA, R. S. O aporte das fontes inquisitoriais para uma história da difusão social da leitura e da escrita no Brasil. *Caderno de estudos linguísticos*, Campinas, n. 58, p. 277-298, 2016.

LOBO, T., SARTORI, A.; SOUZA, P. D. dos S. O sequestro da História? Discurso oficial e práticas de ensino sobre a história linguística do Brasil e a historicidade do português brasileiro. *LaborHistórico*, Rio de Janeiro, v. 7, número especial, p. 282-310, 2021.

LOPES, C. R. dos S. *A norma brasileira em construção*: fatos linguísticos em cartas pessoais do século XIX. Rio de Janeiro: UFRJ, Pós-Graduação em Letras Vernáculas: FAPERJ, 2005.

LUCCHESI, D. *Língua e sociedade partidas*: a polarização sociolinguística no Brasil. São Paulo: Contexto, 2015.

LUCCHESI, D. O conceito de 'transmissão linguística irregular' e o processo de formação do português do Brasil. *In*: RONCARATI, Cláudia; ABRAÇADO, Jussara. *Português brasileiro: contacto linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003. p. 272-284.

LUCCHESI, D. Variação e norma: elementos para uma caracterização sociolinguística do português do Brasil. *Revista Internacional de Língua Portuguesa*, n. 12, p. 17-28, 1994.

MARCUSSI, A. A. O dever catequético: a evangelização dos escravos em Luanda nos séculos XVII e XVIII. *Revista 7 Mares*, Niterói, n. 2, p. 64-79, 2013.

MATTOS E SILVA, R. V. Português brasileiro: raízes e trajetórias (para a construção de uma história). *Discursos*, n. 3, 75-92, 1993.

MATTOS E SILVA, R. V. Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa. *In*: ALKMIM, T. M. *Para a história do português brasileiro*: novos estudos. Coleção *Para a história do português brasileiro*, v. 3. São Paulo: Humanitas, 2002. p. 443-464.

MATTOS E SILVA, R. V. Uma compreensão histórica do português brasileiro: velhos problemas revisitados. *Estudos linguísticos e literários*, Salvador, n. 25/26, 2000a, p. 251-283.

MATTOS E SILVA, R. V. Uma interpretação para a generalizada difusão da língua portuguesa no território brasileiro. *Gragoatá*, Niterói, n. 9, p. 11-27, 2000b.

MATTOSO, K. M. Q. *Ser escravo no Brasil*. 3. ed. São Paulo: Brasiliense, 1990.

MONTEIRO, J. M. 2001. *Tupis,* tapuias *e historiadores*: estudos de história indígena e indigenismo. Tese (Livre Docência em Etnologia) - Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 2001.

NARO, A.; SCHERRE, M. Sobre as origens do português popular do Brasil. *DELTA*, São Paulo, v. 9, número especial, p. 437-454, 1993.

NARO, A.; SCHERRE, M. *Origens do português brasileiro*. São Paulo: Parábola, 2007.

NASCIMENTO, A. A. V. N. *Patriarcado e religião*: as enclausuradas clarissas do Convento do Desterro da Bahia, 1677-1890. Salvador: Conselho Estadual de Cultura, 1994.

OLIVEIRA, K. *Negros e escrita no Brasil do século XIX*: sócio-história, edição filológica de documentos e estudo linguístico. Tese (Doutorado em Letras) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2006.

PAGOTTO, E. Ebulição e sedentarização linguística: o lugar da economia de subsistência na formação do português brasileiro. *Diadorim*, Rio de Janeiro, v. 20, número especial, p. 53–63, 2018.

PETTER, M. Ampliando a investigação do *continuum* afro-brasileiro do português. *Papia:* Revista brasileira de estudos crioulos e similares, Brasília, n. 25, p. 305-317, 2015.

PETTER, M. As línguas africanas no Brasil: o papel do quimbundo na formação do português brasileiro. *In*: TEIXEIRA, E. P.; ARAÚJO, S. S. de F. *Diálogos entre Brasil e Angola*: o português d´aquém e d´além-mar. Feira de Santana: UEFS Editora, 2017. p. 21-48.

PETTER, M. Línguas africanas no Brasil. *In:* CARDOSO, S. A.; MOTA, J.; MATTOS E SILVA, R. V. *Quinhentos anos de história linguística do Brasil*. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. p. 359-380.

PETTER, M. O *continuum* afro-brasileiro do português. *In:* GALVES, C.; GARMES, H.; RIBEIRO, F. do R. África-Brasil: caminhos da língua portuguesa. Campinas: Editora da UNICAMP, 2009. p. 151-173.

PETTER, M.; CUNHA, A. S. As línguas africanas no Brasil. *In:* PETTER, M. *Introdução à linguística africana*. São Paulo: Contexto, 2016. p. 221-250.

PRADO JR., C. *Formação do Brasil contemporâneo – Colônia*. São Paulo: Companhia das Letras, [1946], 2011.

PRIMEIRA visitação do Santo Officio às partes do Brasil pelo licenciado Heitor Furtado de Mendonça: Denunciações de Pernambuco 1593-1595. São Paulo: Ed. Paulo Prado, 1929.

PRIMEIRA visitação do Santo Officio às partes do Brasil pelo licenciado Heitor Furtado de Mendonça: Confissões da Bahia 1591-92. São Paulo: Ed. Paulo Prado, 1922.

PRIMEIRA visitação do Santo Officio às partes do Brasil pelo licenciado Heitor Furtado de Mendonça: Denunciações da Bahia 1591-93. São Paulo: Ed. Paulo Prado, 1925.

PRIMEIRA visitação do Santo Ofício às partes do Brasil: Confissões de Pernambuco 1594-1595. Recife: Universidade Federal de Pernambuco. Ed. José Antônio Gonçalves de Melo, 1970.

PRIMEIRA visitação do Santo Ofício às partes do Brasil: Denunciações e Confissões de Pernambuco 1593-1595. Recife: FUNDARPE, 1984.

QUARTO Livro das Denunciações da Primeira Visitação do Santo Ofício da Inquisição do Brasil, a qual fez o licenciado Heitor Furtado de Mendonça, por especial Comissão de Sua Alteza (1593-1595). In: ARQUIVO NACIONAL DA TORRE DO TOMBO, mf. 3077. Cópia microfilmada. Disponível em: http://digitarq.dgarq.gov.pt. Acesso em: 30 maio 2011.

RODRIGUES, A. D. As línguas gerais sul-americanas. *Papia*: Revista brasileira de estudos crioulos e similares, São Paulo, v. 4, 2, p. 6-18, 1996.

RODRIGUES, A. D. Tupi, tupinambá, línguas gerais e português do Brasil. *In:* NOLL, V.; DIETRICH, W. *O português e o tupi no Brasil.* São Paulo: Contexto, 2010. p. 27-47.

SANTOS, Fabricio Lyrio. *Da catequese à civilização*: colonização e povos indígenas na Bahia. Cruz das Almas: Editora UFRB, 2014.

SANTOS, M. R. A. dos. *Fronteiras do sertão baiano*: 1640-1750. Tese (Doutorado em História Social). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2010.

SARTORI, A. *Cartas de amor na Bahia do século XX*: normas linguísticas, práticas de letramento e tradições do discurso epistolar. Dissertação (Mestrado em Letras e Linguística) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2010.

SARTORI, A. *Pela pena do Santo Ofício*: difusão social da escrita nas Capitanias de Pernambuco, Itamaracá e Paraíba em finais de Quinhentos. Tese (Doutorado em Língua e Cultura) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2016.

SEGUNDA Visitação do Santo Ofício às partes do Brasil pelo Inquisidor e Visitador Marcos Teixeira: Livro das Confissões e Ratificações da Bahia, 1618-1620. Introdução de Eduardo de Oliveira França e *Sônia Siqueira. Anais do Museu Paulista, t. XVII, 1963.*

SILVA, L. G. Sobre a 'etnia crioula': o Terço dos Henriques e seus critérios de exclusão na América portuguesa do século XVIII. *In:* VENÂNCIO, R. P.; GONÇALVES, A. L.; CHAVES, C. M. das G. *Administrando impérios*: Portugal e Brasil nos séculos XVIII e XIX. Belo Horizonte: Fino Traço, 2012. p. 71-96.

SILVA NETO, S. da. *Introdução ao estudo da língua portuguesa no Brasil*. Rio de Janeiro: Presença/Instituto Nacional do Livro, [1950] 1986.

SIQUEIRA, S. *A Inquisição portuguesa e a sociedade colonial*. São Paulo: Ática, 1978.

VIEIRA, A. Voto do padre Antônio Vieira sobre as dúvidas dos moradores de São Paulo acerca da administração dos índios. *In:* VIEIRA, A. *Vozes saudosas, da eloquência, do zelo, e eminente sabedoria...* Lisboa: Officina de Miguel Rodrigues, 1736. p. 141-166.

ZOREK, B. de M. *Modelos explicativos da economia colonial brasileira*: um problema historiográfico. 31º Simpósio Nacional de História, 2021, Rio de Janeiro. *Anais...* Rio de Janeiro: ANPUH, 2021. Disponível em: https://www.snh2021.anpuh.org/resources/anais/8/snh2021/1627586532_ARQUIVO_b26ed4f4b36d3ac178feb89590ddc317.pdf. Acesso em: 16 ago. 2024.

http://doi.org/10.5007/1984-8420.2025.e100380

O PORTUGUÊS NO BRASIL COLONIAL: REFLEXÕES SOBRE EDIÇÕES DE MANUSCRITOS HISTÓRICOS BRASILEIROS

PORTUGUESE IN COLONIAL BRAZIL: REFLECTIONS ON EDITIONS OF BRAZILIAN HISTORICAL MANUSCRIPTS

Mariana Fagundes de Oliveira Lacerda | Lattes | marianafagundes@uefs.br
Universidade Estadual de Feira de Santana

Alícia Duhá Lose | Lattes | alicialose@gmail.com
Universidade Federal da Bahia

Zenaide de Oliveira Novais Carneiro | Lattes | zoncarneiro@uefs.br
Universidade Estadual de Feira de Santana

Resumo: Abordamos, neste artigo, o diálogo necessário entre a Paleografia, a Filologia e a Linguística Histórica. Esse diálogo se intensifica no projeto Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS), desenvolvido pelo Núcleo de Estudos de Língua Portuguesa (NELP) da Universidade Estadual de Feira de Santana (UEFS), a partir de 2017, quando iniciamos a fase 2 (colonial), atendendo aos objetivos do projeto Guarda-Chuva CE-DOHS: um *corpus* para uma caracterização linguístico-gramatical do português brasileiro - fase colonial e fase pós-colonial, criado em 2010. Neste momento, o CE-DOHS fase 2 já dispõe de coleções documentais em edição semidiplomática, a partir de rigoroso tratamento paleográfico-filológico; além disso, os primeiros estudos de exploração linguístico-gramatical dessa documentação encontram-se realizados. Como sabemos, ainda não é possível atestar a existência do português do Brasil no período colonial brasileiro; mas investigar como se configurava o português no Brasil nesse período pode contribuir para uma reconstrução, por aproximação, da história social linguística do português brasileiro. Nosso objetivo, portanto, no presente texto, é apresentar o ingente trabalho de organização de *corpora* históricos, numa perspectiva interdisciplinar, da maneira como vem sendo realizado no âmbito do projeto CE-DOHS, em parceria com o *Modus Scribendi* – Grupo de Pesquisas Paleográficas, Filológicas e Históricas, da Universidade Federal da Bahia (UFBA); dessa organização dependem descrições linguísticas de base empírica, as quais promovem a discussão de questões teóricas solidamente fundamentadas.

Palavras-Chave: Português no Brasil; *Corpus* colonial; Tratamento paleográfico-filológico.

Abstract: In this article, we address the necessary dialogue between Paleography, Philology and Historical Linguistics, a dialogue that intensifies in the Electronic Corpus of Historical Documents of the Sertão (CE-DOHS) project, of the Portuguese Language Studies Center (NELP), of State University of Feira de Santana (UEFS), from 2017, when we started phase 2 (colonial), meeting the objectives of the Umbrella CE-DOHS project: a corpus for a linguistic-grammatical characterization of Brazilian Portuguese - colonial phase and post-colonial phase, created in 2010. At this moment, CE-DOHS phase 2 already has documentary collections in a semi-diplomatic edition, based on rigorous paleographic-philological treatment, as well as the first linguistic-grammatical exploration studies of this documentation found carried out. As we know, it is not yet possible to attest to the existence of Brazilian Portuguese in the Brazilian colonial period; but investigating how Portuguese was configured in Brazil, during this period, can contribute to an approximation reconstruction of the linguistic social history of Brazilian Portuguese. Our objective, therefore, in this text, is to present the enormous work of organizing historical corpora, from an interdisciplinary perspective, as it has been carried out within the scope of the CE-DOHS project, in partnership with Modus Scribendi – Paleographic Research Group, Philological and Historical, from the Federal University of Bahia (UFBA); this organization depends on empirically based linguistic descriptions, which promote the discussion of solidly based theoretical issues.

Keywords: Portuguese in Brazil; Colonial corpus; Paleographic-philologicaltreatment.

1 Introdução ou sobre o trabalho para muitas mãos...

Desde 1997, uma grande equipe de pesquisadores – que só tem aumentado ao longo dos anos dedica-se, no âmbito do Projeto Nacional para a História do Português Brasileiro (PHPB), à reconstrução, por aproximação, da história social linguística do PB (Mattos e Silva, 2004), atuando em diferentes agendas de trabalho (Castilho, 2024). Entre elas, a agenda de *corpus* diacrônico, com a qual o projeto Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS)[1-2] – do Núcleo de Estudos de Língua Portuguesa (NELP)[3] da Universidade Estadual de Feira de Santana (UEFS) – tem contribuído de forma bastante significativa.

[1] Disponível em: http://www.uefs.br/cedohs/. Link de acesso ao Grupo no CNPq: http://dgp.cnpq.br/dgp/espelhogrupo/481138.

[2] Marquilhas e Hendrickx (2016) discorrem sobre "Avanços nas Humanidades Digitais" e citam, entre os projetos de *corpora* históricos reconhecidos internacionalmente, o CE-DOHS (p. 21). O capítulo faz parte do livro *Manual de Linguística Portuguesa*, organizado pelas professoras Ana Maria Martins e Ernestina Carrilho, publicado pela De Gruyter.

[3] Disponível em: https://nelp.uefs.br/.

Como afirmou Mattos e Silva (2004, p. 61), a busca pela reconstrução da história do português brasileiro "será trabalho para muitas mãos, durante muito tempo". E esse trabalho inclui, fundamentalmente, a constituição de *corpus*, em que buscamos os dados que iluminam a linguística teórica. A mão do paleógrafo e a mão do filólogo,juntas à mão do historiador das línguas, garantem a autenticidade dos materiais, dos *corpora*, cujo uso "permite a realização de descrições linguísticas de base empírica e promove, com isso, a discussão de questões teóricas solidamente fundamentadas" (Bacelar do Nascimento, 2004, p. 1).

Em publicação de 2008a, Mattos e Silva ressalta que

> não se pode nem se deve utilizar qualquer edição de texto do passado para a análise histórico-diacrônica: a edição tem de ser feita com rigor filológico e com objetivo claro de servir a estudos linguísticos; há edições úteis ao historiador ou ao estudioso da literatura ou chamado grande público, mas que, contudo, não devem ser usadas para estudos de história linguística (Mattos e Silva, 2008a, p. 15).

Como vemos, esse cruzamento de olhares – diplomático, paleográfico, filológico, sócio-histórico e linguístico – no trabalho de constituição de *corpus* em Linguística Histórica *Stricto Sensu* (Mattos e Silva, 2008) é essencial, condição *sine qua non* para o estudo da mudança linguística no tempo real. A história das línguas, como ressalta Mattos e Silva (2008b, p. 52), citando o historiador José Mattoso, é feita, similar à História, da "incomensurável mistura de elementos de toda espécie".

> O historiador medieval José Mattoso, no seu livro "A escrita da história: teoria e métodos" diz: "Prefiro... tentar descobrir as harmonias resultantes dessa fantástica sinfonia que é a História, feita da incomensurável mistura de elementos de toda espécie, tão dispersos e contraditórios com a própria vida, mas de cuja rede infinitamente complexa e fascinante procuram os eixos, os encontros e desencontros" (1988: p. 10). O que disse o historiador para a História aplica-se, a meu ver, à história da(s) língua(s) (Mattos e Silva, 2008b, p. 52).

Pisando nesse terreno complexo e fascinante da história das línguas, abordamos, neste artigo, um diálogo necessário entre a Paleografia, a Filologia e a Linguística Histórica. Esse diálogo, que se intensifica no projeto CE-DOHS/NELP/UEFS a partir de 2017, quando iniciamos a fase 2 (colonial), atende aos objetivos do projeto Guarda-Chuva CE-DOHS: um *corpus* para uma caracterização linguístico-gramatical do portu-

guês brasileiro - fase colonial e fase pós-colonial, criado em 2010 por Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda. Abordamos, ainda, a importância da pesquisa em rede, como a que vem sendo desenvolvida no âmbito do PHPB, há 28 anos, tendo como coordenador Ataliba Teixeira de Castilho; e no âmbito do NELP, criado em 1997 por Zenaide de Oliveira Novais Carneiro e Norma Lucia Fernandes de Almeida e coordenado por Mariana Fagundes de Oliveira Lacerda[4], desde 2017.

Nosso objetivo, portanto, no presente texto, é apresentar o ingente trabalho de organização de *corpora* históricos, numa perspectiva interdisciplinar, da maneira como vem sendo realizado no âmbito do projeto CE-DOHS, em parceria com o *Modus Scribendi* – Grupo de Pesquisas Paleográficas, Filológicas e Históricas da Universidade Federal da Bahia (UFBA), coordenado por Alícia Duhá Lose e Lívia Borges Souza Magalhães.

2 OPHPB e a enormidade do empreendimento

Em janeiro de 2024, por ocasião do XX Congresso Internacional da Associação Brasileira de Linguística e Filologia da América Latina (ALFAL), no âmbito do Projeto 3 - História do Português Brasileiro, da Europa à América, realizado na Faculdade de Humanidades e Arte da Universidade de Concepción, no Chile, o professor Ataliba Teixeira de Castilho (USP/UNICAMP) proferiu a Conferência de encerramento, intitulada *Historiando a Língua Portuguesa no Novo Mundo.* No texto, ele afirma:

> A Linguística Histórica perdeu espaço entre nós durante o período mais forte de atuação do Estruturalismo e do Gerativismo, movimentos basicamente a-históricos. É verdade, entretanto, que, mesmo no interior dessas teorias, algumas vozes anunciavam a presença da "velha senhora", como gosto de qualificar a Linguística Histórica. (...). No Brasil, foi profética a voz de Tarallo (1984), que proclamou o renascimento da "fênix", excelente metáfora para uma ciência que retorna sempre. Nos anos 1980, sem que tivesse havido uma combinação prévia, três respeitadas linguistas fizeram renascer a Linguística Histórica no domínio da língua portuguesa: a) Clarinda Maia, da Universidade de Coimbra; b) Rosa Virgínia Mattos e Silva, da Universidade Federal da Bahia; c) Mary Kato, da Universidade Estadual de Campinas. O quadro desenhado acima me motivou a propor a criação de um projeto coletivo de pesquisas, tendo

[4] Conferir linha do tempo do NELP em: https://nelp.uefs.br/o-nelp/historico/. O NELP/UEFS completou 25 anos em 2023, sendo considerado um centro de Linguística Histórica no Semiárido baiano. O PHPB, desde sua criação, em 1997, tem contado com a equipe de pesquisadores do NELP e recebido contribuições significativas às agendas de trabalho. Conferir Lacerda (2024), NELP 25 anos https://nelp.uefs.br/nelp-25-anos/ e Livros Publicados/Coleções de *Corpora* e Coletâneas Linguísticas https://nelp.uefs.br/livros-publicados/.

> por objetivo, inicialmente, historiar o Português de São Paulo (Castilho, 1998) (Castilho, 2024, p. 2)[5].

O Projeto para a História do Português Paulista (PHPP) realizou o seu I Seminário em 1997, e, nessa oportunidade, por iniciativa dos participantes, o projeto regional assumiu uma dimensão nacional, criando-se o Projeto para a História do Português Brasileiro (PHPB). Castilho (2024, p.3) destaca que "O 'para' foi uma sugestão oportuna de Mattos e Silva, que nos advertia sempre sobre a enormidade do empreendimento". Um projeto nacional, com 26 anos de existência, que se acha atualmente entre os maiores projetos em rede de pesquisa da Linguística brasileira.

Sem dúvida, um empreendimento enorme a reconstrução, por aproximação, da história social linguística do PB. No texto supracitado de 2024, Castilho refere-se às seguintes agendas de trabalho do PHPB:

> Desde sua criação, o PHPB foi se consolidando como um projeto nacional, desenvolvendo as seguintes agendas de trabalho: a) Corpus diacrônico; b) História social; c) Mudança gramatical; d) Linguística diacrônica do texto; e) Tradições discursivas; f) Semântica diacrônica; g) História do léxico. Nos primeiros 10 anos, as atividades se concentraram na organização do *corpus* diacrônico e na mudança gramatical, a partir de perspectivas teóricas mais sócio-funcionais ou formais. Com a ampliação das equipes e a necessidade de consolidar os resultados do projeto, novas temáticas foram incorporadas ao Projeto (Castilho, 2024, p. 4).

A agenda de *corpus* diacrônico, desde sempre um grande desafio, trata-se de uma das principais agendas do projeto, tendo a Comissão de *Corpus*, em 1º de junho de 2010, definido "o conjunto do *Corpus* Mínimo Comum Diacrônico (séculos XVIII, XIX e XX)". Segundo Castilho (2024, p. 7)[6],"[a] opção foi compor três conjuntos de textos divididos da seguinte maneira: *Corpus* comum mínimo – manuscritos; *Corpus* comum mínimo – impressos; *Corpus* comum diferencial".

A Bahia, como abordamos nas Seções3, 4 e 5, adiante, está entre as equipes regionais do PHPB, destacando-se no trabalho de constituição de *corpus*, entre outras agendas. O Quadro 1 apresenta os membros das equipes regionais do projeto.

[5] O autor esclarece a menção às três respeitadas linguistas: "a) Clarinda Maia, da Universidade de Coimbra, com seus estudos sobre o Galego-Português: Maia (1986). Ver também Maia (2022). b) Rosa Virgínia Mattos e Silva, da Universidade Federal da Bahia, com seu estudo sobre o Português trecentista, a que se seguiu o *Programa de História do Português* (doravante, PROHPOR): Mattos e Silva (1989). c) Mary Kato, na Universidade Estadual de Campinas, juntamente com Fernando Tarallo, quando focalizaram o Português Brasileiro em trabalho apresentado em um congresso internacional: Kato; Tarallo (1988). Depois, eles deram início a um projeto de história do português brasileiro, de que surgiram várias teses e ensaios, já publicados: Roberts, Kato (1993); Galves, Kato e Roberts (2019)" (Castilho, 2024, p. 2).

[6] Plataforma de *Corpora* PHPB disponível em: https://sites.google.com/site/corporaphpb/home.

Quadro 1: Equipes regionais do PHPB.

Alagoas	Thiago Trindade (Universidade Federal de Alagoas).
Bahia	Pedro Daniel dos Santos Souza (Universidade do Estado da Bahia), em substituição a Tânia Lobo, que substituiu Rosa Virgínia Mattos e Silva.
Ceará	Aurea Suely Zavam (Universidade Federal do Ceará).
Mato Grosso	Elias Alves de Andrade (Universidade Federal do Mato Grosso).
Minas Gerais	Jânia Ramos (Universidade Federal de Minas Gerais).
Paraíba	Roseane Nicolau, em substituição a Camilo Rosa (Universidade Federal da Paraíba).
Paraná	Vanderci Aguilera (Universidade Estadual de Londrina).
Pernambuco	Cléber Alves de Ataíde, em substituição a Valéria Gomes (Universidade Federal Rural de Pernambuco), que substituiu Marlos de Barros Pessoa (Universidade Federal de Pernambuco).
Rio de Janeiro	Maria Eugênia Duarte, em substituição a Dinah M. I. Callou (Universidade Federal do Rio de Janeiro).
Rio Grande do Norte	Alessandra Castilho da Costa, em substituição a Marco Antonio Martins (Universidade Federal do Rio Grande do Norte).
Santa Catarina	Izete L. Coelho, em substituição a Gilvan Muller de Oliveira (Universidade Federal de Santa Catarina).
São Paulo	Manoel Mourivaldo Santiago Almeida (Universidade de São Paulo), em substituição a Clélia C. S. Jubran (Universidade Estadual Paulista / São José do Rio Preto), que substituiu Ataliba T. de Castilho (Universidade de São Paulo).
Sergipe	Sandro Márcio Drumond Alves Marengo, em substituição a Renata Ferreira Costa (Universidade Federal de Sergipe).

Fonte: Elaborado pelas autoras, com base em Castilho (2024, p. 5).

As equipes regionais do PHPB se reúnem em seminários nacionais desde 1998; nesse ano, o seminário ocorreu em São Paulo, e o último, em 2019, ocorreu em Sergipe, em comemoração aos 20 anos do projeto.

Os seminários nacionais e regionais do PHPB produziram atas[7]. Por exemplo, no volume IV (Duarte; Callou, 2002), sobre *corpus*, têm destaque, entre outros, os seguintes temas:

a) O controle sócio-histórico da documentação, tanto quanto possível;
b) A identificação do perfil biográfico de quem escreve;

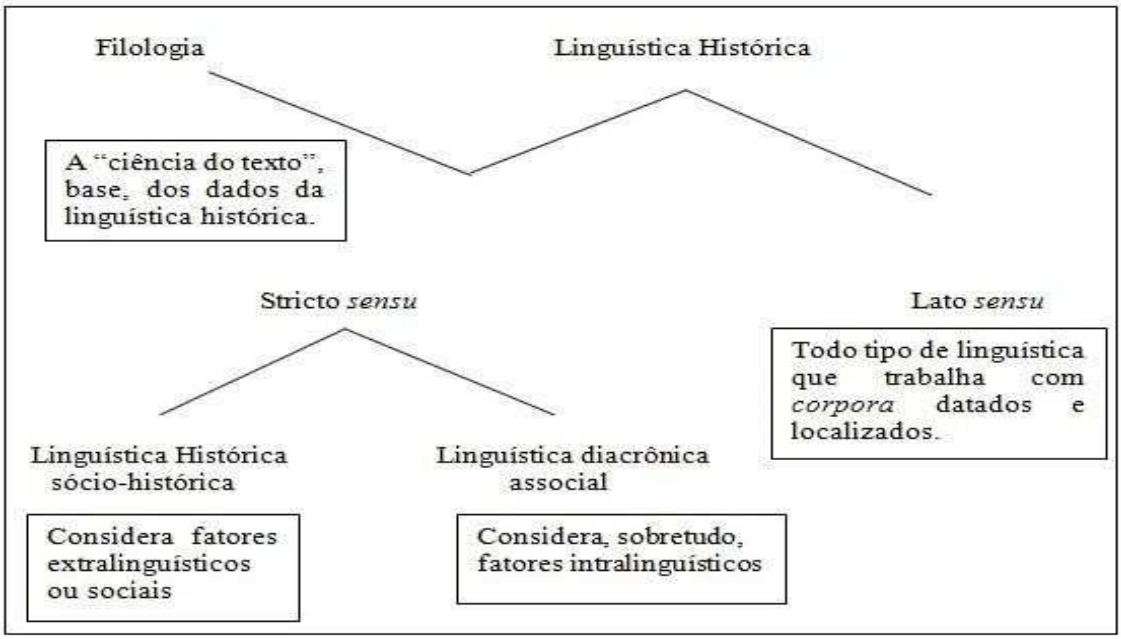
[7] Produziram-se, até o presente, 16 atas, além dos 10 volumes do PHPB, publicados pela editora Contexto.

c) O produtor do texto (autenticidade do material e grau de domínio da técnica de escrever de cada autor/*escriptor*);
d) O contexto de produção do texto (as relações sociais que cercam o ato de escrever na época em foco).

Os controles sócio-histórico, paleográfico e filológico, no trabalho de constituição de *corpus* diacrônico no âmbito do PHPB, sempre foram tidos como condição *sine qua non*. Questões sobre a autenticidade dos documentos, sobre sua leitura, sobre sua transcrição, entre outras, estavam, desde o princípio, na *mesa* do PHPB.

Demonstrando a dependência da Linguística Histórica *Stricto Sensu* em relação à Filologia, disciplinas parceiras na coletividade científica[8], Mattos e Silva (2008a) constrói um esquema apresentado na Figura 1.

Figura 1: Parceria entre a Linguística Histórica *Stricto Sensu* e a Filologia



Fonte: Mattos e Silva (2008a, p. 10).

Sem dúvida, olhares entrelaçados, tendo em vista edições que satisfaçam as necessidades de uma Linguística Histórica empírica (Maia, 2012, 538).

[8] Conferir a discussão sobre a parceria entre Linguística Histórica e Filologia feita por Lose (2017) e Lose e Souza (2022).

3 A agenda de *corpus* diacrônico na Bahia: como formigas obreiras

O PHPB enfrentou grandes desafios no trabalho de constituição de *corpus* diacrônico, especialmente na primeira década. Segundo Lobo,

> [a] segunda metade da década de 1990 inaugura, nos estudos histórico-diacrônicos do português brasileiro, uma fase que se caracterizará pela laboriosa tarefa de construção de uma filologia de textos escritos no Brasil, ponto de partida incontornável para a descrição e análise das mudanças linguísticas que foram configurando o chamado português brasileiro. A inexistência de uma tradição filológica relativa ao período compreendido entre os séculos XVII e o XIX era, então, observável não apenas no Brasil, mas também em Portugal (Lobo, 2009, p. 307-308).

Dado esse cenário, Lobo (2009), que foi uma das coordenadoras do PHPB-Bahia, esclarece que, na primeira década, a constituição de *corpus* diacrônico na Bahia passou por algumas fases: a fase ingênua; a fase da profissionalização; e a fase dos profissionais.

A fase ingênua é caracterizada pela edição de textos escritos no Brasil, sem, ainda, o conhecimento necessário sobre os arquivos, sobre os fundos documentais, sobre a caracterização sociolinguística dos informantes (Lobo, 2009).

A fase da profissionalização remete ao período em que

> n**ão** bastava editar textos escritos no Brasil e, desavisadamente, tomá-los como representativos de um português brasileiro, já que, no Brasil do passado, a pena esteve nas mãos de portugueses, "brancos" brasileiros, africanos, índios e imigrantes – todos, com maior ou menor perícia, escrevendo em português. Dentre as diversas questões a enfrentar, uma sobressaía-se: era preciso identificar, o mais precisamente possível, "quem escreveu" (Lobo, 2009, p. 312)

Por último, a fase dos profissionais é entendida por Lobo (2009, p. 314) desta forma: "Aqui se alcança um grau de maturidade no rastreamento de fontes primárias as mais diversas. Despontam os profissionais".

Como ressalta Mattos e Silva (2008b, p. 39-40), a constituição de *corpus* tem início na prospecção documental, um trabalho de "formigas obreiras" pelos arquivos, em vários pontos do Brasil:

> A Filologia se fez presente no PHPB: no Seminário de Campos do Jordão, um grupo, coordenado por Heitor Megale, estabeleceu as normas para as edições a serem utilizadas no *corpus* documental que seria a base das análises lingüísticas. Começou-se com textos impressos no século XIX, sobretudo os jornais, que, no Brasil, se multiplicaram naquele século, uma vez que, só

> a partir de 1808, foi permitida pela metrópole a imprensa no Brasil. Dessa pesquisa em jornais, publicou-se o livro *E os preços eram commodos... Anúncios de jornais brasileiros. Século XIX*, sob a responsabilidade de Marymárcia Guedes e Rosane de Andrade Berlinck, corria o ano 2000. Depois foi-se para os Arquivos Brasileiros, públicos ou privados. A equipe da Bahia publicou, em 2001, Cartas Baianas Trecentistas, sob a responsabilidade da pesquisadora Tânia Lobo e com a colaboração dos bolsistas de Iniciação Científica – Klebson Oliveira, Permínio Ferreira e Uilton Oliveira. Outros *corpora* vêm sendo constituídos, como base, principalmente, de dissertações e teses, como, por exemplo – *As cartas do Século XIX* para as freiras do Desterro, base da tese de Tânia Lobo (inédita), em que a autora distingue as cartas de brasileiros das dos portugueses; as *Cartas* de Afrânio Gonçalves de comerciantes portugueses no Brasil, que também editou outros documentos para confrontar com as *Cartas*; as atas e outros documentos da Sociedade Protetora dos Desvalidos, editadas por Klebson Oliveira, no seu Mestrado, e ampliadas, em número, na sua tese de Doutorado, e assim por diante. Verificou-se, no correr da pesquisa, que era o século XIX o mais fértil nos Arquivos Brasileiros, diminuindo a incidência de documentos quando se recua no tempo. Como formigas obreiras, têm-se pesquisado Arquivos em vários pontos do Brasil (Mattos e Silva, 2008b, p. 39-40).

Novos olhares, especialmente da Paleografia, sobre os *corpora* constituídos no âmbito do PHPB Bahia[9] ao longo dos anos, bem como no âmbito de outras equipes regionais do PHPB nacional, poderão lançar luz sobre aspectos específicos dos documentos, numa rede de pesquisa colaborativa, produtiva e que prima pelo rigor e pela honestidade científicos. Como bem destacou Maia (2012, p. 538),

> [s]ob o ponto de vista qualitativo, a autenticidade é uma exigência decisiva: se se pretende utilizar os materiais para o estudo de história da língua, as edições devem satisfazer as necessidades de uma linguística histórica empírica e, por esse motivo, devem reflectir fielmente as características linguísticas dos manuscritos, uma vez que nelas se reflectem as marcas da variação da língua da época (Maia, 2012, p. 538).

Portanto, é sempre necessário, e desejável, que os materiais para exploração de dados linguísticos, na Linguística Histórica, recebam os tratamentos diplomático, paleográfico e filológico com o devido rigor científico, além de ser realizado, tanto quanto possível, controle sócio-histórico rígido, de modo que os estudos linguísticos realizados com esses *corpora* tenham confiabilidade. Isso porque a confiabilidade dos dados depende da autenticidade dos materiais.

A promoção dessa rede de pesquisa, entre historiadores das línguas, paleógrafos e

[9] O Grupo de Pesquisa *Modus Scribendi* – Grupo de Pesquisas Paleográficas, Filológicas e Históricas (CNPq-UFBA), por exemplo, está revisitando a documentação da Sociedade Protetora dos Desvalidos (SPD), com que trabalhou Oliveira (2006), trazendo novos documentos e novos elementos para discussão.

filólogos, em um diálogo respeitoso, considerando os diferentes contextos de produção de pesquisa, espaços e tempos, favorece o melhor uso, em Linguística Histórica, dos dados disponíveis. O trabalho interdisciplinar, coletivo, que, nas palavras de Mattos e Silva[10], "desenvolve a solidariedade exigente e permite o intercâmbio fraterno entre os saberes de cada um", vai formando pesquisadores com maior grau de maturidade em qualquer campo do conhecimento.

4 CE-DOHS: fontes para a história do português brasileiro

Faz parte do NELP/UEFS, da Bahia, como dissemos na introdução deste texto, o projeto guarda-chuva intitulado CE-DOHS: um *corpus* para uma caracterização linguístico-gramatical do português brasileiro - fase colonial e fase pós-colonial, criado em 2010 e coordenado por Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda:

> Trata-se de um projeto guarda-chuva – vinculado ao Núcleo de Estudos de Língua Portuguesa (NELP), da UEFS, com 25 anos de existência – que se dedica à constituição de *corpora* históricos de português brasileiro (fase colonial e fase pós-colonial), a partir de cuidadosa prospecção documental e estabelecendo rigorosos controles sócio-histórico, paleográfico e filológico, tendo em vista colaborar com os estudos sobre a história social linguística do PB (Mattos e Silva, 2004), em parceria com o Projeto Nacional para a História do Português Brasileiro (PHPB) e com o Programa para a História da Língua Portuguesa (PROHPOR). Este projeto, que tem início em 2010, trata-se da continuação do projeto, intitulado Contribuições para a Constituição de um Banco de Textos e de um Banco de Dados para o Estudo da História do Português do Brasil, do Século XVII ao XX e iniciado no NELP/UEFS, em 1997, sob a coordenação de Ilza Ribeiro. Vê-se, daí, portanto, a longa trajetória que tem o NELP/UEFS no que diz respeito a estudos sobre o PB nas fases colonial e pós-colonial, firmando-se como uma das principais instituições responsáveis por essas agendas no Brasil, tendo reconhecimento internacional. (...) São significativos os resultados alcançados até o presente, especialmente no âmbito de dissertações de mestrado e teses de doutorado. Consistem em análises sobretudo de natureza qualitativa (especialmente para o período colonial, haja vista a raridade dos documentos e os dados escassos), as quais dependem crucialmente de interpretações de base teórica, refinadas pela Paleografia (Carneiro; Lacerda, 2024, Início).

Segundo Ribeiro (1997), no seu projeto – motivador da criação do NELP –, intitulado Contribuições para a Constituição de um Banco de Textos e de um Banco de Dados para o Estudo da História do Português do Brasil, do Século XVII ao XX, com os direcionamentos propostos e assumidos pelo PHPB – entre os quais a constituição de um *corpus* diacrônico do PB –,

[10] Citação de Mattos e Silva no *menu* Histórico, do site do PROHPOR UFBA.

> está bem pensada e planejada uma reconstrução dos processos lingüísticos e socioculturais condicionadores da formação do português no Brasil, ou português brasileiro, a partir de uma organização sistemática e de estudos lingüísticos de diferentes documentos (estilos diversos) escritos no Brasil, em diversas regiões e em diferentes épocas de sua história. Houaiss (1995: 31) diz que "não preenchemos ainda os requisitos da pesquisa e conhecimento com que se possa elaborar uma história da língua portuguesa no Brasil". Certamente essas pesquisas mudarão o estado da questão (Ribeiro, 1997, p. 3-4).

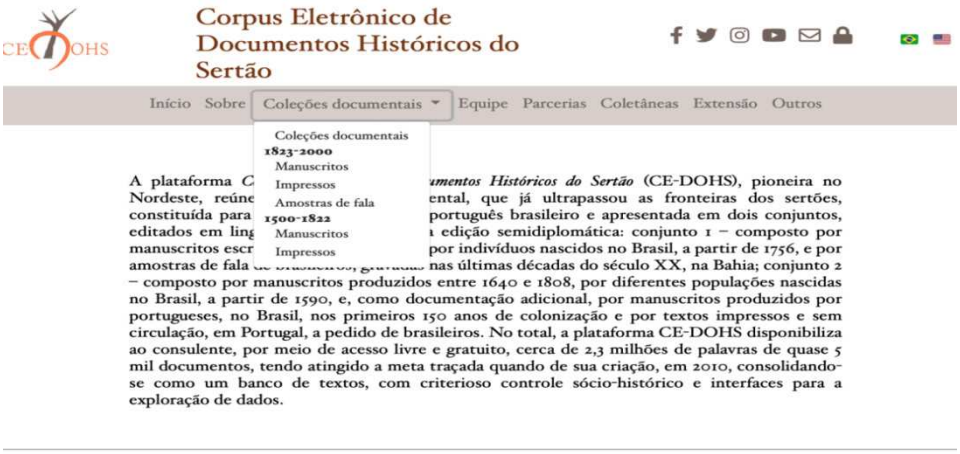
De fato, muito já foi feito, ao longo de mais de duas décadas, pela equipe doPHPB nacional, e importantes contribuições lançaram luz à história do PB.

Atendendo à agenda de *corpus* diacrônico, o PHPB, ao longo de sua história, reuniu documentação sobretudo do período pós-colonial, dos séculos XIX e XX. O CE-DOHS/NELP/UEFS, a partir de 2017 – já disponibilizando rico material da fase 1 (pós-colonial) –, começou a investir na formação de *corpus* ilustrativo do período colonial, intensificando a parceria com a Paleografia e com a Filologia para o tratamento dos textos de épocas ainda mais recuadas[11]; conferir figura 2. Destaca-se a parceria com o Grupo de

[11] Esse estreitamento no diálogo da Linguística Histórica com a Paleografia e a Filologia tem sido promovido pelo CE-DOHS e pelo *Modus Scribendi* também em eventos acadêmicos, como: (i) VI Congresso Internacional de Linguística Histórica (Faculdade de Letras da Universidade de Coimbra, 2023) – "Edições de manuscritos históricos brasileiros: a série Documentos para a história linguística do Brasil colônia", por Alícia Duhá Lose, Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro; (ii) XX Congresso Internacional da Associação Brasileira de Linguística e Filologia da América Latina - Projeto 3 da ALFAL, História do Português Brasileiro, da Europa à América (Faculdade de Humanidades e Arte da Universidade de Concepción, Chile, 2024) – "O português no Brasil colonial: reflexões sobre as edições de manuscritos históricos brasileiros", por Alícia Duhá Lose, Zenaide de Oliveira Novais Carneiro, Mariana Fagundes de Oliveira Lacerda e Lara da Silva Cardoso; (iii) XI Seminário de Estudos Filológicos: Filologia e Ensino - Perspectivas Críticas e Editoriais (Universidade Estadual de Feira de Santana, Feira de Santana, 2024) – "O CE-DOHS na rede de pesquisa PHPB: recuando ao período colonial", por Mariana Fagundes de Oliveira Lacerda, Zenaide de Oliveira Novais Carneiro e Alícia Duhá Lose; (iv) Simpósio da Contemporaneidade à Antiguidade: temas em debate (Centro de Estudos Clássicos e Humanísticos da Universidade de Coimbra) – "Reflexões sobre edições de manuscritos históricos brasileiros: o cruzamento de olhares paleográfico e linguístico", por Alícia Duhá Lose, Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro; (v) IX Simpósio Mundial de Língua Portuguesa (Universidade da Madeira, Ilha da Madeira, Portugal, 2024) – "Fontes linguísticas para estudos do português no Brasil colonial: o exemplo do Testamento de André Vidal de Negreiros (séc. XVII), por Alícia Duhá Lose, Mariana Fagundes de Oliveira Lacerda, Zenaide de Oliveira Novais Carneiro e Victória Santana da Silva Araújo; e (vi) 30º Jornada do GELNE, em dezembro de 2024 (Universidade Estadual de Feira de Santana, Feira de Santana) – "O trabalho de organização de *corpora* históricos, a pesquisa empírica e a periodização da história sociolinguística do Brasil, por Mariana Fagundes de Oliveira Lacerda, Zenaide de Oliveira Novais Carneiro e Alícia Duhá Lose. Em outubro de 2025, no evento SociolinguísticaS, na Universidade Federal da Paraíba (UFPB), apresentaremos o trabalho, intitulado *Plataforma CE-DOHS: língua, história e tecnologia*, na Mesa sobre *Repositório de dados de e para pesquisa*, sob a coordenação da professora Márcia dos Santos Machado Vieira, da Universidade Federal do Rio de Janeiro (UFRJ).

Pesquisa *Modus Scribendi* – Grupo de Pesquisas Paleográficas, Filológicas e Históricas, da Universidade Federal da Bahia (UFBA), anteriormente mencionado.

Figura 2: Página das Coleções Documentais do CE-DOHS



Fonte: Disponível em: http://www.uefs.br/cedohs/view/colecoes_documentais.html.

A fase 2 (colonial) teve início em 2017 com a execução da pesquisa intitulada *O Livro do Gado do Brejo do Campo Seco (Bahia): edição semidiplomática e descrição de índices grafo-fonéticos*, dissertação de Elaine Brandão Santos, desenvolvida no Programa de Pós-Graduação em Estudos Linguísticos (PPGEL) da UEFS.Trata-se de um documento do final do período colonial, da fazenda do Brejo do Campo Seco (BA).

Segundo Santos Filho, o *Livro do Gado* e o *Livro de Razão* (este editado por Silva [2021], em tese intitulada *O sertão por escrito no Livro de Razão: um microcosmo sócio-histórico e linguístico da Bahia rural oitocentista*), constituem "[v]elhos memoriais de 200 anos, manuseados e ensebados quotidianamente pelas mãos de três gerações, guardados depois pela quarta e pelas seguintes, metamorfosearam-se hoje em preciosas relíquias, preciosas testemunhas de remoto passado"(2012, p. 110).

Esses dois livros de fazenda – o *Livro do Gado*, com 57 folhas, escrito de 1755 a 1832, e o *Livro de Razão*, com 195 folhas, escrito de 1795 a 1838 – tiveram suas edições fac-similar e semidiplomática[12] publicadas em 2022na coleção *O Sertão por Escrito: edição filológica dos livros do arquivo do Sobrado do Brejo (Bahia)*, coordenada por Zenaide de Oliveira Novais Carneiro e Mariana Fagundes de Oliveira Lacerda[13].

[12] O CE-DOHS/NELP/UEFS usa as normas de transcrição do PHPB (Castilho, 2020) e do Centro de Estudos, Pesquisa e Documentação Paleográfica do Memória e Arte (CEPEDOP).

[13] A orelha da coleção foi escrita pela paleógrafa e filóloga Alícia Duhá Lose (UFBA/UEFS).

Figura 3: Capas dos dois volumes da Coleção *O Sertão por Escrito*



Fonte: Disponível em: http://www.uefs.br/cedohs/view/coletaneas.html/.

Neste momento, como resultado de parceria de pesquisa entre o CE-DOHS/NELP/UEFS e o Grupo *Modus Scribendi*/UFBA – a partir de um cruzamento de olhares paleográfico, filológico, sócio-histórico e linguístico –, encontra-se, no prelo, a ***série*** *Documentos para a História Linguística do Brasil Colônia*, sob a coordenação geral de Zenaide de Oliveira Novais Carneiro, Mariana Fagundes de Oliveira Lacerda e Alícia Duhá Lose, já anunciada na página *online* do CE-DOHS[14], com publicação pela UEFS Editora, inicialmente com 16 volumes[15]. São eles:

Quadro 2: Volumes a sair na série Documentos para a História Linguística do Brasil Colônia, pela UEFS Editora

Volume 1	Navarro, Eduardo; Lose, Alícia Duhá; Carneiro, Bruna Trindade Gomes; Lacerda, Mariana Fagundes de Oliveira; Carneiro, Zenaide de Oliveira Novais. **Cartas dos Índios Camarões escritas em Tupi Antigo (1645)**.
Volume 2	Araújo, Victória da Silva Santana; Lacerda, Mariana Fagundes de Oliveira; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Cardoso, Lara da Silva. **Testamento de André Vidal de Negreiros (1678)**.

[14] Disponível em: http://www.uefs.br/cedohs/view/coletaneas.html.

[15] A maior parte dos documentos trabalhados até agora está localizada em Portugal, no Arquivo Histórico Ultramarino (AHU), no Arquivo Nacional da Torre do Tombo, na Biblioteca Central da Universidade de Coimbra; mas também na Holanda, no Arquivo Nacional dos Países Baixos. Foi microfilmada e digitalizada, inicialmente, pelo Projeto Resgate Barão do Rio Branco. Em cada volume, consta a informação precisa de localização da documentação nele editada. Para os trabalhos da Série, todavia, análises físicas têm sido realizadas em contato direto com os documentos, e novas imagens têm sido produzidas, pelo que muito agradecemos aos arquivos mencionados, que têm compreendido a necessidade desse contato físico com os originais e das novas imagens, viabilizando um acesso, em nível de exceção.

Volume 3	Paschoal, Izaías Araújo das Neves; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. **Documentos dos terços de homens pretos e pardos (1650-1793).**
Volume 4	Brito, Rosana Carvalho; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. **Documentos de Comissários baianos da Inquisição portuguesa no Brasil (1700-1791).**
Volume 5	Santos, Elane Santos e; Lacerda, Mariana Fagundes de Oliveira; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais. **Documentos de Comissários baianos do Santo Ofício no Brasil (1698-1809).**
Volume 6	Santos, Emília Carneiro dos; Lacerda, Mariana Fagundes de Oliveira; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Cardoso, Lara da Silva. **Documentos da família Vieira Ravasco (1617-1725).**
Volume 7	Souza, Márcia Silva de; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira; Cardoso, Lara da Silva. **Documentos de mamelucos (1694-1723).**
Volume 8	Lemos, Maiara da Silva; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. **Documentos escritos por portugueses no Brasil (1551 e 1736).**
Volume 9	Carneiro, Bruna Trindade Gomes; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Lacerda, Mariana Fagundes de Oliveira. **Um Códice em Língua Geral (1750-1758).**
Volume 10	Alkimin, Ilma Magalhães; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira; Carneiro, Zenaide de Oliveira Novais. **Sermões de Eusébio de Matos (1629-1692).**

Volume 11	Souza, Ionáia Maria de Jesus; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira; Carneiro, Zenaide de Oliveira Novais. **Cartas, requerimentos, pareceres e afins (1712-1758).**
Volume 12	Jesus, Patrícia dos Santos; Lacerda, Mariana Fagundes de Oliveira; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Cardoso, Lara da Silva. **O Borrador de Antônio Gomes Ferrão Castelo Branco (1700-1800).**
Volume 13	Silva, Alan Souza da; Lose, Alícia Duhá; Carneiro, Zenaide de Oliveira Novais; Lacerda, Mariana Fagundes de Oliveira. **Documentos da Jornada do Rio Grande (1597-1598)**.
Volume 14	Dantas, Andreza Alves; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. **Da documentação das Academias Brasílicas na Bahia Setecentista (1724-1725).**
Volume 15	Carvalho, Gabriela Fortalesa de; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. **Cartas selecionadas do padre jesuíta português, Manuel da Nóbrega (1517-1570), escritas no Brasil entre 1549-1560.**
Volume 16	Soares, Daniela Oliveira de Souza; Carneiro, Zenaide de Oliveira Novais; Lose, Alícia Duhá; Lacerda, Mariana Fagundes de Oliveira. **Correspondências entre José da Silva Lisboa e Balthasar da Silva Lisboa, escritas na Bahia, entre 1780 e 1805**.

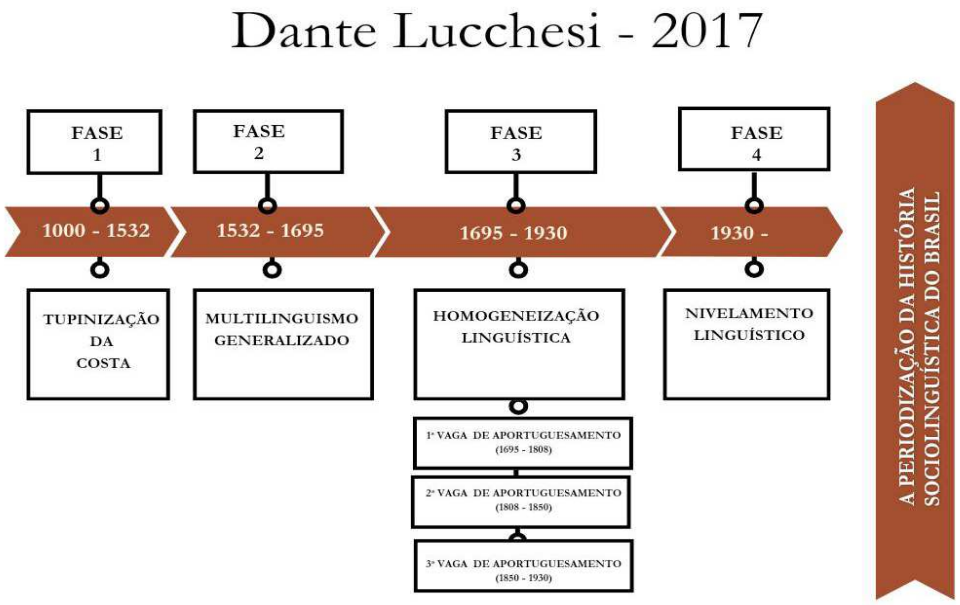
Fonte: Elaborado pelas autoras.

O trabalho de constituição desse *corpus* colonial do CE-DOHS, a nosso ver, representa, junto ao PHPB, um avanço de sua agenda de *corpus* diacrônico, tanto em razão da ampliação do banco de textos – procurando cobrir etapas da periodização da história sociolinguística do Brasil – como em razão do diálogo, com maior grau de maturidade, com a Diplomática, a Paleografia e a Filologia.

O CE-DOHS tem considerado a proposta de periodização da história sociolinguística do Brasil defendida por Lucchesi (2017)[16] e sistematizada na Figura 4.

[16] Conferir também Lucchesi e Callou (2020) sobre os cenários sociolinguísticos do Brasil colônia.

Figura 4: Proposta de periodização da história sociolinguística do Brasil, de Lucchesi (2017)



Fonte: Elaborado pelas autoras.

Conforme o autor,

> [a]o lado de uma vasta pesquisa empírica, possibilitada pelo ingente trabalho de organização de *corpora* históricos, esse movimento também suscitou uma reflexão mais geral sobre a história da língua no Brasil, buscando compreender não apenas a formação da realidade linguística, mas também os fatores que lhe foram determinantes. É nessa perspectiva que se insere a proposta aqui formulada, pois uma periodização deve necessariamente relacionar-se a uma visão de conjunto da história sobre a qual se debruça. Quanto mais explícita for essa relação, mais a periodização alcançará seu objetivo maior: balizar a compreensão de conjunto da história que divide e orientar as pesquisas empíricas no campo. A periodização aqui apresentada busca, assim, contribuir para a compreensão da história da língua no Brasil, particularmente no que concerne aos seus condicionamentos sócio-históricos. Nesse sentido, adota-se aqui como ponto de partida a perfeita síntese de Rosa Virgínia Mattos e Silva (2004: 14), que definiu a história sociolinguística do Brasil como a passagem de "um multilinguismo generalizado a um multilinguismo localizado" (Lucchesi, 2017, p. 348-349).

Com a execução da fase 2 (colonial), o banco CE-DOHS vem cobrindo de forma mais satisfatória a fase 3(da homogeneização linguística, conforme apontada por Lucchesi, 2017), com documentos dos séculos XVII e XVIII, que se somam aos documentos da fase 1 (pós-colonial) de execução do projeto Guarda-Chuva.

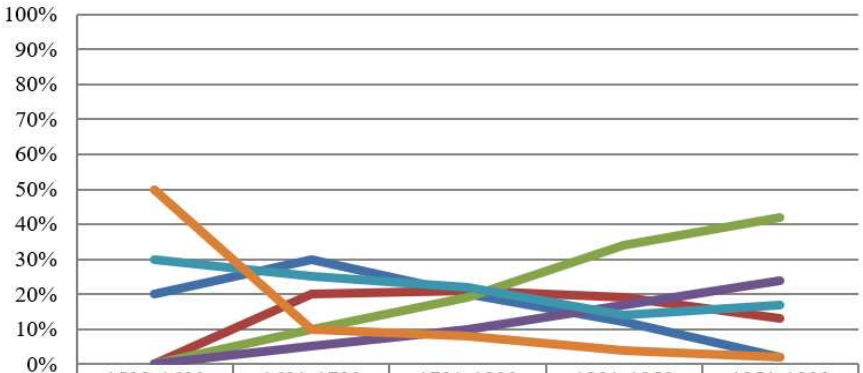
Considerando as proposições de Mattos e Silva – que participou da criação do PHPB e foi coordenadora do PHPB Bahia por muitos anos –, sintetizadas por Lobo

(2015, p. 70), o CE-DOHS/NELP/UEFS desenvolve suas pesquisas sobre a história social linguística do PB:

a) A história linguística do Brasil não se restringe à história da língua portuguesa no Brasil, nem à história do português brasileiro;
b) O português brasileiro emerge em contexto multilíngue: o contato linguístico é, pois, elemento constitutivo da sua formação;
c) Na cena linguística do Brasil colonial, destacam-se três atores principais: o português europeu, as línguas gerais indígenas e o português geral brasileiro;
d) Africanos e afrodescendentes foram os principais difusores da língua portuguesa no Brasil e os principais formatadores do português brasileiro em sua variante social majoritária – o português popular brasileiro;
e) O passado sócio-histórico-linguístico do Brasil deverá ser interpretado para a compreensão do português brasileiro "heterogêneo e variável, plural e polarizado" da atualidade.

Além dessas proposições, o CE-DOHS/NELP/UEFS também considera dados da demografia histórica do Brasil (Mussa, 1991), que revelam os atores principais na cena linguística. O Gráfico 1, adaptado de Mussa (1991), apresenta as porcentagens de raça presentes nas diversas fases de colonização.

Gráfico 1 – Demografia histórica do Brasil. Adaptado de Mussa (1991, p. 163)[17]



	1538-1600	1601-1700	1701-1800	1801-1850	1851-1890
Africanos	20	30	20	12	2
Negros brasileiros	0	20	21	19	13
Mulatos	0	10	19	34	42
Brancos brasileiros	0	5	10	17	24
Europeus	30	25	22	14	17
Índios integrados	50	10	8	4	2

Fonte: Lemos (2023, p. 29).

[17] A respeito da tabela de Mussa, ressaltamos a importância de utilizar, nos dias atuais, diferentemente do autor, "miscigenado" em lugar de "mulato", dado o teor racista presente neste termo, tanto em sua etimologia quanto em sua formação discursiva.

O NELP não temo *corpus* ideal nem trabalha com esse conceito. O que oferecemos é um banco, com lacunaridade[18], para uma aproximação da história linguística dos diferentes grupos envolvidos nessa complexa formação do Brasil, sabendo que, no que diz respeito aos grupos sociais subalternizados, as lacunas são ainda maiores, haja vista que "estes raramente deixaram testemunhos autógrafos", e a escrita de sua história linguística pauta-se, em muitos aspectos, "em uma reconstrução a partir de "indícios", e, necessariamente, do presente em direção ao passado"[19] (Lobo, 2001, p. 109). Como disse Mattos e Silva (2008a), trata-se de um trabalho de arqueologia da linguagem, pesquisa indiciária.

Na descrição do perfil sociocultural dos escreventes, entendendo a escrita a partir de quem a escreveu em um determinado contexto sócio-histórico, o CE-DOHS/NELP/UEFS vai além das categorias tradicionais de perfil, investindo esforços para o que Lopes *et al* (2010, p. 242) denominaram "cruzamento de perspectiva":

> Acreditamos no potencial analítico de um cruzamento de perspectivas, que relacione a trajetória de vida dos sujeitos (que vai mais além das categorias tradicionais de perfil), o contexto de produção dos textos (em que momento foi escrito, o que foi escrito, para quem foi escrito, em que condições e com que finalidades foi escrito) e o mapeamento e descrição das redes de escrita (diálogos estabelecidos e possíveis interlocutores). Isto seguramente nos permitirá localizar a produção escrita de um indivíduo num contexto de produção mais amplo, o que, por sua vez, nos garante a possibilidade de uma conceituação alterna de perfil sociocultural (Lopes *et al*, 2010, p. 242).

Esse cruzamento de olhares e cruzamento de perspectivas estão sempre sob o escopo teórico-metodológico da Paleografia e da Filologia, entendido como condição *sine qua non* para a formação de *corpora* em Linguística Histórica e no âmbito do PHPB.

Fazendo os controles sócio-histórico e paleográfico da documentação, o CE-DOHS descreve o "quem" e o "para quê" (Petrucci, 2003), além de descrever, no âmbito

[18] Segundo Paixão de Souza (2006, p. 25-26), "a documentação possível sobre o passado das línguas chegou até nós através da escrita. Isso traz antes de tudo o problema da lacunaridade do registro escrito: de partida, o estudo das línguas sem tradição de escrita, naturalmente, não é possível por essa via; e tampouco é possível estudar, pela via do registro, todas as etapas da evolução dos idiomas com escrita, em vista das lacunas documentais". A autora continua: "Ora, esses registros representam um fragmento dos acontecimentos. Mais que isso: um fragmento daquilo que um determinado contexto histórico julgou relevante registrar; que um segundo momento histórico julgou importante preservar; e que um terceiro momento histórico considerou pertinente examinar" (Paixão de Souza, 2006, p. 36).

[19] No CE-DOHS, as coleções documentais representativas das normas socialmente estigmatizadas do PB são: Correspondências Amigas; Cartas em Sisal; Cartas Marienses, todas do século XX. Sobre as Cartas em Sisal, editadas por Santiago (2012), foi publicada uma coletânea de estudos morfossintáticos, organizada por Santiago, Lacerda e Carneiro (2023), com financiamento do Programa de Auxílio Interno aos Programas de Pós-Graduação *Stricto Sensu* (AUXPPG/UEFS).

da Paleografia tradicional[20], o "quando" e o "onde". São variáveis que, segundo Mattos e Silva (2004), amenizam, com o recurso também **à historiografia**, problemáticas metodológicas da Sociolinguística Histórica, como os problemas da Representatividade e da Autenticidade (Hernández-Campoy; Shilling, 2012).

É com grande esforço e uma equipe especializada em campos diversos que o PHPBBahia temse dedicado, ao longo de anos, à constituição de *corpus* diacrônico, entre outras agendas do PHPB nacional, sempre buscando aprimorar esse trabalho, dando aos *corpora* o tratamento teórico-metodológico necessário.

> A formação de banco de dados, como o Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS), do Núcleo de Estudos de Língua Portuguesa (NELP) da Universidade Estadual de Feira de Santana (UEFS), exige grande esforço e tempo dedicados à sua metodologia de organização: prospecção documental em fontes confiáveis; caracterização sócio-histórica de acervos; controles paleográfico e filológico; edições em diferentes formatos; processamento, armazenamento e disponibilização de dados. Como se vê, um trabalho de pesquisa que envolve diferentes etapas e uma equipe com formação especializada em campos diversos, como a Paleografia, a Filologia, a História Social da Cultura Escrita, a Linguística Computacional, entre outros (Santiago *et al*, 2021, p. 314).

Essas etapas precedem os estudos linguístico-gramaticais, pois "[é] nos textos remanescentes do passado, afinal, que encontramos os dados diacrônicos para a pesquisa em Linguística Histórica" (Mattos e Silva, 1998, p. 106).

5 O tratamento paleográfico-filológico no âmbito do CE-DOHS fase 2 (colonial): uma ilustração

A Paleografia é a ciência que estuda as escritas antigas, decifrando-as, lendo-as,interpretando-as e situando-as nas coordenadas espaço-temporais. A Paleografia Diplomática, por sua vez, é uma subdisciplina da Paleografia, que se concentra especificamente no estudo de documentos oficiais e administrativos, com o objetivo de analisar e interpretar a escrita e a formatação desses documentos, para entender melhor seu contexto histórico, jurídico e administrativo. Portanto, à Paleografia Diplomática interessam os elementos materiais usados para produção da escrita; os modelos caligráficos usados por diferentes comunidades de escrita e em determinados locais e períodos; as fórmulas textuais empregadas para variados tipos de texto, em contextos específicos; e tudo o mais que possa auxiliar na correta interpretação da escrita, na identificação de sua validade e veracidade,

[20] Conferir livro *Paleografia e suas interfaces*, organizado por Lose e Sacramento (2018).

na sua datação e na identificação das mãos que produzem os textos.

A Filologia, por sua vez, é a ciência que estuda as línguas e seus escritos, analisando e interpretando a evolução, estrutura e significado dos escritos ao longo do tempo. Uma parte importante da Filologia **é a** Ecdótica, ou Crítica Textual, que se concentra na preparação de edições de textos, a partir de critérios coerentes e adequadamente aplicados, fazendo com que um texto seja deslocado no tempo e no espaço, sem perder sua história.

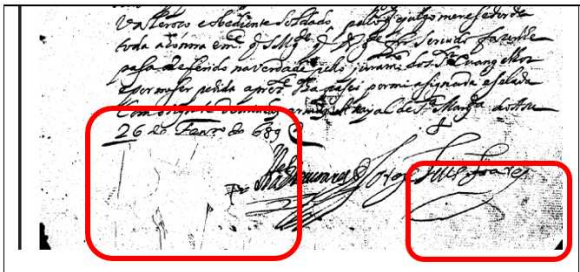
Para o trabalho que vem sendo realizado em parceria com a Linguística Histórica, temos aplicado, para seleção do material a ser transcrito e editado, as metodologias da Paleografia Diplomática, disciplina vital para historiadores da língua que trabalham com textos de épocas passadas, proporcionando ferramentas e métodos para a análise rigorosa e contextualização de registros antigos.

No caso da s**érie** *Documentos para a história linguística do Brasil Colônia*, já aqui referida, tomemos como exemplo o conjunto de textos do Terço dos Homens Pretos e Pardos, instituição encabeçada por Henrique Dias, na capitania de Pernambuco, no âmbito das contendas para expulsão dos holandeses das terras brasileiras. Nas mãos de Izaías Araújo das Neves Paschoal, sob a orientação conjunta de Zenaide de Oliveira Novais Carneiro e Alícia Duhá Lose,

> os documentos manuscritos deixados pelos integrantes dos Terços de Homens Pretos e Pardos passaram por um extenso processo de edição semidiplomática e análise paleográfica – dos escritos e das mãos que levantaram a pena – de modo a garantir a originalidade e a confiabilidade para um estudo linguístico. Sem o apoio da Crítica Textual, da Diplomática e da Paleografia, não seria possível garantir a data de produção dos documentos, tampouco confirmarmos se o assinante fora, realmente, o *scriptor* responsável para que pudéssemos colher os dados linguísticos (Paschoal, 2024, p. 20).

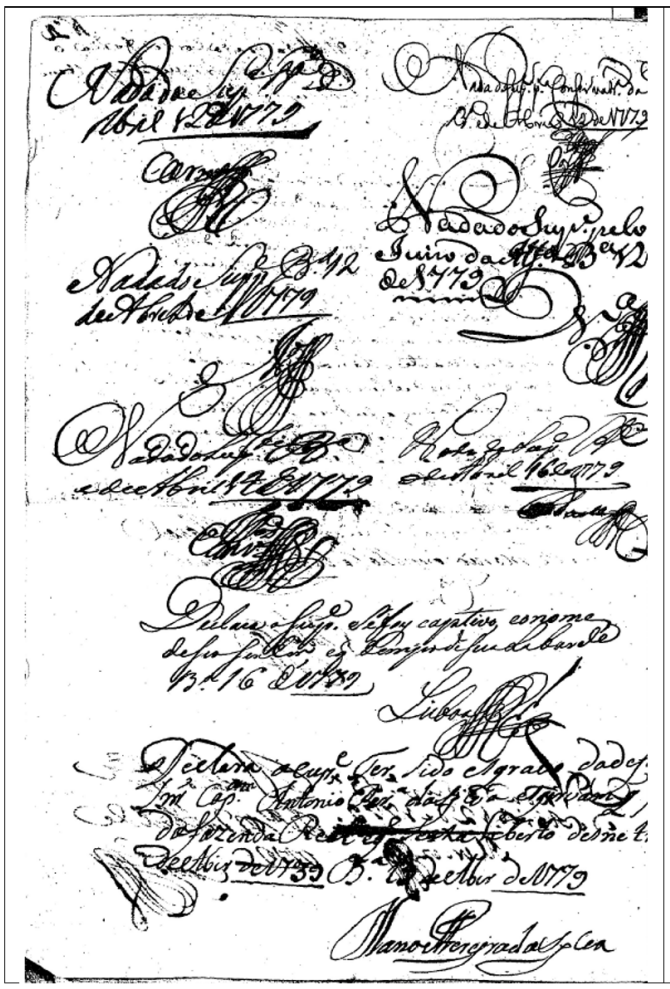
Todos os documentos trabalhados neste conjunto estão custodiados pelo Arquivo Ultramarino, em Lisboa, e foram disponibilizados através do Projeto Resgate Barão do Rio Branco. Todas as análises, por conseguinte, foram feitas a partir das imagens digitais dos microfilmes de cada documento. Não obstante, procedemos **à crítica** paleográfico-diplomática externa: aquela que foca nos aspectos materiais e físicos, analisando características do suporte, dos instrumentos de escrita, dos elementos constitutivos do traçado das letras, dos modelos caligráficos e da *mise-en-page*. Foi possível analisar também as estruturas diplomáticas dos textos e os elementos de validação neles presentes. As Figuras 5, 6 e 7, a seguir, apresentam a sistematização dessas informações.

Figura 5: Visualização das vergaturas do papel, provavelmente de trapoe do selo de chapa colado sobre a data e a assinatura

	Elementos identificados e analisados no documento: - vergaturas do papel; - selo de chapa (papel recortado decorativamente e com duas partes dobradas). As partes desdobradas cobrem parcialmente a data e a primeira assinatura.

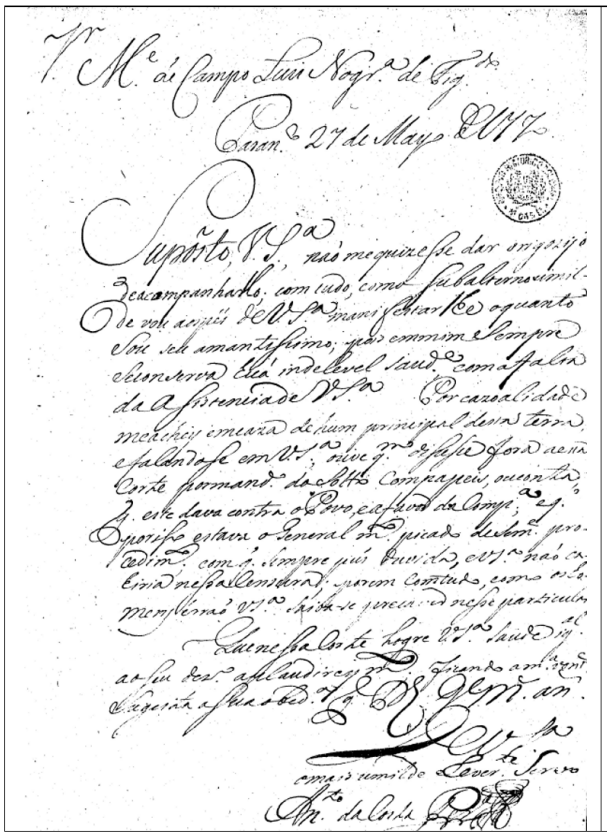
Fonte: ATHPP-JLS-RN-CER-1689.2

Figura 6: Assinaturas, rubricas e guardas, acompanhando os "nada consta" e as declarações

	Elementos identificados e analisados no documento: - escritos a bico de pena; - letras humanísticas cursivas; - variadas mãos com características grafemáticas peculiares e distintivas (como inclinação, peso, *ductus*, altura das hastes, tamanho do módulo e disposição na regradura); - elementos diplomáticos do texto (*validatio*: assinaturas, rubricas e guardas), dispostos, conforme o costume, para o tipo documental e para o período.

Fonte: ATHPP-ASJ-BA-ATE-1774

Figura 7: Elementos comprobatórios de autoria mecânica e intelectual



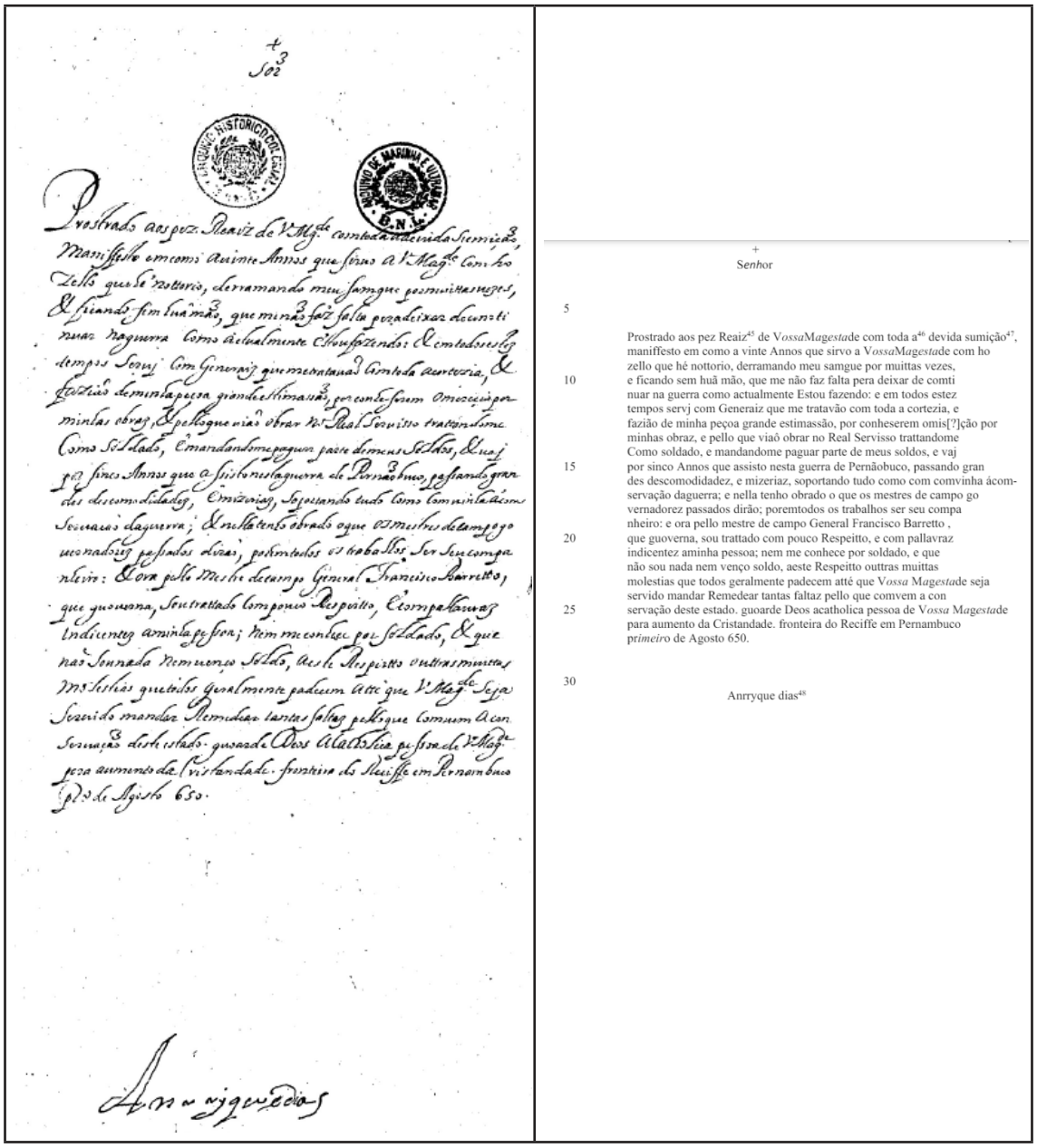
Elementos identificados e analisados no documento:

- escrito a bico de pena;
- letra humanística cursiva, com inclinação dextrógira;
- escrita do texto com a mesma inclinação, peso, *ductus*, altura das hastes, tamanho do módulo e disposição na regradura presentes na escrita da assinatura;
- elementos diplomáticos do texto (protocolo, escatocolo, *validatio*), dispostos comforme o costume para o tipo documental e para o período.

Fonte: ATHPP-ACP-PE-OFI-1770

Depois das análises paleográfico-diplomáticas, foram selecionados, com confiabilidade científica, 28 documentos notariais, manuscritos entre 1650 e 1793, no período do Brasil Colonial, por homens pretos e pardos que integravam os Terços de Homens Pretos e Pardos. Tais documentos foram lidos, transcritos e editados a partir de critérios conservadores: mantendo-se as características linguísticas do texto original, inclusive no procedimento de desenvolvimento de abreviaturas; as letras omitidas foram incluídas em destaque, obedecendo a forma estendida encontrada noutro ponto do documento ou em documentos coetâneos.

Figura 8: Fac-símile e edição semidiplomática de carta datada de agosto de 1650, passada por Henrique Dias para Vossa Majestade, na qual descreve a má administração do Mestre de Campo Francisco Barreto, que o trata com palavras indecentes e não lhe paga o devido soldo.



+
Soz[3]

Prostrado aos pez Reaiz de V.Mg.de com toda a devida sumição, Manifesto em como a vinte Annos que sirvo a V.Mag.de com ho Zello que hé nottorio, derramando meu samgue por muittas vezes, E ficando sem huã mão, que me não faz falta pera deixar de comti nuar na guerra como actualmente Estou fazendo: E em todos estez tempos servj com Generaiz que me tratavão com toda a cortezia, E fazião de minha peçoa grande estimassão, por conheserem omis[?]ção por minhas obraz, E pello que viaõ obrar no Real Servisso trattandome Como soldado, E mandandome paguar parte de meus soldos, E vaj por sinco Annos que assisto nesta guerra de Pernãobuco, passando gran des descomodidadez, E mizeriaz, soportando tudo como com comvinha ácom servação daguerra; E nella tenho obrado o que os mestres de campo go vernadorez passados dirão; poremtodos os trabalhos ser seu compa nheiro: E ora pello mestre de campo General Francisco Barretto, que guoverna, sou trattado com pouco Respeitto, E com pallavraz indicentez aminha pessoa; nem me conhece por soldado, E que nãs sou nada nem venço soldo, aeste Respeitto outtras muittas molestias que todos geralmente padecem atté que V.Mag.de Seja Servido mandar Remedear tantas faltaz pello que comvem a con Servação deste estado. guoarde Deos Acatholica pessoa de V.Mag.de pera aumento da Cristandade. fronteira do Reciffe em Pernambuco p.mr.o de Agosto 650.

Anrryque dias

+
Senhor

Prostrado aos pez Reaiz[45] de V*ossa Magesta*de com toda a[46] devida sumição[47], maniffesto em como a vinte Annos que sirvo a V*ossa Magesta*de com ho zello que hé nottorio, derramando meu samgue por muittas vezes, e ficando sem huã mão, que me não faz falta pera deixar de comti nuar na guerra como actualmente Estou fazendo: e em todos estez tempos servj com Generaiz que me tratavão com toda a cortezia, e fazião de minha peçoa grande estimassão, por conheserem omis[?]ção por minhas obraz, e pello que viaõ obrar no Real Servisso trattandome Como soldado, e mandandome paguar parte de meus soldos, e vaj por sinco Annos que assisto nesta guerra de Pernãobuco, passando gran des descomodidadez, e mizeriaz, soportando tudo como com comvinha ácom servação daguerra; e nella tenho obrado o que os mestres de campo go vernadorez passados dirão; poremtodos os trabalhos ser seu compa nheiro: e ora pello mestre de campo General Francisco Barretto , que guoverna, sou trattado com pouco Respeitto, e com pallavraz indicentez aminha pessoa; nem me conhece por soldado, e que não sou nada nem venço soldo, aeste Respeitto outtras muittas molestias que todos geralmente padecem atté que V*ossa Magesta*de seja servido mandar Remedear tantas faltaz pello que comvem a con servação deste estado. guoarde Deos acatholica pessoa de V*ossa Magesta*de para aumento da Cristandade. fronteira do Reciffe em Pernambuco p*rimeir*o de Agosto 650.

Anrryque dias[48]

Fonte: Elaborada pelas autoras, a partir de Paschoal (2024, p. 244).

A Figura 8 ilustra o trabalho de edição semidiplomática realizado por Paschoal (2024) na sua dissertação de mestrado, defendida e aprovada no PPGEL da UEFS, tendo uma banca examinadora composta pela professora Mariana Fagundes de Oliveira Lacerda e pelo professor e historiador Luiz Geraldo Silva.

6 Possibilidades de estudos linguísticos do *corpus* colonial do CE-DOHS

Uma vez garantida a autenticidade dos materiais – com esse trabalho conjunto da Paleografia, da Filologia e da Linguística Histórica –, que é uma exigência decisiva (Maia,

2012, p. 538), eles podem ser explorados, em diferentes níveis de análise e diferentes quadros teóricos, seja na perspectiva da Linguística Histórica Sócio-Histórica, seja na perspectiva da Linguística Diacrônica (Mattos e Silva, 2008a).

Em relação ao *corpus* colonial do CE-DOHS, Cardoso, Carneiro e Lacerda (2021, p. 350-351) listaram as seguintes possibilidades de estudos linguísticos, entre outras:

> Estudo de painel (Labov, 1994; Rumeu, 2008); estudo de tendências (Labov, 1994); análises contrastivas entre acervos; contraste com *corpora* portugueses do mesmo período; contraste entre as normas usadas pelos colonos e as normas linguísticas propostas pelos portugueses para uso da LP no século XVIII; estudo das marcas de inabilidade dos manuscritos do *corpus*; estudo de fenômenos linguísticos diversos, visando à identificação de marcas específicas do português do Brasil Colonial; ampliação do *corpus* diferencial do PHPB (Cardoso; Carneiro; Lacerda, 2021, p. 350-351).

O CE-DOHS fase 2 (colonial), com seis anos de execução, já apresenta alguns frutos, entre eles: dissertações, teses, artigos, capítulos, a coleção *O Sertão por escrito* e, em breve, a série de documentos e coletâneas de estudos linguístico-gramaticais. Mas nada disso seria possível sem aqueles que vêm discutindo, com espírito científico e ética, a questão do *corpus* no PHPB desde muito tempo. Nesse contexto, destacamos o papel de Ilza Ribeiro desde a criação do NELP/UEFS, com seu projeto inaugural *Contribuições para a Constituição de um Banco de Textos e de um Banco de Dados para o Estudo da História do Português do Brasil, do Século XVII ao XX*.

A respeito das pesquisas de mestrado e doutorado, já foram defendidas cinco dissertações e cinco teses; em andamento, acham-se duas dissertações e duas teses. É importante destacar que todos esses estudos (14, no total)[21] atendem a três agendas do PHPB (constituição de *corpus* diacrônico, estudo sócio-histórico e estudo linguístico); e seguem um mesmo modelo de elaboração: os mesmos tratamentos diplomático, paleográfico, filológico e sócio-histórico. O estudo linguístico realizado segue uma orientação de natureza sobretudo qualitativa, feito em uma abordagem descritivo-interpretativa, comum em Linguística Histórica e que se trata de uma sistematização metodológica que busca "apresentar uma descrição organizada dos factos lingüísticos", tendo em vista "uma gramática descritiva, que opere sobre inventários que se definam como representativos" (Mattos e Silva, 1989, p. 44).

[21] Encontra-se em organização a série *Dissertações e teses para a história sociolinguística do Brasil*, com a publicação de pesquisas que atendem o projeto Guarda-Chuva CE-DOHS: um *corpus* para uma caracterização linguístico-gramatical do português brasileiro - fase colonial e fase pós-colonial. Nessa série, serão apresentados os resultados de exploração dos *corpora* editados, tanto a exploração paleográfico-filológica como as explorações sócio-histórica e linguístico-gramatical.

Entre os dez trabalhos já defendidos, seis abordam fenômenos gramaticais: as dissertações de Santos (2022), Araújo (2023) e Paschoal (2024) e as teses de Brito (2024), Santos (2024) e Lemos (2025)[22]; as dissertações de Santos (2019) e Souza (2022) e as teses de Silva (2021) e Carneiro (2025) abordam índices grafo-fonéticos[23].

Sobre os seis estudos descritivo-gramaticais já defendidos no PPGEL da UEFS e como parte do CE-DOHS fase 2 (colonial), foram selecionados os seguintes fenômenos: a concordância verbal, a sintaxe das orações relativas, o complemento verbal *lhe*, investigados nas dissertações de Santos (2022), Araújo (2023) e Paschoal (2024), respectivamente; e os sintagmas possessivos, os dativos e os clíticos, investigados nas teses de Brito (2024), Santos (2024) e Lemos (2025), respectivamente. Durante o desenvolvimento de suas pesquisas, os seis autores enfrentaram o problema da invariação, em maior ou menor grau, apontados por Hernández-Campoy e Schilling (2012) como um dos principais problemas teórico-metodológicos em estudos linguísticos sócio-históricos. Esse problema acontece porque, em geral, textos escritos fornecem pouca possibilidade de variação linguística, sobretudo quando o *corpus* possível são documentos notariais, como é o caso dos documentos explorados nas três dissertações em questão.

A pesquisa de Araújo (2023), por exemplo – orientada por Mariana Fagundes de Oliveira Lacerda e Alícia Duhá Lose –, cujo foco é a sintaxe das relativas no Testamento de André Vidal de Negreiros, do século XVII,(tema que se destaca, sendo considerado decisivo na emergência do português brasileiro (Tarallo, 1983, 1993), em um processo de "garimpar", chegou aos seguintes resultados, localizando três ocorrências de relativas não-padrão. São eles:

> Embora tenham sido levantadas, no total, 153 sentenças relativas, a maior parte desses números se trata do emprego de estratégias de relativização padrão. Grande parcela dessa constatação deve-se à quantidade de registros de cláusulas cujas funções relativizadas são as de Sujeito e de Objeto Direto, que, juntas, alcançam um total de 110 ocorrências padrão. Além disso, nas posições mais baixas, também são predominantes tais estratégias. (...) São apenas três as sentenças relativas de estratégia não-

[22] Resultados desses estudos gramaticais, entre outros, fazem parte do livro, intitulado *Oito estudos gramaticais sobre o período colonial brasileiro* (no prelo, Pontes Editores), sob a organização de Mariana Fagundes de Oliveira Lacerda, Zenaide de Oliveira Novais Carneiro e Elane Santos e Santos, com recursos SEPESQ/UEFS, e do livro, intitulado *Para a história do português no Brasil colonial: revisitando problemas e ampliando análises* (no prelo, UEFS Editora), sob a organização de Rosana de Carvalho Brito, Elane Santos e Santos, Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro, com recursos SEPESQ/UEFS.

[23] Resultados dessas pesquisas de natureza grafo-fonética fazem parte do livro, intitulado *Corpora histórico-diacrônicos do português brasileiro: da descrição grafemática aos indícios da oralidade* (no prelo, Pontes Editores), sob a organização de Huda da Silva Santiago, Adilson Silva de Jesus, Mariana Fagundes de Oliveira Lacerda e Zenaide de Oliveira Novais Carneiro, com recursos SEPESQ/UEFS.

> padrão identificadas no *corpus*. Duas delas cuja função relativizada é a de Objeto Direto e apenas uma cuja função relativizada é a de Adjunto (Araújo, 2023, p. 105-106).

Araújo (2023) evidencia que "vestígios de tinta" de relativas não-padrão, como as relativas cortadoras e as relativas resumptivas, podem ser encontrados em textos escritos durante o Brasil Colonial. É mesmo um trabalho de "formiga obreira", aqui e ali atrás de dados, em busca de indícios que ajudem a contar a história do PB. Os *corpora* de Santos (2022), Araújo (2023) e Paschoal (2024) são datados dos séculos XVII e XVIII, quando ainda não era possível atestar a existência do português do Brasil; mas investigar como se configurava o português no Brasil pode contribuir para uma reconstrução, por aproximação, da história social linguística do PB. É nisto que o CE-DOHS tem investido esforços desde 2017: enfrentar a raridade documental, quando se trata do período colonial no Brasil, e procurando fazer o melhor uso dos dados disponíveis.

7 Considerações finais: o trabalho com *corpus* não é cômodo, mas é necessário

Desde o início do projeto, em 1997, Mattos e Silva adverte a todos sobre a "enormidade do empreendimento" (Castilho, 2004, p. 3), referindo-se ao PHPB. Além disso, sobre a agenda de *corpus* diacrônico do projeto, a autora complementa:

> O trabalho com *corpus*não é "commodo" – pelo contrário, é por vezes penoso, sempre curioso, às vezes divertido (...) Sem dúvida é um trabalho necessário, como base para a reconstrução do passado lingüístico do português que aqui se formou, o português brasileiro (Mattos e Silva, 2004, p. 120).

A fala de Mattos e Silva (2004) destaca a importância de fazer um trabalho que não é cômodo, mas que é necessário, como necessário é o diálogo entre a Linguística Histórica, a Diplomática, a Paleografia e a Filologia. Sem esse diálogo não é possível o trabalho com *corpus* para estudos linguísticos e discussões teóricas com solidez.

Ao longo de 28 anos, os pesquisadores do PHPB, nas diferentes fases de trabalho,"-como quixotes ou como loucos, ou apenas como brasileiros interessados em compreender um aspecto fundamental da sua história pregressa" (Mattos e Silva, 2004, p. 67), têm buscado executar as agendas de trabalho desse enorme empreendimento, trazendo, com cada equipe regional, como a equipe da Bahia, elementos *Para uma História do Português Brasileiro*. Elementos esses que sempre podem, numa perspectiva interdisciplinar, como defendemos aqui, ser revisitados e refinados.

Como "quixotes ou como loucos" (Mattos e Silva, 2004, p. 7) ou "como formigas obreiras" (Mattos e Silva, 2008b, p. 40), os pesquisadores do PHPB continuam empenhando esforços no seu grande empreendimento... E o NELP/UEFS tem sido um braço forte, especialmente, mas não apenas, na agenda de constituição de *corpus* diacrônico; como afirma Castilho, no prefácio da série colonial (no prelo), "O NELP, com quase 30 anos de trabalho intenso, tem já uma enorme folha de serviços em favor da Linguística Histórica no Brasil."

Referências

ARAÚJO, V. S. S. *O Testamento de Vidal de Negreiros (século XVII)*: edição semidiplomática e elementos para uma descrição da sintaxe das relativas. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana, UEFS, 2023.

BACELAR DO NASCIMENTO, M. F. *O lugar do corpus na investigação linguística*. Disponível em: http://www.clul.ul.pt/equipa/berlim-2000-nascimento.pdf.. Acesso em: 20 abril 2004.

BRITO, R. C. *Pelas mãos de Comissários do Santo Ofício na Bahia setecentista:* edição semidiplomática e estudo do artigo definido diante de pronome possessivo. Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2024.

CARDOSO, L. S.; LACERDA, M. F. O.; CARNEIRO, Z. O. N. Relações entre a constituição de corpora diacrônicos do português brasileiro e os princípios da Ciência Aberta. *ABRALIN em Cena 17*: dados linguísticos, 2023.

CARDOSO, L. S.; CARNEIRO, Z. O. N.; LACERDA, M. F. O. Para um estudo da formação do português brasileiro: descrição, representatividade e potencialidades do corpus colonial do CE-DOHS. *In: LaborHistórico*, v.7, 2021. p.330-355.

CARNEIRO, Z. O. N.; LACERDA, M. F. O. (Org). *CE-DOHS - Corpus Eletrônico de Documentos Históricos do Sertão - Início*. Disponível em: http://www.uefs.br/cedohs/view/home.html. Acesso em: 14 março 2024.

CARNEIRO, Z. O. N.; LACERDA, M. F. O. (Coordenação Geral). *Coleção O Sertão por Escrito*: edição filológica dos livros do arquivo do Sobrado do Brejo (Bahia). Feira de Santana: Editora UEFS, 2022, 2 volumes. p. 498.

CARNEIRO, Z. O. N.; SILVA, A. J.; SOUZA, E. H. P. M. *Livro de Razão (XVIII-XIX):* edições fac-similar e semidiplomática. Feira de Santana: Editora UEFS, 2022, v.1.

CARNEIRO, Z. O. N.; LACERDA, M. F. O.; LOSE, A. D. (Coordenação Geral). *Série Documentos para a HistóriaLinguística do Brasil Colônia*. Feira de Santana: UEFS Editora (No prelo).

CARNEIRO, B. T. G. U*m códice em língua geral:* edição, estudo paleográfico e sócio-história da Amazônia (1750-1758). Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2025.

CASTILHO, A. T. Historiando a língua portuguesa no novo mundo. *In: XX Conferência final no Congresso Internacional da Associação Brasileira de Linguística e Filologia da América Latina (ALFAL), Projeto 3 - História do Português Brasileiro, da Europa à América*. Faculdade de Humanidades e Arte. Universidade de Concepción, Chile, 2024.

CASTILHO, A. Projeto *Para a História do Português Brasileiro (PHPBB), Normas de Transcrição de Documentos Manuscritos e Impressos - Edição Semidiplomática*. Disponível em: https://sites.google.com/site/corporaphpb/home/normas-de-edicao-do-phpb-2a-versao. Acesso em: 1 março 2023.

DUARTE, M. E. L.&CALLOU, D. M. I. (Org.). *Para a História do Português Brasileiro,* v. IV: Notícias de corpora e outros estudos. Rio de Janeiro: Universidade Federal do Rio de Janeiro / Faperj, 2002.

HERNÁNDEZ-CAMPOY, Juan M.; SCHILLING, Natalie. The Application of the Quantitative Paradigm to Historical Sociolinguistics: problems with the generalizability principle. *In:* HERNÁNDEZ-CAMPOY, Juan Manuel; CONDE-SILVESTRE, Juan Camilo. *The Handbook of Historical Sociolinguistics*. New Jersey: Blackwell Publishing, 2012. p. 61-79.

LACERDA, M. F. O. O Núcleo de Estudos de língua Portuguesa da Universidade Estadual de Feira de Santana: centro de Linguística Histórica no Semiárido baiano. *In: Revista Confluência*. 2024. p. 147-199.

LACERDA, M. F. O.; CARNEIRO, Z. O. N.; LOSE, A. D. O trabalho de organização de *corpora* históricos, a pesquisa empírica e a periodização da história sociolinguística do Brasil. *In: 30º Jornada do GELNE*. Universidade Estadual de Feira de Santana. Feira de Santana, 2024.

LACERDA, M. F. O.; CARNEIRO, Z. O. N.; LOSE, A. D. *O CE-DOHS na rede de pesquisa PHPB: recuando ao período colonial. In: XI Seminário de Estudos Filológicos: Filologia e Ensino: perspectivas críticas e editoriais*. Universidade Estadual de Feira de Santana, Feira de Santana, 2024.

LACERDA, M. F. O.; CARNEIRO, Z. O. N. (Org). *Núcleo de Estudos de Língua Portuguesa*.

Disponível em: https://nelp.uefs.br/. Acesso em: 14 mar 2024.

LACERDA, M. F. O.; SANTOS, E. B. *Livro do Gado (XVIII-XIX):* edições fac-similar e semidiplomática. Feira de Santana: Editora UEFS, 2022, v.1. p.118.

LEMOS, D. M. *ELIHS, um projeto para além do litoral:* da constituição de *corpus* oral de comunidades afro-brasileiras **à análise estatística da variação na concordância nominal de número na variedade da comunidade afro-brasileira de Volta do** Angico(BA). 2023. Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2023.

LEMOS, M. da S. *A língua vinda d'álém-mar:* colocação de clíticos na escrita de portugueses no Brasil colonial. Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2025.

LOBO, T. C. F. Rosa Virgínia Mattos e Silva e a história social linguística do Brasil. *Estudos de Linguistica galega*, v. 7, p. 69-82, 2015.

LOBO, T. C. F. L. Arquivos, acervos e a reconstrução histórica do português brasileiro. *In:* OLIVEIRA, K. CUNHA E SOUZA, H.; SOLEDADE, J. *Do português arcaico ao português brasileiro*: outras histórias. Salvador, Edufba, 2009. p. 305-327.

LOPES, C. R. S.; MARCOTULIO, L. L.; RUMEU, M. C. B.; LIMA, A. X. Reflexões metodológicas para a análise sociocultural de redatores em *corpora* históricos. *In: Gragoatá*, Niterói, n. 29, 2010.

LOSE, A. D. *Grupo de Pesquisa Modus Scribendi- Grupo de Pesquisas Paleográficas, Filológicas e Históricas*. Disponível em: https://letras.ufba.br/modus-scribendi-grupo-de-pesquisas-paleograficas-filologicas-e-historicas. Acesso em: 1 mar 2024.

LOSE, A. D. *Centro de Estudos, Pesquisa e Documentação Paleográfica do Memória e Arte (CEPEDOP)*. Disponível em: https://www.memoriaarte.com.br/. Acesso em: 1 mar2024.

LOSE, A. D.; LACERDA, M. F. O.; CARNEIRO, Z. O. N.; ARAÚJO, V. S. S. Fontes linguísticas para estudos do português no Brasil colonial: o exemplo do Testamento de André Vidal de Negreiros (séc. XVII). *In: IX Simpósio Mundial de Língua Portuguesa*. Universidade da Madeira, Ilha da Madeira. Portugal, 2024.

LOSE, A. D.; CARNEIRO, Z. O. N.; LACERDA, M. F. O.; CARDOSO, L. S. *O português no Brasil colonial: reflexões sobre as edições de manuscritos históricos brasileiros. XX Conferência final no Congresso Internacional da Associação Brasileira de Linguística e Filologia da América Latina (ALFAL), Projeto 3 - História do Português Brasileiro, da Europa à América*. Faculdade de Humanidades e Arte. Universidade de Concepción, Chile, 2024.

LOSE, A. D.; LACERDA, M. F. O.; CARNEIRO, Z. O. N. *A crítica paleográfica como base da seleção dos corpora. In: XI Seminário de Estudos Filológicos: Filologia e Ensino:* perspectivas críticas e editoriais. Universidade Estadual de Feira de Santana, Feira de Santana, 2024.

LOSE, A. D.; LACERDA, M. F.; CARNEIRO, Z. O. N. Reflexões sobre as edições de manuscritos históricos brasileiros: a relação entre o olhar paleográfico e o olhar linguístico. *In: Simpósio da Contemporaneidade à Antiguidade*: temas em debate. Centro de Estudos Clássicos e Humanísticos da Universidade de Coimbra, Coimbra, 2024.

LOSE, A. D.; LACERDA, M. F.; CARNEIRO, Z. O. N. Edições de manuscritos históricos brasileiros: a série Documentos para a história linguística do Brasil colônia. *In: VI Congresso Internacional de Linguística Histórica*. Faculdade de Letras da Universidade de Coimbra, Coimbra, 2023.

LOSE, A. D.; SOUZA, SACRAMENTO, A. de. Para uma filologia na pesquisa em linguística histórica. Letras. Revista do Programa de Pós-Graduação em Letras da Universidade de Santa Maria, Santa Maria, n. 60, 2022. p. 11-31. Disponível em: https://doi.org/10.5902/2176148542058. Acesso em: 3 abril 2023.

LOSE, A. D. L.; SACRAMENTO, A. *Paleografia e suas interfaces*. Salvador: Memória e Arte, 2018.

LOSE, A. D. Edições de documentos históricos: a quem interessam? a quem se destinam? *Revista da ABRALIN*, n. 16, v. 2, São Paulo, 2017. p. 71-86.

LUCCHESI, D.; CALLOU, D. Os cenários sociolinguísticos do Brasil colonial. *In:* Castilho, A. (Coordenador Geral). História do português brasileiro. v. 9. CALLOU, D.; LOBO, T. *História social do português brasileiro*: da história social à história linguística. São Paulo: Contexto, 2020. p. 156-181.

LUCCHESI, D. A periodização da história sociolingüística do Brasil. *In: Revista D.E.L.T.A*. 33.2, 2017. Disponível em: https://www.scielo.br/j/delta/a/NGxLPBSqNXYNGhFtwqrrwgh/?format=pdf&lang=pt. Acesso em: 14 fev 2024.

MAIA, C. Linguística Histórica e Filologia. *In:* LOBO, T.; CARNEIRO, Z; SOLEDADE, J. et al. (Org.). *Rosae linguística histórica, história das línguas e outras histórias*. Salvador: EDUFBA, 2012. p. 533-541.

MARQUILHAS, R.; HENDRICKX, I. Avanços nas humanidades digitais. *In:* MARTINS, A. M.; CARRILHO, E. *Manual de Linguística Portuguesa*. MRL Series. De Gruyter, 2016. p. 1-26.

MATTOS E SILVA, R. V. *Estruturas trecentistas*: elementos para uma gramática do português arcaico. Lisboa: Imprensa Nacional: Casa da Moeda, 1989.

MATTOS E SILVA, R. V. *Caminhos da linguística histórica*: ouvir o inaudível. São Paulo, SP: Parábola Ed., 2008a.

MATTOS E SILVA, R. V. Teorias da mudança lingüística e a sua relação com a(s) história(s) da(s) língua(s). *In: Revista de Estudos Lingüísticos da Universidade do Porto*. v. 3. 2008b. Disponível em: https://ler.letras.up.pt/uploads/ficheiros/6874.pdf. Acesso em: 1 abril 2024.

MATTOS E SILVA, R. V. *Ensaios para uma sócio-história do português brasileiro*. São Paulo: Parábola Editorial, 2004.

MATTOS E SILVA, R. V. Sobre desencontros e reencontros: Filologia e Lingüística no Brasil no século XX. *Estudos Lingüísticos e Literários*, Salvador, n. 21-22, 1998. p. 97-108.

OLIVEIRA, K. (2006). *Negros e escrita no Brasil do século XIX*: sócio-história, edição filológica de documentos e estudo linguístico. Tese de Doutorado. Programa de Pós-Graduação em Letras e Linguística. UFBA, 2006.

PASCHOAL, I. A. N. *Documentos dos Terços de Homens Pretos e Pardos (1650-1793)*: edição semidiplomática e estudo do complemento verbal lhe. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana, UEFS, 2024.

PETRUCCI, A. *La ciencia de la escritura*: primeralección de paleografía.Buenos Aires: Fondo de Cultura Económica de Argentina, 2003.

RIBEIRO, I. M. *Constituição de um Banco de Textos e de um Banco de Dados para o Estudo da História do Português do Brasil, do Século XVII ao XX*. 1997. Disponível em: https://nelp.uefs.br/o-nelp/historico/. Acesso em: 13 abril 2024.

SANTIAGO, H. S.; LACERDA, M. F. O.; BRITO, R. C.; CARNEIRO, Z. O. N. CE-DOHS: um banco de dados sociolinguísticos para a história do português brasileiro. *In: LaborHistórico*, vol 7, número especial, Rio de Janeiro. 2021. p. 311-329.

SANTOS, E. B. *O Livro do Gado do Brejo do Campo Seco (Bahia):* edição semidiplomática e descrição de índices grafo-fonéticos. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana, UEFS, 2019.

SANTOS, E. C. *Documentos da Família Vieira Ravasco (XVII-XVIII):* edição semidiplomática e elementos para uma descrição da concordância verbal. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana, UEFS, 2022.

SANTOS, E. S. *Entre Linhas e Testemunhos:* edição semidiplomática de manuscritos de comissários baianos do Santo Ofício e estudo descritivo de argumentos dativos. Tese (Doutorado em Estudos Linguísticos). Feira de Santana, UEFS. 2024.

SANTOS FILHO, L. *Uma comunidade rural do Brasil antigo*: aspectos da vida patriarcal no sertão da Bahia nos séculos XVIII e XIX. São Paulo: Editora Nacional, 1956.

SILVA, A. J. *O sertão por escrito no Livro de Razão*: um microcosmo sócio-histórico e linguístico da Bahia rural oitocentista. Tese (Doutorado em Língua e Cultura). Salvador, UFBA, 2021.

SOUZA, M. S. de. *Documentos de Mamelucos no Brasil Colonial*: edição e descrição dos aspectos gráficos. Dissertação (Mestrado em Estudos Linguísticos). Feira de Santana (UEFS). 2022.

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. *In:* ROBERTS,I.; KATO,M. (Orgs.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora da UNICAMP,2018[1993].

TARALLO, Fernando. *Relativization strategies in Brazilian Portuguese*. Tese (Doutorado em Linguística) – University of Pennsylvania, Philadelphia, 1983.

http://doi.org/10.5007/1984-8420.2025.e101224

ANÁLISE DO *STATUS* FONOLÓGICO DAS CONSOANTES RÓTICAS, LATERAIS E NASAIS DUPLAS DO PORTUGUÊS ARCAICO

ANALYSIS OF THE PHONOLOGICAL STATUS OF DOUBLE RHOTIC, LATERAL AND NASAL CONSONANTS IN ANCIENT PORTUGUESE

Débora Aparecida dos Reis Justo Barreto | Lattes | debi_barreto@hotmail.com
Universidade Estadual Paulista

Gladis Massini-Cagliari | Lattes | gladis.massini-cagliari@unesp.br
Universidade Estadual Paulista

Resumo: A finalidade deste artigo é a de estudar os fenômenos fonológicos do português dos trovadores, analisando especificamente as consoantes róticas, laterais e nasais duplas, retratadas na escrita como <rr>, <ll/lh> e <nn/nh>, existentes em 250 cantigas medievais – 100 da vertente religiosa e 150 da vertente profana. Nosso objetivo é o de averiguar se, no nível fonológico, tais consoantes duplas podiam ser interpretadas como geminadas em contexto intervocálico naquela época da história da língua portuguesa. A análise dos dados coletados foi desenvolvida por meio das teorias fonológicas não-lineares. A metodologia se embasa na análise das variações gráficas, muito frequentes nos documentos remanescentes daquele período, e no estudo do comportamento fonológico das referidas consoantes dentro da sílaba e da palavra. Os dados coletados retrataram que, em posição intervocálica, róticas, laterais e nasais duplas apresentavam o mesmo *status* no período arcaico, ou seja, podiam ser interpretadas como geminadas do ponto de vista fonológico da língua, pois compreendem elementos que valem por dois. Tal interpretação pôde ser validada em decorrência de não termos encontrado ditongos antes de tais segmentos duplos nem palavras proparoxítonas com <rr>, <ll/lh> e <nn/nh> no ataque da última ou da penúltima sílaba.

Palavras-chave: Consoantes duplas; Rótica; Lateral; Nasal; Português arcaico.

Abstract: This article aims to study phonological phenomena from medieval troubadours' Portuguese, specifically analysing double rhotic, lateral and nasal consonants, graphically represented by <rr>, <ll/lh> and <nn/nh>. This study considers 250 medieval cantigas – 100 religious and 100 secular ones. Our goal is to verify whether these double

consonants could be interpreted as geminate in phonological level, in intervocalic context, at that historical period. The analysis of the selected data is based on non-linear phonological theories. The methodology is based on the analysis of graphic variations and in the study of the phonological behaviour of these kind of consonants inside the syllable and word-internally. Data show that, between vowels, double rhothic, lateral and nasal consonants present the same status in Ancient Portuguese, that is, they can be interpreted as geminate in phonological level, because they correspond to elements that "count" as two. This interpretation could be validated since we could not find diphthongs before these double segments and proparoxytone words with <rr>, <ll/lh> and <nn/nh> in the onset of the last or penultimate syllable.

Keywords: Double consonants; Rhotics; Laterals; Nasals; Ancient Portuguese.

Introdução e objetivo

Este artigo foi feito em homenagem a Ataliba Teixeira de Castilho, o maior congregador entre os linguistas brasileiros[1], tendo coordenado grandes projetos coletivos, como o Gramática do Português Falado (Castilho, 2003) e o Para a História do Português Brasileiro (PHPB), ao qual esteve ligada uma das coautoras deste trabalho (Massini-Cagliari, 2019). Neste último projeto, uma das características marcantes da pesquisa desenvolvida no contexto do grupo encarregado da investigação da mudança fônica (Hora, Battisti e Monaretto, 2019) era o equilíbrio entre, por um lado, a confiança advinda das edições de documentos históricos realizadas a partir de um acurado trabalho filológico e, por outro, a busca de novas metodologias, embasadas em teorias fonológicas mais recentes, que permitissem avançar além dos conhecimentos advindos dos estudos de gramática histórica de tradição filológica. Este era o espírito passado pelo Professor Ataliba, nas suas aulas de sintaxe na Universidade Estadual de Campinas (Unicamp), em que uma das autoras iniciou seu mergulho na pesquisa histórica, tendo enveredado por outros caminhos, que a levaram, e também à outra coautora, à investigação das origens da Fonologia do Português, no período medieval de sua formação.

Trazendo um novo olhar metodológico (Massini-Cagliari, 2013) aos estudos medievais, mas ao mesmo tempo partindo do imenso respeito ao cuidado filológico, este

[1] Na entrevista que concedeu a Carlos Fioravanti (2017), para a *Revista Fapesp*, sobre a necessidade de realização de grandes projetos sobre a língua portuguesa, Castilho afirma: "O estudo de uma língua tem de ser feito em grupo. Esse é o corolário do Nurc, da Gramática do Português Falado e do Projeto para a História do Português Brasileiro, que também estimula a convivência de gente com visões diferentes, abrigando sociolinguistas, gerativistas, funcionalistas e cognitivistas. [...] Eu respeito o pensamento diferente. E ao respeitar você junta as pessoas. Ninguém quer ficar levando lambada dos outros. Ciência não existe para isso, mas para unir as pessoas na descoberta do conhecimento.".

artigo visa sumarizar a análise empreendida pelas pesquisadoras acerca do processo de geminação das consoantes duplas do Português Arcaico (PA). Ao longo dos anos, foram feitos, separadamente, estudos acerca de três consoantes duplicadas[2] daquele período da história: *rr*, *ll/lh* e *nn/nh*[3]. Esses segmentos apresentam similaridades relevantes quando preenchem o ambiente intervocálico da palavra no português arcaico[4].

Assim, este estudo propõe uma comparação com relação às três consoantes duplas acima expressas acerca do *status* fonológico desses segmentos da língua do medievo quando estão em posição intervocálica. Convém ressaltar que este estudo faz uso de dados da escrita para analisar aspectos fonológicos do período ora focalizado, tendo em vista que a etapa arcaica do português não apresentava, obviamente, tecnologia voltada à gravação e à transmissão de produções orais. Assim, o único modo de conseguir extrair pistas prosódicas a respeito daquela língua é por meio da análise da estrutura métrica de poemas, obrigatoriamente baseada nas características rítmicas do falar que a ela oferece suporte[5].

Por isso, para a concretização deste trabalho, foram selecionadas 250 cantigas medievais galego-portuguesas: as 100 primeiras *Cantigas de Santa Maria* (CSM), da vertente religiosa, e 150 poemas profanos, 50 de cada gênero canônico[6]. O uso de obras poéticas, como já dito, é de vital importância para a elaboração deste estudo, pois a escrita do PA era de base alfabética, isto é, os documentos remanescentes daquela época não expressavam uma notação especial para os fenômenos prosódicos da língua. Logo, um estudo de natureza fonética/fonológica por meio de textos em prosa é praticamente impossível.

---

[2] As análises não incluíram <ss> porque não há dúvidas de que o dobro da letra é um fenômeno apenas gráfico, não havendo possibilidade de esse grafema representar uma consoante fonologicamente dupla.

[3] As escritas nn e ll aparecem nos cancioneiros contemporâneos aos trovadores; já as escritas nh e lh aparecem nos cancioneiros copiados em períodos posteriores. Segundo Williams (1975 [1938], p. 36), no início da história escrita do português, a nasal palatal [ɲ] apresentava diferentes representações na escrita, como <n>, <ni> e <nn>, por **não haver correlato gráfico no alfabeto latino para a nasal palatal, pois esta é um som inovador do português**. O mesmo pode ser apurado em relação à lateral palatal [ʎ], que também apresentava distintas representações, como <l>, <li> e <ll>. A grafia <nn> veio, através da Galícia, da Espanha. Conforme o estudioso, a primeira ocorrência de *nh* data de meados dos anos 1263/1273. A teoria aceita pelos pesquisadores para a origem das formas gráficas <nh> e <lh> é a de que tais grafias foram emprestadas do provençal. Entretanto, não há uma unanimidade com relação às razões que motivaram tal empréstimo.

[4] Usaremos tal expressão, comumente referida em outros estudos como "galego-português", para nos referir ao ancestral medieval do Português falado no Brasil (PB), em homenagem a Mattos e Silva (1989).

[5] Vários estudos realizados na área de Fonologia não-linear, como Prince (1989), Halle (1989), Hayes (1989), entre outros, provaram que um estudo voltado à Fonologia a partir de obras elaboradas no passado é realizável e possível, sobretudo quando se desenvolve uma descrição em um nível *mais abstrato*, ou seja, mais fonológico e não fonético.

[6] *Cantigas de amor*, *de amigo* e *de escárnio e maldizer*. Os critérios usados para selecionar os poemas profanos se fundamentam em Massini-Cagliari (2015) e são três: 1) representatividade (escolhendo textos de autores de épocas diferentes); 2) local; e 3) classe social.

Composições poéticas, por outro lado, caracterizam-se por conter elementos segmentais (lugar dos acentos nos versos, número de sílabas poéticas etc.), informações por meio das quais é possível depreender padrões e reincidências acentuais e rítmicas da língua medieval (Massini-Cagliari, 2015).

O *status* fonológico das consoantes duplas *rr*, *ll/lh* e *nn/nh* sempre foi uma questão muito controversa na literatura, não somente hoje em dia, mas também em períodos anteriores ao atual. Sobre as róticas, o que se encontra em textos da área é o registro de duas interpretações: a língua portuguesa tem dois fonemas róticos, um fraco e um forte; ou o português conta com um fonema rótico que, para uns, é o r-forte e, para outros, o r-fraco. No entanto, para o PA, como exibiremos neste trabalho, o que se tem é um único fonema rótico com duas variantes, uma simples, grafada como <r> e <rr>, que pode aparecer em diferentes posições dentro da palavra, e uma geminada no contexto intervocálico, representada na escrita como <rr> (Barreto, 2019).

Já acerca das palatais *ll/lh* e *nn/nh*, inovações do português, uma vez que esses elementos não existiam no latim (Câmara Jr., 1975), tem-se que tais consoantes, quando intervocálicas, se comportam da mesma maneira que a rótica dupla, isto é, podem ser interpretadas, na perspectiva fonológica, como segmentos geminados no PA, pois valem por dois na estrutura silábica interna daquela época da história da língua (Barreto, 2023). A complexidade de tais consoantes também foi constatada por Wetzels (2000) no que diz respeito ao Português Brasileiro (PB).

Este artigo, portanto, pretende sumarizar e entender como se constitui o quadro de consoantes duplas geminadas do PA. As consoantes róticas, laterais e nasais dobradas do PA[7] exibem alguns traços comuns quando ocupam a posição intervocálica do termo e têm o mesmo comportamento fonológico desse ambiente específico da palavra. Por meio dos dados coletados diretamente dos cancioneiros arcaicos, este estudo visa demonstrar o caráter complexo desses segmentos e expor as conclusões às quais chegamos após tantos anos de estudo.

1. *Corpus* poético e variação gráfica no PA

Como exposto, selecionamos 250 textos medievais galego-portugueses para a realização deste artigo, 100 da vertente religiosa e 150 da linha profana. As CSM foram escritas na segunda metade do século XIII pelo rei de Leão e Castela, Dom Afonso X. A edição

[7] Cabe dizer que este artigo não visa à construção de uma análise diacrônica da língua, pois focaliza somente o PA trovadoresco. Trata-se, assim, de uma caracterização sincrônica de um momento do passado (Mattos e Silva, 1989).

geral feita por Mettmann (1986) tem 427 cantigas em louvor da Virgem Maria, sendo que sete das 427 são poemas repetidos (Leão, 2007; Massini-Cagliari, 2015). Segundo Leão (2002), as CSM retratam um panorama da vida religiosa do povo da Península Ibérica, visto que narram os hábitos do período, as doenças, os jogos, as profissões etc. Essas obras foram escritas para serem cantadas, a fim de entreter um público seleto, possivelmente constituído por cortesãos.

Parkinson (1998), seguido por Mongelli (2009) e XX (2015), pontua que as CSM se localizam em quatro códices e representam um processo de ampliação e de evolução, pois, inicialmente, a quantidade de poemas feitos foi a de 100 (pertencentes ao códice Toledo, To). A partir daí, as cantigas juntas compuseram um cancioneiro e passaram a ser organizadas como tal, isto é, adquiriram números, títulos e índice. O To é, pois, o mais antigo dos códices, além de ser o menor. O códice Rico (T) é o segundo a ser elaborado em virtude do desejo do rei de ampliar o códice inicial. É tido como o mais abundante em conteúdo artístico. O códice Florença (F), por sua vez, é uma cópia e constitui com o códice T o que ficou conhecido como *Códices das histórias*. Por fim, o códice Escorial Músicos (E) é o mais completo dos códices e é visto como uma cópia menos decorada do T.

A vertente profana da lírica medieval galego-portuguesa é bastante expressiva. Massini-Cagliari (2007) ressalta que o conjunto de textos profanos é formado por mais de 1.700 cantigas, cuja autoria cabe a cerca de 160 trovadores. Essa quantidade de poesias, no entanto, é apenas o que conseguiu sobreviver até os dias atuais. Acredita-se que a produção daquele momento tenha sido extremamente maior. Assim, o que se conservou da produção profana são oito testemunhos datados entre o final do século XIII e o século XVI.

Os cantares profanos dividem-se em três gêneros canônicos: *cantigas de amor*, *de amigo* e *de escárnio e maldizer*. Tais obras englobam diferentes épocas, localidades, categorias sociais, nacionalidades e temáticas, o que demonstra sua importância e representatividade. As obras que sobreviveram até hoje se encontram em três grandes cancioneiros e em cinco folhas avulsas em que há uma ou mais cantigas.

O primeiro dos cancioneiros profanos é o Cancioneiro da Ajuda (CA), datado do fim do século XIII e começo do século XIV. É o manuscrito mais contemporâneo aos trovadores, detém apenas *cantigas de amor* e é o único de origem ibérica (Massini-Cagliari, 2007). O segundo dos cancioneiros profanos é o Cancioneiro da Biblioteca Nacional de Lisboa (CBN). Tal códice é o único testemunho de aproximadamente 250 poemas, o que

faz com que ele seja o mais completo dos três cancioneiros, com, em média, 1.560 cantigas de mais de 150 trovadores. O terceiro dos cancioneiros é o Cancioneiro da Vaticana (CV), considerado como o irmão do CBN, pois ambos partilham um grande número de obras em comum. O CBN e o CV têm cantigas dos três gêneros. Ademais, os dois são cópias realizadas na Itália na primeira metade do século XVI.

Sobre a língua da época trovadoresca, que se situa entre os séculos XIII e XV (Mattos e Silva, 2006), é necessário dizer que o PA ainda não apresentava uma ortografia padrão instituída por lei, o que fazia com que a variação gráfica dominasse a documentação escrita daquela etapa, que sofria flutuação não apenas na grafia, mas também na morfologia e na sintaxe.

Segundo Mattos e Silva (2006), as variações presentes nos textos da Idade Média podem ser tidas como indicadores das mudanças que vieram a acontecer e que, a partir da padronização, passaram a ser suprimidas pelos materiais escritos, visto que somente uma parte das variantes usadas é de fato eleita para integrar o sistema ortográfico da língua. Logo, a padronização da grafia eliminou toda a diversidade de grafias existente no PA.

A falta de um padrão ortográfico naquela época fazia com que uma mesma palavra fosse, como mostraremos neste artigo, representada de diferentes formas. Tal diversidade de grafias é uma das características mais marcantes do português medieval. Uma mesma palavra era grafada de maneiras distintas, inclusive, dentro de um mesmo cantar, o que mostra que a alternância gráfica era uma atitude comum, empreendida inclusive pelo mesmo poeta dentro de uma mesma poesia. Às vezes, a variação escrita englobava um termo que aparecia mais de uma vez no interior da mesma estrofe, por exemplo.

Faraco (2008) pondera que o perfil da comunidade feudal derivou, em matéria de língua, em uma vasta diversificação. Em resposta ao variado mapa linguístico daquele período, emergiu um projeto padronizador, que visava instituir uma norma gráfica, eliminando, assim, as variações na representação da escrita. Mattos e Silva (2006) salienta que a inexistência de uma padronização ortográfica em PA pode refletir certas maneiras de uso daquela língua no cotidiano medieval. Logo, ao investigar a variação gráfica da língua dos trovadores, é possível constatar o *uso primeiro* da língua arcaica, pois a existência de vestígios das vozes do passado pode fornecer uma percepção do que era produzido naquele período. Assim, a variação do PA espelha traços do oral no escrito, traços esses que foram apagados da representação gráfica quando se estabeleceu a ideia (errada) de que existe uma única forma *correta* de escrever (e de empregar a língua), que seria o conjunto de regras prescritas nas gramáticas (Bagno, 2003).

2. Metodologia

O método de estudo assumido neste artigo baseia-se na análise da representação da grafia das consoantes róticas, laterais e nasais duplas coletadas nas 250 poesias que compõem o *corpus* selecionado e no estudo do comportamento desses elementos em todas as posições em que estão no interior da sílaba e da palavra. Além disso, consideram-se a separação em sílabas poéticas e o posicionamento da palavra no verso, que dão indícios da realização fonética da silabação na época e da posição do acento, fatores bastante relevantes para a determinação da existência (ou não) de geminação fonológica.

Primeiramente, empreendeu-se a coleta de todas as palavras com as referidas consoantes duplas nos 250 cantares galego-portugueses. Essa primeira coleta foi feita por meio do emprego de edições atualizadas das composições poéticas, a fim de facilitar o entendimento das palavras, a ordenação dos dados e a compreensão dos textos. Isso posto, adotamos as edições de Mettmann (1986) para as cantigas religiosas e de Lopes e Ferreira et al. (2011-) para as profanas.

Segundamente, após feita a coleta de dados, realizou-se a conferência de todos os termos encontrados nas edições fac-similadas das poesias, que são a reprodução fotográfica dos códices originais. Como, para desenvolver esta pesquisa, nos valemos dos dados de variações na escrita, o uso do fac-símile se faz primordial, visto que é somente a partir dele que temos acesso ao que, de fato, foi escrito pelos escribas daquela época da história.

Terminado o referido mapeamento, as ocorrências coletadas no *corpus* foram analisadas qualitativamente com base nos modelos fonológicos não-lineares. Na análise, usou-se apenas o que foi encontrado nos fac-símiles, ou seja, as variações gráficas (diferentes representações de uma mesma palavra). Para a realização deste estudo, alguns pontos foram de vital importância, a saber:

- Gênero em que a ocorrência foi encontrada (isto é, se ela estava em uma CSM ou em uma cantiga profana). Tal fato é importante, porque a escrita dos cancioneiros da vertente religiosa e o CA é mais antiga, o que faz com que estes guardem representações gráficas mais recuadas no tempo;
- Ambiente ocupado pela palavra dentro dos versos, pois termos podem aparecer grafados colados a outros que estão nas imediações, dados os padrões de segmentação adotados na época;
- Contexto da variação dentro da sílaba e da palavra;
- Segmentos antes e depois da variante, pois elementos que estão no entorno das

consoantes focalizadas podem influenciar ou afetar seu comportamento. Por exemplo: inserção de outras letras na palavra ou troca de grafemas; e

- Frequência dos tipos de variações, porque um tipo recorrente de ocorrência pode indicar que dada grafia era adotada por diferentes escribas em mais de uma época.

Para este artigo, faz-se de grande importância, ademais, a verificação do comportamento dos segmentos duplos dentro da palavra, uma vez que o que há antes e após a consoante dobrada pode influenciar seu comportamento. Por exemplo, para se ter uma consoante do tipo geminada, que vale por duas, é necessário que ela esteja em ambiente intervocálico. Assim, codas ocupadas antes de <rr>, <ll/lh> e <nn/nh> inviabilizam que tais elementos se portem como geminados; o mesmo pode ser observado quando ocorrem ditongos antes de consoantes duplas, o que também faz com que a coda da sílaba anterior não seja leve. Em outras palavras, para se ter uma consoante do tipo geminada, a coda da sílaba antecedente precisa estar vazia.

Como se vê, como a consoante do tipo geminada vale por duas, ela apresenta dois pontos de ancoragem na estrutura interna da sílaba, o que faz com que a coda da sílaba precedente deva estar vazia para poder abrigar parte desse segmento duplo. Portanto, para além das variações da grafia, o contexto em que a consoante dobrada aparece é um dos pontos determinantes para que, neste estudo, possamos determinar seu estatuto fonológico no PA.

3. Fundamentação teórica

A fundamentação teórica desta pesquisa tem como alicerce os modelos fonológicos não-lineares, em especial dois desses modelos: métrico (Hayes, 1995) e autossegmental (Goldsmith, 1976). Na presente seção, retrataremos uma breve discussão acerca dos conceitos de sílaba e de geminação. Nosso objetivo não é o de esgotar tais reflexões, mas apresentar um panorama geral, tendo em vista o caráter sucinto deste artigo.

Embora existam modelos teóricos mais modernos na área de Fonologia, optamos por tal perspectiva teórica neste trabalho pelos novos modelos não apresentarem diferentes *insights* em relação à geminação. Portanto, mesmo se tratando de pesquisas elaboradas nas décadas de 80 e 90, as teorias fonológicas não-lineares são a melhor opção para este estudo.

Na Teoria Métrica, segundo os estudos de Selkirk (1982), os constituintes silábicos estão organizados de maneira hierárquica. Então, tem-se o ataque (A), também conhecido como *onset* (O), e a rima (R). A rima constitui-se por um núcleo (Nu) e por uma coda

(Co). O ataque silábico, para Câmara Jr. (1985 [1970]), é o ponto de aclive, o núcleo é o ápice e a coda é o declive. Assim sendo, o que se tem é um ponto de ascensão, um pico e um momento de perda de força.

De acordo com a teoria formulada por Selkirk (1982), qualquer posição, exceto o núcleo, pode ser vazia, isto é, ataque e coda são periféricos e opcionais. Ademais, há um relacionamento mais estreito entre a vogal do núcleo e a consoante da coda do que entre essa vogal e a consoante do ataque, já que que núcleo e coda preenchem juntos a rima da sílaba. O núcleo do PB é sempre uma vogal e não pode estar desocupado para que uma sílaba exista. Já ataque e coda são contextos em que há consoantes no PB, sendo que ambos podem ser simples, ramificados ou vazios.

Outra questão relevante para nós diz respeito ao peso silábico. Collischonn (2005) relata que sílabas pesadas são aquelas em que há mais de um elemento na rima, isto é, são sílabas que, necessariamente, têm rima ramificada. Já sílabas leves têm somente um segmento na rima, mais precisamente uma vogal no núcleo. As sílabas V e CV são, logo, leves. Para o peso da sílaba, o ataque não é relevante, ou seja, ataques ramificados, como *pl* e *br*, por exemplo, não fazem com que a sílaba seja pesada.

Hyman (1975) assume a unidade silábica como uma unidade de peso, denominada como mora, cujo símbolo é μ. De acordo com essa proposta teórica, uma sílaba pesada tem duas moras e uma sílaba leve tem uma mora. Logo, sílabas CVV são sempre bimoraicas ou pesadas e sílabas CVC podem variar segundo a estrutura de cada língua, ou seja, falares que contam os segmentos apenas do núcleo a consideram monomoraica e os que contam os segmentos da rima a entendem como bimoraica (Massini-Cagliari, 1999). O máximo de moras que uma sílaba pode ter é duas, logo, mesmo que uma consoante esteja unida a sílabas CVC ou CVV, ela será integrada à última mora e não fornecerá mais peso à sílaba.

Um outro conceito importante para este estudo é um fenômeno conhecido como duração ou alongamento compensatório. Esse fenômeno fonológico consiste na permanência da duração de um dado segmento quando ele sofre supressão, isto é, quando uma regra fonológica ocasiona, por vários motivos, o apagamento de determinado elemento, a duração dele pode ser preservada e associada a um outro segmento que está nas imediações (Hogg; McCully, 1987).

Acerca da geminação, Goldsmith (1990) explica que, se duas consoantes iguais estão na posição intervocálica, elas pertencem a sílabas diferentes, isto é, um dos segmentos preenche a posição de coda da sílaba anterior enquanto o outro ocupa o ataque da sílaba

subsequente. Sobre o peso de tais consoantes, tem-se que elas têm uma mora, pois parte do elemento preenche a coda, travando a sílaba anterior, e parte preenche o ataque, contexto que não apresenta mora e, portanto, não fornece mais peso à sílaba (Perlmutter, 1995).

As geminadas têm duas propriedades fonológicas tidas como invioláveis para Cedeño e Morales-Front (1999): integridade e inalterabilidade. Logo, não é possível separar os elementos que compõem uma geminada por regras de epêntese e não é possível alterar uma parte sem que, necessariamente, se altere a outra. Consoantes geminadas, portanto, são um único elemento que vale por dois (Perlmutter, 1995), o que faz com que seja impossível inserir segmentos no interior dos segmentos duplos, que se mostram resistentes aos processos fonológicos que, por diferentes razões, poderiam se aplicar a eles.

4. Levantamento e análise dos dados

Nas 250 cantigas que constituem nosso *corpus*, foram coletadas:

- 672 palavras grafadas com <rr>;
- 1.893 palavras grafadas com <ll/lh>; e
- 899 palavras grafadas com <nn/nh>.

Todos os dados de nasais duplas foram localizados no meio do termo. Contudo, o mesmo não foi apurado em relação aos casos de consoantes róticas e laterais dobradas. Das 672 palavras escritas com róticas duplas, 45 delas apresentavam <rr> no começo da palavra (627 no meio da palavra, portanto). Já das 1.893 palavras escritas com laterais duplas nas obras do medievo, 823 dados compreendem pronomes pessoais oblíquos/clíticos (*lle/lhe*, por exemplo), restando 1.070 ocorrências de laterais no interior da palavra.

No que diz respeito às róticas dobradas da língua arcaica, foram encontrados vários tipos de variação na amostragem analisada, no entanto, neste artigo, em virtude de seu caráter sucinto, nos deteremos em três tipos, a saber:

Quadro 1 – Tipos de variação das róticas duplas do PA

Tipo de variação	Palavra sem variação	Palavra com variação	Cancioneiro com variação	Cancioneiro sem variação

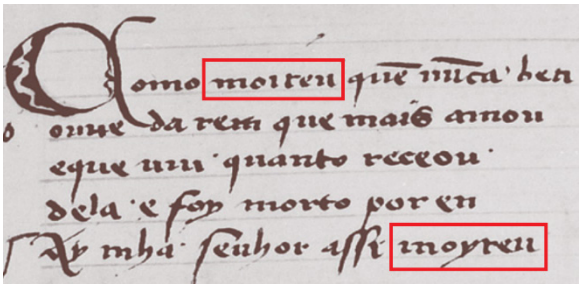
RR-IR-YR[8]	morrer	moireu/ moyreu[9]	CBN	CA
RR-LR	carreira	calreyra	CBN	CV
SR-RR	Israel	Irrael	To/E	-

Fonte: elaboração própria.

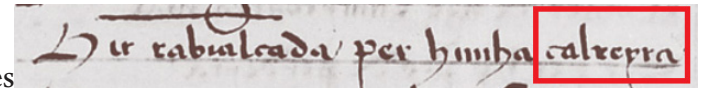
Os tipos de variação gráfica apresentados no quadro 1 compreendem argumentos a favor da geminação das róticas duplas do PA, pois demonstram que tais consoantes eram interpretadas pelos escribas do momento trovadoresco como formadas por duas unidades. O conceito de peso silábico é imprescindível para o entendimento dessas variações, pois a troca de um elemento por outro, nos três dados, ocorreu na primeira parte da geminada, isto é, na consoante que ocupa a coda da sílaba anterior, local que contribui para o peso da sílaba e carrega uma mora.

A seguir, exibimos um exemplo de cada tipo de variação retirado dos fac-símiles.

Figura 1 – Dado *moireu/moyreu* – Trecho da cantiga *Como morreu quem nunca bem*, de Paio Soares de Taveirós



Fonte: Edição fac-similada do códice da Biblioteca Nacional de Lisboa – Colocci-Brancuti (1982, p. 150).

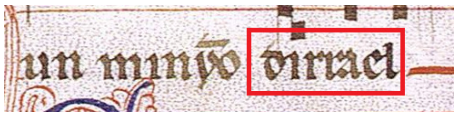
Figura 2 – Dado *calreyra* – Trecho da cantiga *Maria Negra vi eu, em outro dia*, de Pero Garcia Burgalês 

Fonte: Edição fac-similada do códice da Biblioteca Nacional de Lisboa – Colocci-Brancuti (1982, p. 625).

[8] Foram encontrados vários dados da variação RR-IR-YR, entretanto, optamos por eleger apenas um exemplo para este estudo. Para mais informações sobre essa variação, consultar Barreto (2019).

[9] Nos cantares aqui analisados, a variação <i>/<y> aparecia com bastante frequência, uma vez que tais vogais eram alternadas na época por motivos estilísticos. Tais segmentos, assim, correspondiam ao mesmo som e não causavam, no PA, mudança no significado das palavras.

Figura 3 – Dado *d'Irrael* – Trecho da CSM 4



Fonte: Edição fac-similada do códice Toledo (2003, p. 12r).

Com relação ao primeiro tipo de variação (RR-IR-YR), o que se verifica é que, ao grafar, nos códices, <rr> como <ir/yr>, cria-se um ditongo na primeira sílaba da palavra (*oi/oy*). Assim, o que se tem é uma palavra, no caso *m**oi**reu/m**oy**reu*, com dois ambientes ocupados na rima, um no núcleo e um na coda[10], o que faz com que a primeira sílaba da palavra seja pesada, com coda preenchida. Como *morreu* e *moireu/moyreu* são variações de uma mesma palavra, é preciso que haja uma correspondência de peso entre essas representações gráficas. Para que isso se verifique nos casos da referida variação, <rr> precisa ser interpretado como um segmento geminado, uma vez que *mor* e *moi/moy* apresentariam o mesmo peso silábico, com <r> e a semivogal do ditongo (*i* ou *y*) ocupando a coda da primeira sílaba da palavra e tornando-a pesada.

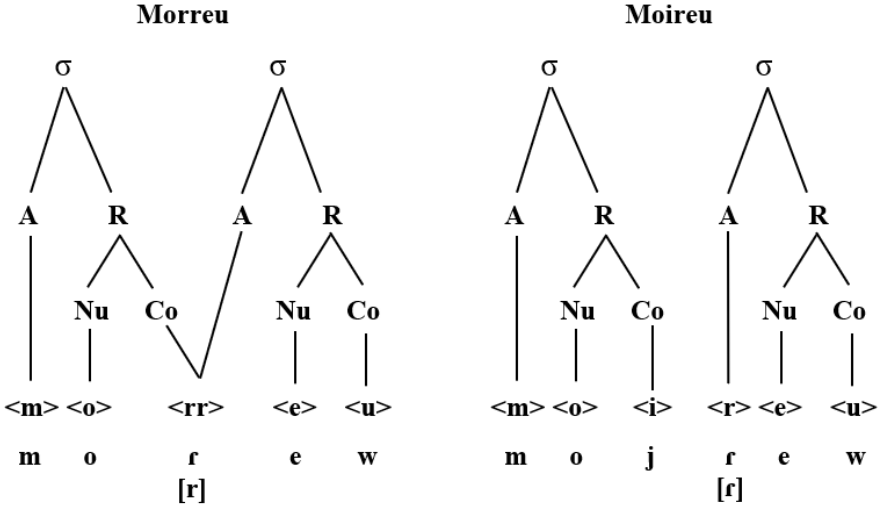
Dessa maneira, considerar que a rótica dupla intervocálica é um elemento geminado é a única possibilidade capaz de explicar esse tipo de variação encontrado, pois o encadeamento de dois sons diferentes (uma vogal *i/y* somada a uma consoante grafada como <r>) se equivale em peso (uma mora na coda) e em função a uma rótica representada como dupla.

Tal constatação não advoga que, necessariamente, existisse apenas um som rótico no PA trovadoresco. O que essa interpretação revela é que, mesmo que existissem dois sons róticos na época em questão, o tepe e a vibrante múltipla (como há ainda hoje no português europeu), eles não estavam em oposição fonológica e consistiam na produção de um mesmo fonema da língua. Essa hipótese foi postulada para o PB e para o Português Europeu (PE) por Câmara Jr. (1953) e simboliza uma continuação ao sistema latino, em que havia apenas um fonema rótico, que podia ou não ser geminado.

Para ilustrar o exposto, apresentaremos o exemplo (1), que demonstra que, como *morreu* e *moireu/moyreu* são variações da mesma palavra, o peso silábico é igual nos dois termos. Logo, para que haja equivalência entre as variantes, <rr> se reparte, passando a ocupar dois ambientes, simultaneamente, um na coda e um no ataque da sílaba seguinte.

[10] Neste estudo, assumimos Zucarelli (2002) acerca da configuração dos ditongos do PA. Segundo tal pesquisadora, o núcleo do português medieval apresenta apenas um local de ancoragem, ou seja, a semivogal do ditongo preenche a coda da sílaba, comportando-se como uma consoante que trava a sílaba.

(1)[11]



Acerca do segundo tipo de variação (RR-LR), foi localizado um único dado nas cantigas analisadas: a variação *carreira/calreyra*. Esse caso também simboliza um argumento a favor da geminação da consoante rótica dupla intervocálica do PA, porque revela que, no desdobramento de uma geminada em duas partes, ou, no movimento contrário, quando acontece a união de dois segmentos em uma geminada, o processo conhecido como alongamento compensatório se mostra de grande importância.

Em um processo fonológico, quando um elemento, que está associado a uma mora, sofre apagamento, a mora se mantém e passa a outro segmento. Desse modo, no termo *calreyra*, caso /l/ seja suprimido, a mora ligada a esse som (que apresenta uma mora em razão de estar na coda) permanecerá e será vinculada a consoante seguinte, no caso, à rótica <r>, que se torna longa ou geminada. Como se vê, o peso silábico se conserva constante, embora os segmentos possam ser substituídos por outros por diferentes motivos. Logo, o alongamento compensatório é um mecanismo capaz de compensar o desaparecimento de segmentos de uma dada palavra.

Assim como ocorre com a variação RR-IR-YR, em *ca**l**reyra*, o segmento lateral <l> está na mesma posição que o primeiro <r> da rótica dupla de *ca**r**reira*, ou seja, ocupa o ambiente de coda da primeira sílaba do vocábulo, tornando-a pesada, com rima ramificada. Assim sendo, os grafemas <rr> e <lr> apresentam uma mesma duração.

Por fim, no que se refere ao terceiro tipo de variação (SR-RR), também se tem, nas

[11] Olhando o referido exemplo de cima para baixo, temos: o símbolo **σ** se refere à "sílaba"; já as letras **A** e **R** dizem respeito ao "ataque" e à "rima". A "rima" da sílaba se reparte em "núcleo" (**Nu**) e "coda" (**Co**). Nesse caso, /morew/ e /mojɾew/ consistem na representação fonológica da palavra "morreu". Em seguida, os símbolos < > retratam que se trata da representação ortográfica do termo.

obras investigadas, apenas um caso: a variação *Israel/Irrael*. No referido caso, a desassociação de <s> deu origem a uma rótica dupla por meio do processo de alongamento compensatório. A fricativa <s>, da primeira sílaba de *Israel*, encontra-se na coda e, por isso, carrega uma mora. Seu apagamento não acarreta, entretanto, o desaparecimento da unidade de peso, que é mantida e associada a um outro segmento. Como já pontuamos, quando um elemento é apagado em razão da aplicação de regras fonológicas, sua duração é preservada e se vincula a outro *slot* que está no entorno. Assim sendo, em *Irrael*, a supressão da fricativa <s> gerou o alongamento da rótica, que incorporou a mora da coda anterior e passou a ser dupla ou geminada. Portanto, o aumento temporal da rótica, que passou de simples para dupla, compensou a omissão de <s>.

Os dados de variação retratados com relação às róticas duplas permitem concluir que tal consoante podia ser considerada como geminada, na posição intervocálica, no nível fonológico, no PA. Esse segmento, como demonstram os casos analisados, apresenta uma duração maior do que seu correspondente simples *r*, valendo por dois. Além disso, não foram localizados ditongos antes de <rr>, o que favorece essa interpretação. Logo, no PA, assumimos a existência de apenas um fonema rótico com duas variantes, uma simples (que podia ser escrita como <r> e <rr>[12]) e uma geminada, em contexto intervocálico, representada na escrita como <rr>.

Em relação às laterais duplas do português dos trovadores, é importante ressaltar que os segmentos em questão não ocorrem exclusivamente no contexto intervocálico, visto que podem aparecer depois de sílabas travadas por consoante e no começo da palavra[13].

Assim como em relação às róticas duplas, não foram encontrados dados em que ditongos integravam a sílaba anterior à lateral <ll/lh> e à nasal <nn/nh>, fato que favorece a consideração da geminação dessas consoantes da língua medieval, porque, para ser uma geminada, a coda da sílaba anterior ao segmento duplo precisa, obrigatoriamente, estar vazia para poder abrigar parte desse elemento complexo.

Outro forte argumento a favor da geminação da lateral dupla – comprovado também em relação às róticas e às nasais duplas – é o fato de que não foram localizadas proparoxítonas com tais segmentos duplos. Massini-Cagliari (2015) pontua que a língua arcaica era sensível ao peso das sílabas para a localização do acento tônico. Assim, sílabas

[12] Em decorrência da natureza sucinta deste trabalho, as estudiosas optaram por não apresentar uma análise dos casos em que <rr> não pode ser considerado como um segmento geminado, mas, de forma geral, tais consoantes ocorrem em contextos que não o intervocálico, como início de palavra e após sílaba travada por elemento consonantal. Para informações mais detalhadas a esse respeito, consultar Barreto (2019, 2023).

[13] Para uma análise dos referidos casos, ler Barreto (2023).

pesadas localizadas na última ou na penúltima sílaba da palavra atraem para si o acento principal.

Dessa maneira, a ausência de proparoxítonas com róticas, laterais e nasais duplas reforça a hipótese de geminação dessas consoantes, do ponto de vista fonológico, porque a presença de uma consoante do tipo geminada no interior do vocábulo impede que o acento retroceda para a antepenúltima sílaba.

Para exemplificar, podemos pensar nas palavras *beze**rr**o*, *pare**ll**a* e *Espa**nn**a*. As três têm uma consoante duplicada na última sílaba (gráfica) e são paroxítonas em PA. Ao assumirmos que *rr*, *ll/lh* e *nn/nh* são elementos geminados, temos que a presença dos elementos duplos em foco fez com que a última sílaba das palavras estudadas fosse leve (**CV**), com uma única mora (*ro*; *la*; *na*), o que, de acordo com Massini-Cagliari (2015), é um fator decisivo para a atribuição do acento na língua dos trovadores. Ademais, as palavras aqui focalizadas seguem um dos padrões canônicos do PA, que, conforme as análises empreendidas por Massini-Cagliari (1999, 2015), consiste em paroxítonas terminadas em sílaba leve.

Assim, sendo *rr*, *ll/lh* e *nn/nh* segmentos geminados, a penúltima sílaba das palavras em questão são pesadas, com coda preenchida (*be-**zer**-ro*, *pa-**rel**-la* e *Es-**pan**-na*) e, em razão disso, não pode ser *pulada*, ou seja, o acento não pode figurar na antepenúltima sílaba. Como se vê, a existência das consoantes geminadas impede que o acento esteja na primeira sílaba (*be*; *pa*; *Es*), posto que seria preciso *pular* uma sílaba pesada para conseguir dar origem a uma proparoxítona, o que é irregular, em termos acentuais, na língua dos trovadores.

Sendo assim, ao assumirmos que tais segmentos são geminados, no nível fonológico, na etapa arcaica, explica-se a inexistência de proparoxítonas com róticas, laterais e nasais dobradas no começo da última ou da penúltima sílaba da palavra no PA.

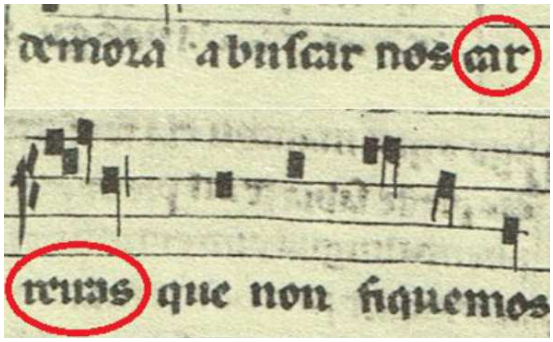
Além da ausência de ditongos antes de segmentos róticos, laterais e nasais duplos no PA e de proparoxítonas com essas consoantes no início da última ou penúltima sílaba, a divisão das sílabas de alguns termos nos cancioneiros consiste em um outro argumento relevante a favor da geminação intervocálica das referidas consoantes dobradas naquele período.

Nos fac-símiles analisados, foram encontrados registros de ocorrências em que o próprio responsável pela cópia do documento grafou <rr>, <nn/nh> e <ll/lh> de forma separada, isto é, retratou uma parte do elemento na posição de coda da sílaba precedente e uma parte no ambiente de ataque da sílaba seguinte. Esse tipo de dado foi encontrado

no interior da palavra entre vogais e na junção de vocábulos em contexto intervocálico. Ademais, foram localizados dados em mais de um cancioneiro, ou seja, há ocorrências nas CSM e nas cantigas profanas.

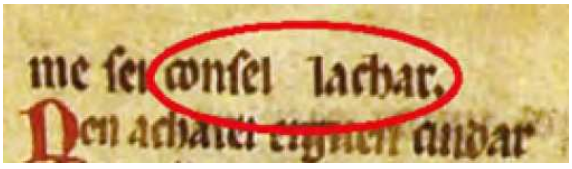
Como são muitos exemplos, optamos por apresentar somente um de cada para ilustrar o exposto. Para expressar a diversidade de códices nos quais os dados de separação silábica foram encontrados, apresentamos casos de cancioneiros diferentes.

Figura 4 – Dado *car reiras* – Trecho da CSM 71
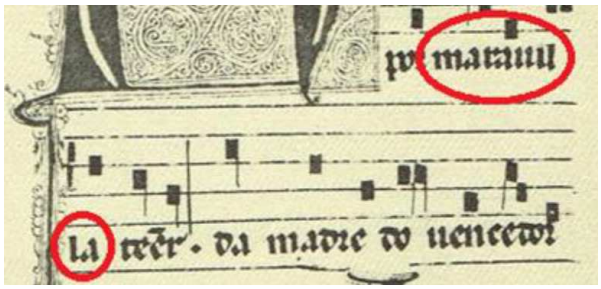


Fonte: Edição fac-similada do códice Escorial Músicos, editada por Anglés (1964, p. 89v).

Figura 5 – Dado *consel achar* – Trecho da cantiga *Muitas vezes em meu cuidar*, de João Soares Somesso



Fonte: Edição fac-similada do código da Biblioteca da Ajuda (1994, p. 16).

Figura 6 – Dado *maravil la* – Trecho da CSM 27



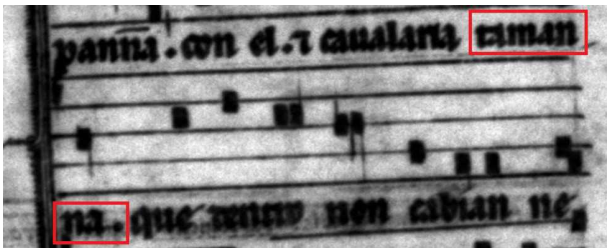
Fonte: Edição fac-similada do códice Escorial Músicos, editada por Anglés (1964, p. 51v).

Figura 7 – Dado *sen nor* – Trecho da cantiga *Nom soube que x'era pesar*, de Vasco Gil



Fonte: Edição fac-similada do códice da Biblioteca da Ajuda (1994, p. 155).

Figura 8 – Dado *taman na* – Trecho da CSM 69



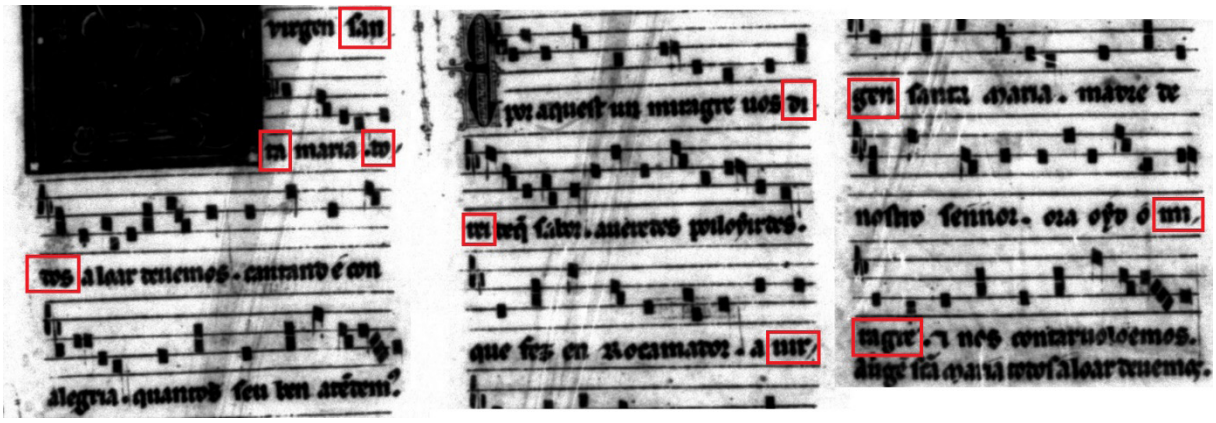
Fonte: Microfilme do códice Escorial Rico, cedido pela Biblioteca do Mosteiro de El Escorial. Este microfilme pertence ao grupo de pesquisa "Fonologia do Português: Arcaico & Brasileiro".

Como se vê, os dados localizados ocorreram em final de verso e em meio de verso. Além disso, todos os casos situam-se na parte inicial da cantiga, espaço destinado às notações musicais das composições poéticas[14]. As figuras 4 a 8 evidenciam que o escriba, provavelmente, desejava destacar na escrita que o segmento duplo ocupava duas sílabas diferentes. É importante ressaltar que os escribas daquela época tinham uma intuição bastante apurada em relação à silabação das palavras da língua, pois segmentavam os termos para encaixá-los nos versos e nas músicas. Em análises realizadas ao longo dos anos, as pesquisadoras averiguaram que a separação das sílabas empreendida pelos escribas seguia, como regra geral, a unidade silábica. Não era comum, assim sendo, uma divisão que separasse, por exemplo, ataque e núcleo ou núcleo e coda. Na figura 9, apresentamos um trecho da CSM 61 em que o escriba segmentou algumas palavras em contexto de final de linha.

[14] Cabe expor que o CA, embora apresente um local reservado para a notação musical, nunca chegou a receber tais inscrições. Conforme Massini-Cagliari (2007), esse códice é incompleto em vários sentidos, pois, além da ausência das notações musicais, a decoração, as rubricas e as miniaturas não foram terminadas.

http://doi.org/10.5007/1984-8420.2025.e101224

Figura 9 – Exemplos de separação silábica – Trecho da CSM 8[15]



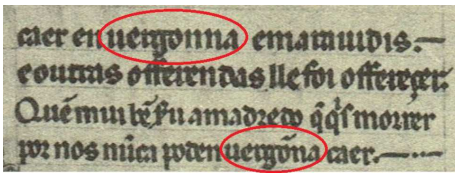
Fonte: Microfilme do códice Escorial Rico, cedido pela Biblioteca do Mosteiro de El Escorial. Este microfilme pertence ao grupo de pesquisa "Fonologia do Português: Arcaico & Brasileiro".

Os dados de segmentação gráfica das róticas, laterais e nasais duplas, logo, mostram que os escribas do PA, muito possivelmente, entendiam essas consoantes dobradas como compostas por duas unidades de tempo. É claro que não é possível descartar a possibilidade das segmentações dos elementos duplos serem resultado de práticas de escrita usadas pelos escribas nos *scriptoria* daquele período, lugares em que esses trabalhadores compartilhavam conhecimentos sobre o PA medieval e interagiam entre si durante a elaboração dos códices arcaicos, no entanto, tendo em vista a quantidade de dados encontrados e o fato de que eles figuram em cantigas e em códices diferentes (o que revela que tais vocábulos foram redigidos por mais de uma pessoa e em períodos distintos), defendemos que os dados são fortes indícios a favor da geminação das róticas, laterais e nasais duplas do PA.

Para além das evidências pontuadas até aqui, foram encontrados nas obras trovadorescas casos de uma variação que ocorreu especificamente com as nasais dobradas e que compreendem mais um indício a favor da interpretação desses segmentos como geminados no nível fonológico da língua. Essa variação consiste na representação gráfica da nasal *nn* como til + *n* e de *nh* como til + *h*. Esses casos apareceram dentro da mesma cantiga em versos diferentes e no mesmo verso de uma cantiga, mas em códices diferentes. A seguir, exibimos um exemplo de til + *n* e de til + *h* para exemplificar o exposto.

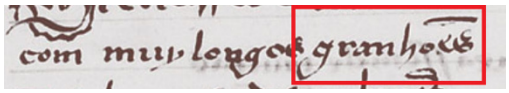
[15] Palavras em destaque: *santa*, *todos*, *direi*, *virgen* (*virgem*) e *miragre* (*milagre*).

Figura 10 – Dados *vergonna* e *vergõna* – Trecho da CSM 63



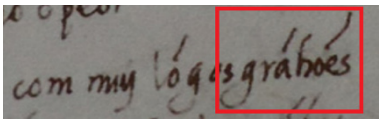
Fonte: Edição fac-similada do códice Escorial Músicos, editada por Anglés (1964, p. 82r).

Figura 11 – Dado *gra**nh**ões* – Trecho da cantiga *O genete*, de Afonso X[16]



Fonte: Edição fac-similada do códice da Biblioteca Nacional de Lisboa – Colocci-Brancuti (1982, p. 491).

Figura 12 – Dado *grã**h**ões* – Trecho da cantiga *O genete*, de Afonso X



Fonte: Edição fac-similada do Cancioneiro Português da Biblioteca Vaticana (1973, p. 74).

O til é um assunto muito complexo nas cantigas arcaicas. Por apresentar várias funções[17] diferentes, a dúvida que se apresenta é: o que a marca de til está simbolizando? Seria uma marca de abreviação ou de nasalização? Massini-Cagliari (2015) explica que há duas hipóteses acerca do referido tema: a primeira é radical, visto que defende que toda ocorrência de til é uma marca de abreviação; a segunda é menos radical, pois entende que, quando o til está sobre vogais, a marca consiste em um sinal de nasalização. Neste estudo, nos filiamos à segunda hipótese.

Ainda segundo Massini-Cagliari (2015), em um modelo derivacional, é possível assumir a nasalização como um traço flutuante nos casos de alocação de til sobre vogais, já que pode se realizar como ataque quando essa posição não estiver previamente preenchida (como em *bõa* > *bo.na*) ou como nasalização da vogal (pelo espraiamento do traço nasal sobre a vogal). Massini-Cagliari (2015) pontua que a melhor realização da nasal

[16] *Granhões/grannões*: quer dizer *barba* (Mettmann, 1972, p. 157).
[17] Acerca das funções do til nos documentos do medievo, consultar Massini-Cagliari (1998).

seria em contexto de ataque e não como traço nasal da vogal, entretanto, conforme os dizeres da pesquisadora, a nasalização se realizaria no ataque quando esse contexto estivesse livre ou vazio. Mas e quando o ataque está preenchido, como em *vergõ**n**a* e *grã**h**ões*?

Em relação à posição de ataque do PA, Massini-Cagliari (2015) destaca que, a exemplo do que se tinha no latim e do que acontece ainda hoje no PB, somente /p, b, t, d, k, g, f, v/ podem aparecer na primeira posição do ataque. No segundo contexto, ou seja, quando se tem um ataque ramificado, apenas consoantes líquidas /l, ɾ/ são permitidas.

Dessa maneira, podemos entender que casos como *vergõna* e *grãhões* representam casos em que a nasal <n> foi "trocada" pelo til na representação gráfica, que passou a aparecer sobre a vogal anterior à consoante, preservando seu traço nasal. Em razão da configuração do ataque da sílaba da língua arcaica, a nasalização não pode se realizar nesse ambiente, pois já se encontra ocupado por *n* (*vergõ**n**a*) e por *h* (*grã**h**ões*). Então, o traço nasal só pode se realizar no contexto que antecede o ataque, isto é, na coda da sílaba anterior.

Tais dados, muito frequentes no *corpus*, revelam que os escribas do PA viam *nn/nh* como uma consoante que vale por duas. As ocorrências investigadas evidenciam que tal segmento era entendido como duplo em sua estrutura, pois, mesmo quando os escribas não grafavam <nn/nh> nos cancioneiros, registravam tal elemento como a soma de duas unidades (til + *n* ou til + *h*). A variação em questão, como mostrado, também é um importante argumento a favor da existência da geminação da nasal dupla intervocálica do PA.

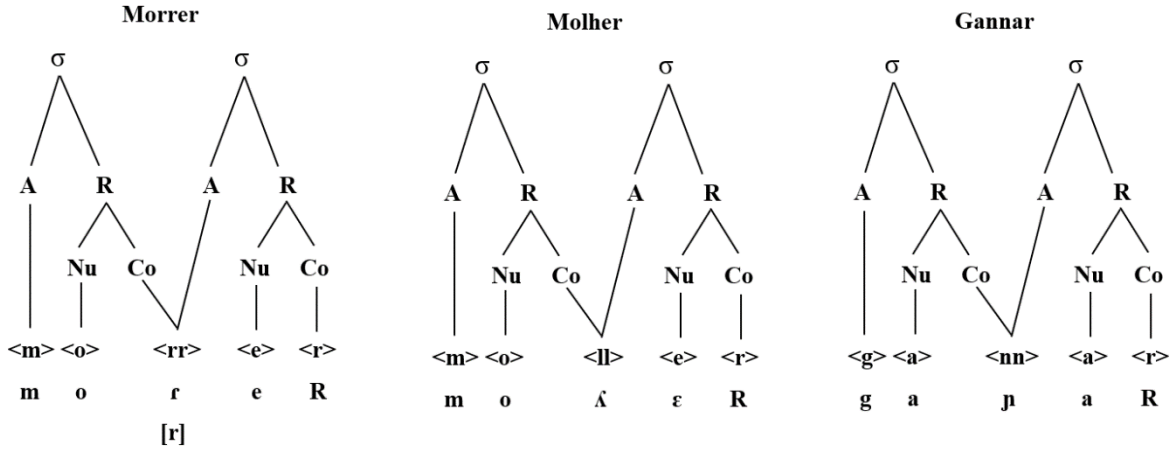
Ainda acerca das nasais *nn/nh*, convém observar que tal configuração complexa também pode ser verificada no português brasileiro, uma vez que, dos segmentos nasais ([m, n, ɲ, ŋ]) do PB, [ɲ] é o único que, com relação ao lugar de articulação, caracteriza-se como sendo, ao mesmo tempo, coronal e dorsal (Cagliari, 1998). Então, enquanto [m] é [+ labial], [n] é [+ coronal] e [ŋ] é [+ dorsal], /ɲ/ é simultaneamente [+ coronal] e [+ dorsal], o que retrata que tal consoante porta uma complexidade articulatória que já tinha sido notada pelos escribas do PA.

Como se vê, os dados de variação gráfica localizados nos fac-símiles medievais retratam que a consideração da existência da geminação é a que melhor explica a natureza das consoantes róticas, laterais e nasais duplas do PA. Por meio das ocorrências analisadas, entendemos que na época arcaica tais elementos podiam ser interpretados, fonologicamente, como geminados. Para o PB atual, Wetzels (2000, p. 6) admite a geminação dessas mesmas consoantes dobradas. De acordo com o pesquisador, além das róticas *rr*, as laterais e as nasais duplas devem ser vistas como geminadas no nível fonológico, tendo em vista que:

> as soantes palatais /ñ[18], ʎ/ do Português Brasileiro (PB) se comportam, sob muitos aspectos, diferentemente das soantes não palatais. Em se tratando da nasalização da vogal precedente, a nasal-palatal se comporta como se fosse uma consoante na coda, embora ela ocorra exclusivamente em posição intervocálica. Acrescentado a isso, as sílabas que precedem uma soante palatal são sempre leves, como pode ser observado não só na completa ausência de rimas pesadas precedendo uma soante palatal intervocálica, como também no algoritmo de silabificação, que cria hiato no caso de seqüências de Vogal + Vogal Alta que precedem /ñ, ʎ/ (*moinho*, *faúlha*), enquanto antes de /m, n, r, l/, os ditongos decrescentes surgem obrigatoriamente (*queima*, *baila*). Além disso, se uma soante palatal ocorre como *onset* de uma sílaba em final de palavra, como em *alcunha*, o acento da palavra nunca cai na antepenúltima sílaba, embora o acento proparoxítono seja um padrão possível no PB.

Nossa análise mostra que as consoantes duplas em questão, em PA, apresentam o mesmo comportamento descrito pelo autor para o PB. Assim, nos filiamos ao estudioso ao defendermos que, na época arcaica do português, as consoantes róticas, laterais e nasais duplas, representadas na escrita como <rr>, <nn/nh>, <ll/lh>, podiam ser interpretadas como geminadas, uma vez que eram segmentos que valiam por dois e preenchiam, simultaneamente, coda e ataque na estrutura silábica, isto é, tinham uma duração maior se comparados às correspondentes simples. Portanto, sendo consoantes do tipo geminada, os moldes silábicos que melhor representam tais elementos são os retratados em (2)[19]:

(2)[20]



[18] Símbolo usado pelo autor para [ɲ].

[19] Apesar de a geminada travar a penúltima sílaba nos exemplos em (2), o padrão acentual das palavras observadas é oxítono, por apresentarem todas elas uma sílaba travada por rótica, na posição final de palavra.

[20] Moldes silábicos das palavras morrer, mulher e ganhar da língua portuguesa moderna. No exemplo (2), /moɾeR/, /moʎɛR/ e /gaɲaR/ consistem na representação fonológica dos termos em questão.

Considerações finais

Os dados de variações gráficas analisados mostraram que os resultados alcançados nesta pesquisa foram relevantes para estabelecer o estatuto fonológico das consoantes róticas, laterais e nasais duplas do momento trovadoresco da língua portuguesa. Por meio da análise das edições fac-similadas das 250 cantigas arcaicas que compõem o nosso *corpus*, foi possível concluir que essas três consoantes dobradas apresentavam o mesmo comportamento fonológico naquela fase da língua e podem ser interpretadas como geminadas, no nível fonológico, quando se situam no ambiente intervocálico, já que preenchiam duas posições na estrutura interna da sílaba do PA e apresentavam uma duração maior do que suas correspondentes simples *r*, *l* e *n*.

Com relação às três consoantes dobradas, não foram localizados dados de ditongos antes de tais segmentos nem palavras proparoxítonas com <rr>, <ll/lh> e <nn/nh> no ataque da última ou da penúltima sílaba. Essas descobertas contribuem para nossa interpretação, posto que, como demonstramos ao longo desta análise, para se ter uma consoante geminada, é preciso que a coda da sílaba anterior ao segmento duplo seja aberta, ou seja, leve, sem rima ramificada, para que o ambiente de coda da sílaba antecedente possa abrigar parte da geminada, que, por ser complexa, compreende um elemento que vale por dois (Perlmutter, 1995).

A finalidade deste artigo, para além de divulgar as descobertas encontradas nas cantigas, o potencial desses documentos para a realização de pesquisas acerca da história do português e a importância dos documentos remanescentes do PA, consiste em evidenciar que essas reflexões apenas foram possíveis por meio das teorias fonológicas não-lineares, que estabeleceram que o peso das sílabas e a estrutura das consoantes (que podem ser vistas como simples ou geminadas) dependem da distribuição dos segmentos no constituinte silábico. Ao mesmo tempo, as considerações também partiram da comparação dos dados em edições críticas e fac-similadas, em direção ao desejado equilíbrio (já ensinado por Ataliba Teixeira de Castilho) entre tradição e inovação nos estudos históricos.

Agradecimentos

Agradecemos à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Processo: 2022/09590-4) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Processo: 304657/2023-9) por viabilizarem a realização deste artigo.

Referências

AFONSO X, O SÁBIO. *Cantigas de Santa María:* edición facsímile do Códice de Toledo (To). Biblioteca Nacional de Madrid (Ms. 10.069). Vigo: Consello da Cultura Galega, Galaxia, 2003.

ANGLÉS, H. *La música de las Cantigas de Santa María del Rey Alfonso el sabio:* fac-símil, transcripción y estudio critico por Higinio Anglés. Barcelona: Diputación Provincial de Barcelona; Biblioteca Central; Publicaciones de la Sección de Música, 1964.

BAGNO, M. Norma lingüística & preconceito social: questões de terminologia. *Veredas.* Juiz de Fora, v. 5, p. 71-83, ju./dez. 2003.

BARRETO, D. A. R. J. *Estudo das consoantes róticas nas cantigas medievais galego-portuguesas.* Araraquara, 2019. Dissertação (Mestrado em Linguística e Língua Portuguesa) – FCL/UNESP.

BARRETO, D. A. R. J. *Análise do comportamento fonológico das consoantes líquidas do português dos trovadores.* Araraquara, 2023. Tese (Doutorado em Linguística e Língua Portuguesa) – FCL/UNESP.

CAGLIARI, L. C. *Fonologia do Português - Análise pela Geometria de Traços.* 2ª ed. revista. Campinas: edição do autor, 1998.

CÂMARA JUNIOR, J. M. *Para o estudo da fonêmica portuguesa.* 1ª ed. Rio de Janeiro: Padrão, 1953.

CÂMARA JUNIOR, J. M. *Estrutura da língua portuguesa.* 15ª ed. Petrópolis: Vozes, 1985. [1ª ed. 1970]

CÂMARA JUNIOR, J. M. *História e estrutura da língua portuguesa.* 1ª ed. Rio de Janeiro: Padrão, 1975.

CASTILHO, A. T. de. *Gramática do Português Falado. Volume 1:* A Ordem. Campinas: Editora da Unicamp, 2003.

CASTILHO, A. T. de. *História do Português Brasileiro:* O Português Brasileiro em seu contexto histórico. São Paulo: Editora Contexto / Fapesp, 2018.

CANCIONEIRO Português da Biblioteca Vaticana (Cód. 4803): Reprodução fac-similada com introdução de L. F. Lindley Cintra. Lisboa: Centro de Estudos Filológicos, Instituto de Alta Cultura, 1973.

CANCIONEIRO da Biblioteca Nacional (Colocci-Brancuti): Cód. 10991. Reprodução fac-similada. Lisboa: Biblioteca Nacional, Imprensa Nacional – Casa da Moeda, 1982.

CANCIONEIRO da Ajuda: Edição fac-similada do códice existente na Biblioteca da Ajuda. Lisboa: Edições Távola Redonda, 1994.

CEDEÑO, R. A. N.; MORALES-FRONT, A. *Fonología generativa contemporánea de la lengua española.* Washington, DC: Georgetown University Press, 1999.

COLLISCHONN, G. A sílaba em português. In: BISOL, L. (Org.). *Introdução a estudos de fonologia do português brasileiro.* 4ª ed. rev. e ampl. Porto Alegre: EDIPUCRS, 2005, p. 101-129.

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

FIORAVANTI, C. Ataliba Teixeira de Castilho: o linguista libertário. *Revista Fapesp*. Edição 259. Setembro/2017. Disponível em: https://revistapesquisa.fapesp.br/ataliba-teixeira-de-castilho-o-linguista-libertario/.

GOLDSMITH, J. A. *Autosegmental Phonology.* Cambridge, MA, 1976. Tese (Doutorado em Linguística) – Cambridge.

GOLDSMITH, J. A. *Autosegmental & metrical phonology.* Oxford: Blackwell, 1990.

HALLE, M. Addendum to Prince's "Metrical Forms". In: KIPARSKY, P.; YOUMANS, G. (Orgs.). *Phonetics and Phonology.* Nova York: Academic Press, 1989, p. 81-86, V. 1: Rhythm and Meter.

HAYES, B. The prosodic hierarchy in meter. In: KIPARSKY, P.; YOUMANS, G. (Orgs.). *Phonetics and Phonology.* Nova York: Academic Press, 1989, p. 201-260, V. 1: Rhythm and Meter.

HAYES, B. *Metrical Stress Theory:* principles and case studies. Chicago: The University Of Chicago Press, 1995.

HOGG, R.; McCULLY, C. B. *Metrical phonology:* a coursebook. Cambridge: Cambridge University Press, 1987.

HORA, D.; BATTISTI, E.; MONARETTO, V. O. (Coord.) *História do Português Brasileiro: Mudança fônica do Português Brasileiro.* São Paulo: Contexto, 2019.

HYMAN, L. M. *Phonology:* theory and analysis. New York: Holt, Rinehart and Winston, 1975.

LEÃO, Â. V. Questões de linguagem nas *Cantigas de Santa Maria*, de Afonso X. *Ensaios:* Associação Internacional de Lusitanistas (AIL). 2002. Disponível em: http://periodicos.pucminas.br/index.php/scripta/article/view/10371/0. Acesso em: 09 jul. 2024.

LEÃO, Â. V. *Cantigas de Santa Maria de Afonso X, o sábio.* Aspectos culturais literários. São Paulo: Linear B; Belo Horizonte: Veredas & Cenários, 2007.

LOPES, G. V.; FERREIRA, M. P. et al. (2011-), *Cantigas Medievais Galego Portuguesas [base de dados online]*. Lisboa: Instituto de Estudos Medievais, FCSH/NOVA. Disponível em: http://cantigas.fcsh.unl.pt. Acesso em: 10 jul. 2024.

MASSINI-CAGLIARI, G. Escrita do *Cancioneiro da Biblioteca Nacional de Lisboa*: fonética ou ortográfica? In: *Filologia e Linguística Portuguesa*, São Paulo, n. 2, 1998, p. 159-178.

MASSINI-CAGLIARI, G. *Do poético ao lingüístico no ritmo dos trovadores:* três momentos da história do acento. Araraquara: FCL, Laboratório Editorial, UNESP; São Paulo: Cultura Acadêmica, 1999.

MASSINI-CAGLIARI, G. *Cancioneiros Medievais Galego-Portugueses.* Fontes, edições e estrutura. São Paulo: WMF Martins Fontes, 2007.

MASSINI-CAGLIARI, G. Inovação Científica em Estudos Medievais: Descobrindo os sons do Português Arcaico. *Revista da Anpoll*. Vol. 1, nº 34, p. 17-50, Florianópolis, Jan./Jun. 2013.

MASSINI-CAGLIARI, G. *A música da fala dos trovadores:* desvendando a prosódia medieval. 1ª ed. São Paulo: Editora UNESP Digital, 2015.

MASSINI-CAGLIARI, G. Acentos em nomes. In: da HORA, D.; BATTISTI, E.; MONARETTO, V. O. (Coord.) *História do Português Brasileiro:* Mudança fônica do Português Brasileiro. São Paulo: Contexto, 2019. p. 198-225.

MATTOS E SILVA, R. V. *Estruturas trecentistas:* elementos para uma gramática do Português Arcaico. Lisboa: Imprensa Nacional – Casa da Moeda, 1989.

MATTOS E SILVA, R. V. *O português arcaico:* fonologia, morfologia e sintaxe. São Paulo: Contexto, 2006.

METTMANN, W. Glossário. In: AFONSO X, O SÁBIO. *Cantigas de Santa Maria*. Coimbra: Universidade, 1972. v.IV: Glossário.

METTMANN, W. (Org.). *Cantigas de Santa María (cantigas 1 a 100):* Alfonso X, el Sabio. Madrid: Castalia, 1986.

MONGELLI, L. M. *Fremosos cantares:* Antologia da lírica medieval galego-portuguesa. São Paulo: Editora WMF Martins Fontes, 2009.

PARKINSON, S. As *Cantigas de Santa Maria*: estado das cuestións textuais. *Anuario de estudios literarios galegos*, Vigo, p. 179-205, 1998.

PERLMUTTER, D. Phonological Quantity and Multiple Association. In: GOLDSMITH, J. A. (ed.). *The Handbook of Phonological Theory.* Oxford, UK: Blackwell, 1995, p. 307-317.

PRINCE, A. S. Metrical Forms. In: KIPARSKY, P.; YOUMANS, G. (Orgs.). *Phonetics and Phonology.* Nova York: Academic Press, 1989, p. 45-80, V. 1: Rhythm and Meter.

WETZELS , E. The Syllable. In: HULST, H.; SMITH, N. (Org.). *The structure of phonological representations (part. II).* Dordrecht: Foris, 1982, p. 337-383.

WETZELS, W. L. Consoantes palatais como geminadas fonológicas no português brasileiro. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 9, n. 2, 2000, p. 5-15, jul./dez.

ZUCARELLI, F. E. *Ditongos e hiatos nas cantigas medievais galego-portuguesas.* Araraquara, 2002. Dissertação (Mestrado em Linguística e Língua Portuguesa) – FCL/UNESP.

http://doi.org/10.5007/1984-8420.2025.e101224

DO PORTUGUÊS CLÁSSICO AO PORTUGUÊS BRASILEIRO: HERANÇA, RAÍZES, DERIVA E CONTATO[1]

FROM CLASSICAL PORTUGUESE TO BRAZILIAN PORTUGUESE: INHERITANCE, ROOTS, DRIFT AND CONTACT

Charlotte Galves | Lattes | charlottemgc@gmail.com
Universidade Estadual de Campinas/CNPq

Resumo: Baseado em dois corpora sintaticamente anotados cobrindo o período do século 16 ao século 19, este artigo argumenta que o contato com outras línguas, em particular as africanas, não precipitou a "deriva lusitana e românica" do português no Brasil, como afirmado por Silva Neto (1977 [1950]) e Naro e Scherre (2007). Ao contrário o estudo da evolução do uso dos determinantes na história do português europeu, e sua comparação com outras línguas românicas, mostra que a deriva dessas línguas foi no sentido de uma generalização da presença dos artigos, em direção contrária à tendência observada no português brasileiro de manter a variação no uso e ampliar os contextos de ausência de determinante. Isso coloca em xeque o papel atribuído ao contato, pelos autores mencionados, de coadjuvante da deriva, ao mesmo tempo que reintegra no estudo da evolução do português no Brasil a descrição da língua que foi importada sem descontinuar durante 3 séculos.

Palavras-Chave: Deriva linguística. Contato linguístico. Determinantes. Português brasileiro. Português clássico.

Abstract: Based on two syntactically annotated corpora covering the period from the 16th to the 19th century, this article argues that the contact with other languages, particularly African ones, did not precipitate the "Lusitanian and Romance drift" of Portuguese in Brazil, as stated by Silva Neto (1977 [1950]) and Naro and Scherre (2007). On the contrary, the study of the evolution of the use of determiners in the history of European Portuguese, and its comparison with other Romance languages, shows that the drift of these languages was towards a generalization of the presence of articles, in the opposite direction to the trend observed in Brazilian Portuguese to maintain variation in use and to expand the contexts of absence of determiners. This calls into question the role attributed to contact, by the mentioned authors, as supporting the drift, at the same time that it reintegrates into the study of the evolution of Portuguese in Brazil the description of the language that was imported without discontinuing for 3 centuries.

Keywords: Linguistic drift. Linguistic contact. Determiners. Brazilian Portuguese. Classical Portuguese.

[1] Este trabalho foi parcialmente financiado pelo CNPq, Processo 313146/2023-3.

1 Introdução

Muito se tem escrito sobre as raízes africanas do português brasileiro (PB no que segue). Eu me incluo entre os estudiosos do PB que argumentam a favor de uma forte marca do contato com línguas africanas em território brasileiro, em particular em trabalhos conjuntos com Juanito Avelar (cf. entre outros, Avelar e Galves, 2014). Também critiquei a postura de Anthony Naro e Marta Scherre quando afirmavam: "Ainda não conseguimos identificar nenhuma característica do português do Brasil que não tenha um ancestral claro em Portugal". (Naro e Scherre, 2007, p. 68). Porém, neste capítulo, focalizarei um fenômeno presente nos textos da época em que o português veio ao Brasil e pode ser considerado um "ancestral" de uma característica do português do Brasil. Trata-se da variação na presença dos determinantes com nomes possessivos, com nomes próprios e com nomes comuns em certos contextos. Além disso, uma vez que o português europeu (doravante PE) evoluiu no sentido oposto, criticarei a ideia presente em Silva Neto (1977 [1950]) e também assumida por Naro e Scherre (2007, p. 47) de que no Brasil a "deriva secular" que atua "ao longo de um vetor de desenvolvimento" "se encontrou com outras forças que reforçavam e expandiam a direção original", essas forças sendo o contato com as línguas africanas em particular. Mostrarei que o contato, no que diz respeito a esse fenômeno, teve como efeito reforçar o uso de sintagmas nominais sem determinante da gramática portuguesa original, indo contra a deriva atuando na língua em Portugal e em outras línguas românicas, onde o determinante passou a ser categórico na maior parte dos contextos. Para isso, trarei dados de dois corpora históricos do português, o *Corpus Tycho Brahe* (doravante CTB) e o *Corpus Post-Scriptum* (doravante CPS). Ambos cobrem o período em que vagas sucessivas de portugueses vieram ao Brasil, ou seja do século 16 ao século 19, porém com uma composição bem diferente. Enquanto o primeiro é composto de textos de diversos tipos (cartas, sermões, peças de teatro, biografias, historiografia, jornais) escritos por letrados, o segundo reúne cartas representativas da escrita cotidiana ("ordinary writing"). Se a competência na língua escrita varia muito nos escreventes desse corpus, desde "mãos inábeis"[2] até pessoas com alto nível de letramento, nenhum deles tem como profissão escrever, e os documentos por eles escritos são cartas pessoais do dia a dia. Essa diferença torna a comparação dos dois corpora muito preciosa, porque nos permite avaliar a representatividade de textos mais próximos da norma escrita.

O capítulo está organizado da seguinte maneira. Na próxima seção, discuto a periodização do português europeu, argumentando que, de um ponto de vista da gramática subjacente, foi o português clássico (doravante PCl) que aportou ao Brasil. A Secão 3

[2] Cf. Marquilhas (2001).

expõe as propriedades do sintagma nominal do PCl no que diz respeito à presença ou não do determinante. A Seção 4 mostra que a evolução do PE foi no sentido do aumento do uso de mais determinantes, indo numa deriva que se encontra em outras línguas românicas, enquanto no PB a evolução foi no sentido da presença de menos artigos, tanto definidos quanto indefinidos. A conclusão volta à questão da relação da deriva com o contato propondo uma articulação dos dois conceitos em termos emprestados tanto à teoria mentalista da gramática gerativa quanto a abordagens sócio-históricas.

2 A periodização gramatical do português europeu: limites e nomeação

A pergunta das gramáticas que vieram ao Brasil só pode ser respondida a partir de uma visão precisa da periodização gramatical do português europeu. O quadro a seguir é baseado em Mattos e Silva (1992), acrescido da proposta de Teyssier (1982)[3] e de Galves (2012) que retoma a abordagem de Galves, Namiuti e Paixão de Sousa (2006), em termos de períodos definidos por gramáticas, no sentido da Gramática Gerativa, ou seja, como competência dos falantes adquirida no processo natural de aquisição de língua materna.

<table>
<tr><th>Época</th><th>Leite de Vasconcelos</th><th>Silva Neto</th><th>Pilar V.Cuesta</th><th>Lindley Cintra</th><th>Teyssier (1982)</th><th>Galves (2012)</th></tr>
<tr><td>Primeiros documentos até 1350-1420</td><td rowspan="2">arcaico</td><td>trovadoresco</td><td>galego-português</td><td>português antigo</td><td>galego-português</td><td>galego-português</td></tr>
<tr><td>até 1536-1550</td><td>português comum</td><td>português pré clássico</td><td>português médio</td><td rowspan="3">português europeu</td><td rowspan="2">português hispânico</td></tr>
<tr><td>Até séc. 18</td><td rowspan="2">moderno</td><td rowspan="2">português moderno</td><td>português clássico</td><td>português clássico</td></tr>
<tr><td>Até hoje</td><td>português moderno</td><td>português moderno</td><td>português europeu moderno</td></tr>
</table>

O que se vê nesse quadro são diferenças entre os estudiosos que dizem tanto respeito à divisão em períodos (2, 3 ou 4), à localização das fronteiras temporais entre os períodos, e finalmente à sua nomeação. Por exemplo a fronteira na virada do século 14 para o século 15 é assumida por todos menos Leite de Vasconcelos, que estabelece um

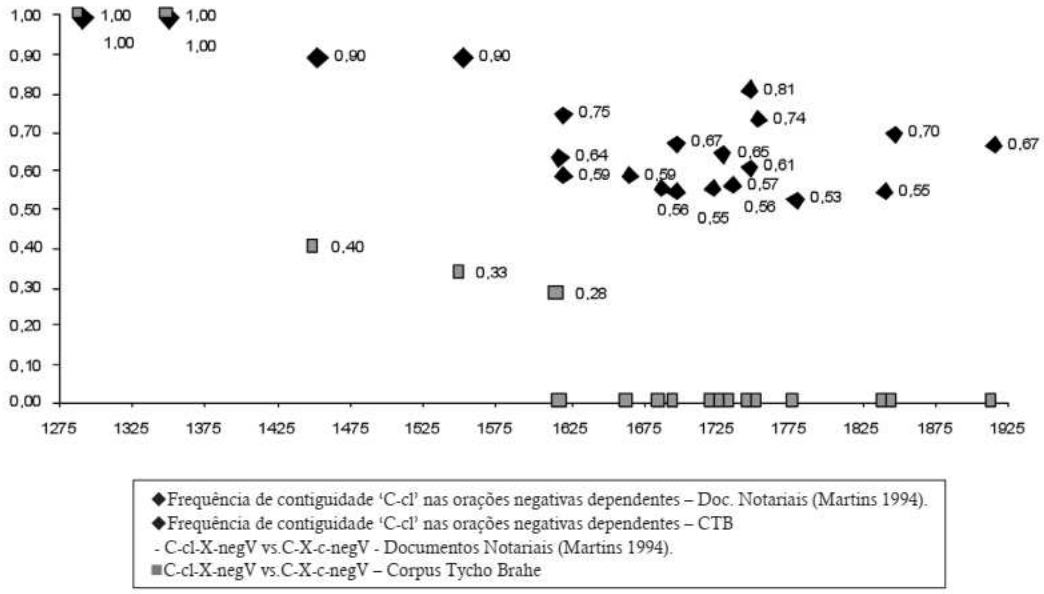
[3] Essa inclusão foi feita por Spatti (2024). No artigo de Mattos e Silva (1992) os autores mencionados no quadro não são datados.

só limite entre dois grandes períodos, o arcaico e o moderno, no século 16. Isso é também assumido pelos três autores seguintes e é certamente a periodização mais tradicional, mas com uma divisão em duas fases distintas, para as quais há uma enorme variação de nomeação (português arcaico, trovadoresco ou galego-português, para a primeira e português comum, pré-clássico e médio para a segunda). Já Teyssier inova ao considerar que o português moderno começa no fim do galego-português. Enfim, argumentei, com base na análise de Galves, Namiuti e Paixão de Sousa (2006), que do ponto de vista estritamente gramatical, o português médio e o português clássico formam uma unidade, e que o primeiro traz, nos textos, as marcas da competição com o galego-português, o que aliás combina bem com a nomeação de "pré-clássico" proposta por Vasquez-Cuesta no quadro acima. O nome de "Português hispânico" para essa unidade gramatical faz referência à noção de Hispânia, nome dado à península ibérica no século 16, onde o português e o castelhano coexistiam de maneira bastante interativa, com influências mútuas (cf. Martins, 2003, Paixão de Sousa, 2004).

Essa fase cobre os séculos 16, 17 e para alguns autores o séc. 18.[4] Portanto, mesmo considerando um limite mais tardio para o final do galego-português, não há dúvida de que o português que aportou inicialmente ao Brasil era o português clássico, trazendo mais ou menos marcas de fases arcaicas conforme a proveniência geográfica.

Desse ponto de vista, a questão de saber se os primeiros portugueses chegados ao Brasil falavam um português quatrocentista (cf. Castilho, 2001) ou quinhentista perde da sua acuidade, uma vez que seriam duas faces da mesma moeda gramatical. É verdade, porém, que há diferenças entre essas duas vertentes do português, que afetam em primeiro lugar a morfologia da língua e a perda de certos itens lexicais como os pronomes clíticos genitivo em/ende e locativos i. Mas outros aspectos arcaicos como o uso do auxiliar *ser* para o passado dos verbos inacusativos ainda se encontram no PCl (cf. Galves *et al.*, em andamento).. Se a competição de formas encontrada em textos quatrocentistas também se encontrava nos falantes do início do séc. 16, talvez por razões de diversidades dialetais, a hipótese seguida aqui é que a gramática deles era a gramática do português clássico, aliás não tão diferente do português antigo em termos paramétricos (cf. Medeiros, 2018), a não ser talvez para a sintaxe de colocação dos clíticos. Mas o gráfico a seguir, de Namiuti (2008) mostra que já no século 15 o fenômeno da interpolação por um elemento que não seja a negação (quadradinhos cinzentos) se torna muito raro nos textos, e por hipótese ainda mais na fala.

[4] Os trabalhos baseados no CTB mostraram que a mudança para o PE se manifesta na primeira geração dos autores do século 18 (cf. entre outros Galves e Paixão de Sousa, 2017).



Fonte: Namiuti (2008)

Talvez um dos aspectos importantes da diferença entre o PA e o PCl seja mais textual do que gramatical, no sentido de que a escrita do primeiro seria mais permeável à língua oral do que a do segundo, em particular a partir do processo de gramaticalização ligado à invenção da tipografia. Isso, de uma certa maneira, aproxima o primeiro período do português em Portugal da língua falada no Brasil.

O centro da discussão de Célia Castilho quando indaga se "seria quatrocentista o português do Brasil" (Castilho, 2001) são os fenômenos de redobramento que não deixaram de ser frequentes no português clássico, onde se acham dois processos, o redobramento clítico (1)[5] e a deslocação à esquerda (2):

(1) a. Eu não posso mandar pedir mercadorias nem contrato sem <u>mas pedir a mim</u> El-Rei de Marrocos, quando houver por trabalho mandar a Portugal a isso uma pessoa de confiança. (PSCR0014,.16) 16c.

b. E assim Nosso Senhor <u>lhe dará a VM</u>, na outra vida, o prémio da glória e, nesta, a saúde que deseja para lhe fazer muitos serviços. (CARDS1063,.2) 17 c.

c. Mas o que eu <u>lhe lembro a você</u> que Francisco Teles de Pina me levará caminho de Évora porque tenho de que me acusar e será por toda essa semana, se Deus quiser. (CARDS4008,.2)

[5] Em (1)a, o redobro é com o pronome clítico me, contraído com o pronome as na forma mas.

d. E eu peço a VM, da minha parte, como amigo que eu sou de VM, que lhe responda a ela. (PSCR1252,.3)

e. Fico entregue das duas patacas, de que dou a VM o agradecimento, e fico muito certa para a VM lhe obedecer, pedindo à senhora do Cardal lhe dê a VM bom sucesso. (CARDS1108,.1)

(2) a. E por isso muito acomodadamente a eles lhes disse o Senhor, que dentro do mesmo século se havia de cumprir (Vieira, 1608, Sermões, 69.126)

b. O freio, vos encomendo muito que mo mandes. (PSCR0014,.21) sec. 16

c. Com o padre Francisco Rodrigues de Abreu, morador em Ribafeita, coadjutor que foi nesta vila de São Pedro do Sul, com ele tive uma amizade onde ele imaginou me persuadisse, o que não foi. (CARDS4016,.1) séc. 17

Note-se que esses processos de reduplicação não se restringem às cartas populares, mas se encontram também em correspondências e outros gêneros de textos, por autores eruditos e literatos, como o Padre Antonio Vieira. Porém, nesses, são mais frequentes as construções em que o elemento redobrado precede o clítico, como em (2)a e b. do que naquelas em que o segue, como em (1). (2)c. mostra por outro lado um caso de redobramento não clítico, com preposição.

A presença de pronomes lembrete também se dá em orações relativas, um fenômeno muito pouco frequente em textos eruditos, onde se restringem a casos de ilhas sintáticas[6] e mais comum nas cartas do CPS, sem restrição.

(3) a. Porque, emfim, não posso negar que sou um daqueles, que, quando o demo os toma uma vez, sempre lhes fica um jeito. (Melo, 1608, 31.286)

b. E eu assim me ia também e, se me disse que VM havia mandado dar cá a Gaspar Pereira dez mil reais, os quais eu lhos dei, e VM nos mande a mim. (PSCR0018,.13) séc. 16

Na próxima seção, apresento um outro fenômeno do PCl que se encontra no PB, a variação no uso dos artigos definidos e indefinidos no português clássico que, de maneira mais geral coloca no centro da discussão a constituição do sintagma nominal nas diversas fases da história do português.

[6] Ou seja, em domínios nos quais o movimento para fora das fronteiras de frase é impossível, como nas orações relativas. Nesse caso, a presença de um pronome pode ser considerada uma estratégia de "último recurso".

3 O sintagma nominal no português clássico

As frases seguintes mostram que os autores do período clássico usam variavelmente o artigo definido com sintagmas nominais possessivos, até numa mesma frase.[7]

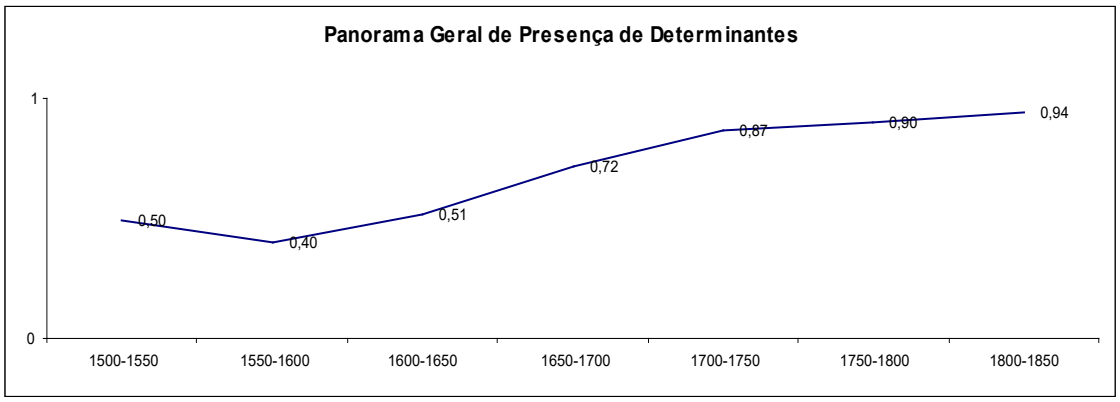
(4) Em várias partes das fronteiras fizeram os castelhanos fumo a fim de que o vento o deitasse para a banda de Portugal, e fizesse fugir nosso gado para as suas defesas: (Galhegos, 1597, 38.542)

(5) por isso, disse galantemente Tertuliano: "Teme um homem matar a sua vaca, porque acaso não coma alguma posta de sua avó". (Bernardes, 1644,13.150)

(6) Lastime- -se a terra dos seus marcos e dos seus muros para se distinguirem as aquisições de seus usurpadores. (Brochado, 1651, 145.1354)

Entre o século 16 e o século 19, o uso do artigo cresce sensivelmente, chegando a um uso quase obrigatório no PE contemporâneo (cf. Raposo e Miguel, 2013, p. 729). A Figura I, de Floripi (2008) mostra essa evolução por períodos de 50 anos.

Figura 1: Frequência de ocorrências do determinante nos DPs possessivos do século 16 a 19, por períodos de 50 anos.



Fonte: Floripi (2008, p.138)

Note-se que a dinâmica de mudança já se dá entre o século 16 e 17. Isso se pode observar também na seguinte tabela, adaptada de Galves *et al.* (em andamento).[8]

[7] Os números aparecendo ao lado do nome dos autores é 1. A sua data de nascimento 2. A página da frase no texto utilizado no corpus 3. o número da frase no texto.

[8] Não passará desapercebido ao leitor atento que as frequências relativas da presença do determinante são mais altas no gráfico de Floripi (2008) do que na tabela de Galves *et al.* (em andamento). Essa discrepância se dá em grande parte pelo fato da primeira ter retirado do conjunto dos dados contabilizados as expressões cristalizadas como "de minha parte", o que não foi feito por Galves *et al*. Outra fonte de diferença pode vir da composição próxima porém diferente dos respectivos corpora, dado que se observa uma variação entre autores, que pode afetar os resultados a nível dos períodos.

Tabela 1: NPs possessivos sem e com determinantes definitos

	16	17	PCl
Sem D	1837 (69%)	1857 (48%)	3694
Com D	817 (31%)	2049 (52%)	2866
Total Poss N	2654	3906	6560

Fonte: Galves *et al.* (em andamento)

A Tabela 1 mostra que a frequência dos artigos definidos antes de possessivo passa de um terço à metade dos dados do século 16 para o 17. Contudo, nos dois séculos, é possível observar uma regularidade na variação dos determinantes. Galves *et al.* (em andamento), na sequência de outros autores (cf. Rinke, 2010), mostram que a presença do artigo definido está relacionada com a especificidade e topicalidade do sintagma nominal. Esses traços estão frequentemente associados, entre outros, à posição sujeito como se vê na tabela a seguir:

Tabela 2. O uso de artigos definidos em NPs possessivos sujeitos

	16	17
Sujeito sem D	192 (54%)	221 (36%)
Sujeito com D	162 (46%)	422 (64%)
Total de Poss N	354	643

Fonte: Galves *et al.* (em andamento)

Comparando as duas tabelas, verificamos que a frequência de determinante com NP sujeito aumenta substancialmente em posição sujeito. 46% contra 31% no século 16, e 64% contra 52% no século 17. Uma evidência suplementar encontra-se na primeira pessoa, a mais alta na escala de definitude. Em 221 ocorrências na primeira pessoa no século 16 e 488 no século 17, 106 e 367, respectivamente, têm artigos definidos. Isso corresponde a 48% para o século 16 e 75% para o século 17.

Quanto à ausência de artigos, Floripi (2008) e Macedo-Costa (2016) observam que se encontra essencialmente com nomes expressando relações de família e outras relações de vários tipos (amigos, servidores, superiores ou inferiores hierárquicos). Macedo-Costa (2016, p. 183) mostra que, até o fim do século 17, somente em posição sujeito esse tipo de nome tem artigos numa frequência superior a 10%. Os outros nomes sem artigo se encontram essencialmente com antecedente [-humano] bem como em expressões cristalizadas como em (8)

(7) A primeira e mais antiga se chama Tamaracá, a qual tomou este nome de uma ilha pequena, onde sua povoação está situada. (Gândavo, 1540, 11.78)
(8) Se morrem ou vivem é só por sua conta, (Brochado, 1651, 151.1405)

Existe também uma interessante variação entre tipos textuais. A diferença entre a correspondência e os sermões do Padre Vieira nos dá um exemplo do papel do discurso e da estrutura informacional no uso dos determinantes. Nos sermões encontram-se artigos definidos em 62% das ocorrências de nomes possessivos contra 50% na correspondência. Isso é em parte devido à alta frequência da primeira pessoa do plural, usada com poucas exceções sem artigo. Além disso, nomes que fazem referência a noções e argumentos importantes nos Sermões como *dia do juizo*, que aparece 141 vezes no texto, tem artigo definido em todas as ocorrências de NPs possessivos (10/10).

(9) Os que morrem hoje, teem o seu dia do Juiso hoje; (Vieira-Sermões: 1608, 69.118)

Podemos inferir desses dados que, nos séculos 16 e 17, o artigo ainda tem um valor discursivo que herdou do demonstrativo que está na sua origem, expressando especificidade, unicidade e, em última análise, ênfase. O uso variável do artigo definido com nomes próprios apresenta características semelhantes. Em (10), vê-se *Dom João de Alvarado* introduzido primeiro sem determinante, e na segunda ocorrência aparecendo com o artigo definido.

(10) (...) determinou Dom António de Mendonça, Viso-Rei da Nova Hespanha, mandar a elas uma Armada por sua conta, de que elegeo por Capitão Dom João de Alvarado, mas depois das despezas feitas faleceo o Dom João de Alvarado de uma queda que deo de um cavalo (Couto: 1542, 07.71)

Os nomes próprios podem ser precedidos também por artigos indefinidos, possivelmente quando não há pressuposição de conhecimento por parte dos leitores. Vê-se nas duas frases seguintes, um mesmo personagem introduzido com artigo indefinido e retomado em seguida com artigo definido.[9]

[9] Note-se que o primeiro nome de *de la Torre* é diferente na segunda ocorrência. A mesma diferença se acha na edição de 1612 das *Décadas*, disponível na Biblioteca digital de Lisboa. O contexto das frases não deixa dúvida quanto ao fato de que as duas menções correspondem à mesma pessoa.

(11) a. mandou o General um Bernardo de la Torre por Capitão de um galeãozinho, chamado São Joanilho, e a Pero Ortiz de la Rueda na galeota, (Couto: 1542, 09.110)
b. e o Fernão de la Torre se ficou negociando, e provendo pera sua jornada, (Couto: 1542,10.116)

O mesmo padrão e tendência se encontram nas cartas do CPS (Galves *et al.,* em andamento). Contudo, a frequência de artigos definidos é bastante menor nesse corpus (15% no século 16 e 22% no século 17). Galves *et al.* notam que isso pode ser explicado pela recorrência de nomes de parentesco. Contudo o artigo definido com esses nomes também se encontra em posições destacadas, como o sujeito e o tópico, criando contrastes como exemplificado em (12a.).

(12) a. A minha sobrinha, filha de minha irmã, a mais moça, tem duas filhas e um filho. (PSCR0049,.16) 16c.
b. E da minha filha não me dizeis nada? (CARDS1005,.17) 17c.

Finalmente, encontramos nos textos da época clássica nomes "nus" tanto plurais quanto singulares.

(13) E pobres em nenhuma mesa faltarão, se houver quem os queira por hóspedes. (S_001_PSD,62.696)
(14) a.Recebi carta de Lisboa, de 29 de Setembro, do padre procurador do Brasil, em que me diz chegara caravela da Baía, (Vieira-Cartas, 1608,0.203)
b. DEPOIS que esta província Santa Cruz se começou de povoar de Portugueses,
sempre esteve instituída em uma governança, na qual assistia governador geral por el-Rei nosso senhor com alçada sobre os outros capitães que residem em cada capitania. (Gândavo, 1540, 14.142)
c. Feito isto, poz ali Capitão novo, (Couto: 1542, 115.1122)
d. Agora que são as 3 horas da tarde vi uma carta de holandês de Amsterdam, interessado na Companhia, escrita a outro holandês desta côrte, na qual diz as palavras seguintes. (Vieira-Cartas: 1608,0.1269)

e. Como da terra se avistou embarcação do Reino, foi grande o alvoroço nos Padres, que tinham chegado adiante. (Barros, 1675, 78.620)

Comparando com o PE moderno (cf. Peres, 2013), constatamos que os nomes nus em PCl são licenciados de maneira mais geral e com interpretações mais diversificadas. Os nomes nus plurais podem aparecer em posição pré-verbal (13) e os nomes nus singulares podem receber interpretações específicas, como mostram os modificadores *de Lisboa*, *de Amsterdam*, *do Reino*. O mesmo se verifica no CPS.

(15) a. Contanto que é ainda, está em Londres até vir recado de Portugal e logo é aqui.
(PSCR0002,.4) séc. 16
b. E não seja carta tão curta como a que me cá mandais, que carta para tão longe
terra não há de ser tão curta. (PSCR0002,.15) séc. 16
c. Pois dizer-vos o que faríamos vendo vossas cartas é escusado, porque em vir
carta nos parecerá vermos a todos e algo nos satisfaremos. (PSCR0005,.7) séc. 16
d. E quando, com o favor de Deus, o possa fazer, não posso subir eu degrau que
não seja em braços de homens. (CARDS0052,7) séc. 16
e. Por este portador me mande VM um rosário de contas, que o não mando pedir por
ter falta dele, mas quero ver cousa da mão de VM, que, o vendo, farei
de conta vejo a VM para o encomendar a Deus, que lhe dê muitos largos anos de
vida para meu amparo. (CARDS1001,.13) séc. 17
f. Por se oferecer portador destas partes, com razão se queixaria VM se em
mim houvesse descuido em não solicitar nesta ocasião suas notícias.
(CARDS1111,.1) séc. 17
g. Dela tem filho e já teve outro que lhe morreu. (CARDS2123,.8) séc. 1710

[10] Aqui o sentido de "um filho" é reforçado por "um outro".

Em todos os casos de (14) e (15), o PE privilegia hoje a presença de um artigo indefinido, e até, no caso de (14)b. de um artigo definido.

A tabela a seguir (Schmitt e Galves, 2016) mostra a evolução da frequência de nomes nus singulares e plurais do século 16 a 19 no português europeu. Em negrito estão os valores para os quais se vê uma mudança. São os singulares nus que passam de 10.83 a 6.54, num movimento oposto aos indefinidos singulares que passam de 7.00 a 11.46, e os definidos plurais que passam de 9.75 a 13.92. Em geral, se fazemos a proporção de nomes nus em relação aos nomes com determinantes, passamos de 18,29% no século 16 a 13,22% no século 19.

Tabela 3: A evolução dos tipos de SNs do século 16 a 19

Object	1500	1600	1700	1800
Bare Sg	**10.83**	**8.91**	**6.87**	**6.54**
Bare Pl	7.46	7.94	6.94	6.68
Ind Sg	**7.00**	**7.28**	**9.90**	**11.46**
Ind Pl	1.33	0.25	0.31	0.56
Def Sg	22.38	28.67	36.41	20.08
Def-pl	**9.75**	**15.07**	**13.79**	**13.92**
Others	52.08	40.79	32.65	47.3
Total	100	100	100	100

Fonte: Schmitt e Galves (2016, hand-out)

Apesar de não haver ainda estudos sistemáticos desse tipo em relação à história do português no Brasil, sabemos que hoje em dia o PB tem uma alta frequência de nomes nus (cf. entre outros Cyrino e Espinal, 2015; Pires de Oliveira e Rothstein, 2011; Muller e Oliveira, 2004; Schmitt e Munn, 2003; Wall, 2018). Um exemplo interessante da diferença entre PB e PE pode ser achado em traduções de textos nas duas vertentes. Aqui o *Cálice de Fogo*, da série Harry Potter:

(16)

a. Os feiticeiros, que supostamente estão mortos, fariam melhor em não dar de caras com as feiticeiras do Ministério da Magia numa estalagem. (PE)

b. Seria melhor que bruxos presumivelmente mortos não esbarrassem em bruxas do Ministério da Magia em hotéis à beira de estradas... (PB)

c. Porque ... os feiticeiros estão a chegar ao país... (PE)
d.Porque ... estão chegando ao país bruxos do mundo inteiro ... (PB)

Por outro lado, em vários dialetos do PB, os nomes nus singulares recebem uma interpretação definida (cf. Lopes e Pagotto, 2023, Wall, 2018), o que não se encontra nos corpora históricos. Esta situação aponta para uma evolução oposta do PB e do PE. A partir de um tronco comum onde nomes nus, singulares e plurais, recebem uma interpretação indefinida, o primeiro evoluiu ampliando os usos, enquanto o segundo teve uma dinâmica contrária, indo para um uso mais sistemático do artigo indefinido. No que tange ao PB, foi sugerido por vários estudos, que o uso mais extenso de nomes nus estava relacionado com o efeito do contato maciço do português com as outras línguas faladas no Brasil durante o período colonial, em particular – mas possivelmente não só (cf. Lopes e Pagotto, 2023) – as línguas africanas. Com efeito, encontra-se o mesmo fenômeno em textos históricos escritos em português na África e no Brasil por africanos, bem como no português africano contemporâneo (cf. Inverno, 2011, a respeito do português angolano).

Na próxima seção, mostraremos que esse efeito do contato no PB não contrasta somente com a evolução do PE, mas também com uma tendência muito geral das línguas românicas, que parece ser uma deriva comum a essas línguas.

4 A deriva românica do artigo

É sabido que os artigos das línguas românicas emergiram a partir dos demonstrativos do latim clássico, que não tinha artigos. Simonenko e Carlier (2020) comentam na introdução do seu estudo da evolução do uso do artigo definido no francês antigo que se trata de uma das instâncias mais robustamente atestadas de gramaticalização. Em outro artigo (Simonenko e Carlier, 2022), as mesmas autoras argumentam que se encontra na história das línguas românicas o "crescimento pan-românico de uma gramática DP" ("pan-Romance rise of a DP grammar"). Entendem por uma gramática DP, um sistema no qual a (in)definitude emerge das propriedades do próprio sintagma nominal e não de propriedades da oração, ligadas à ordem livre dos constituintes. Nesse quadro, e durante um longo período da história, os l-elementos (derivados do latim ill-) compartilham propriedades de demonstrativos e de artigos, e os nomes nus são frequentes. Isso foi amplamente documentado para o francês antigo (séc. 9 a 14) além de ter sido estudado também para o espanhol antigo.[11] Rinke (2010, p.131) também afirma que, seguindo a

[11] Para o espanhol, Simonenko e Carlier (2020) mencionam Roca (2009).

evolução diacrônica das línguas românicas, os artigos definidos com possessivos são primeiro marcadores discursivos que referem anaforicamente a um referente já mencionado, sendo depois gramaticalizados. A descrição trazida na seção anterior mostra que esse padrão se manteve em português durante um tempo mais longo do que nas suas línguas coirmãs, durando ainda nos séculos 16 e 17. Vimos que o artigo definido ainda tem um claro papel discursivo no caso dos sintagmas nominais possessivos e dos nomes próprios, onde sua ausência é o não marcado. Observamos também a frequência de nomes nus onde o padrão europeu moderno requer artigos indefinidos, ou até definidos no caso de indefinidos plurais, como se vê em (16) a,c (cf. Peres, 2013). Para Galves *et al.* (em andamento), esses fenômenos são gerados por uma gramática na qual há duas posições distintas de determinantes, paralelamente à existência de duas posições distintas de complementador na oração.[12] Desse ponto de vista, a diferença entre o século 16 e 17 não seria devida a uma mudança gramatical, uma vez que os autores argumentam que os dois séculos têm a mesma gramática, mas à tendência crescente em usar o l-elemento (que, em português, perdeu o l...) como uma simples marca de definitude.

Essa é a deriva românica e portuguesa. O caminho do PB vai no sentido estritamente oposto, não no sentido de uma desgramaticalização, mas no sentido de reduzir o uso do determinante em lugar de ampliá-lo. No dialeto padrão, o limite dessa redução são os referentes específicos definidos. Em outros dialetos, a redução atinge até sintagmas nominais definidos, como nos exemplos apresentados por Wall (2018, p. 236, 242):

(17) a. Ela tem uma filha [...] filha dela tem 7 anos (PEUL:RJ)
b. Ela arrumou um rapaz. Rapaz gosta muito dela (PEUL:RJ)
c. saxofonista tenta ver o que entrou no seu saxofone (corpus elicitado)

Não faz parte dos objetivos deste capítulo propor uma análise para esses dados, mas o que chama atenção nos exemplos acima a. e b. é o caráter anafórico dos nomes nus "filha" e "rapaz", que aponta para uma diferença fundamental entre o PE e as línguas românicas por um lado e o PB por outro. Não pode ser a mesma "deriva" que leva aos exemplos em (17) e aos exemplos em (16) a,c.

Na conclusão retomo a questão da relação entre deriva e contato.

[12] A mesma estrutura é proposta por Bernstein *et al.* (2022), que apresentam como argumento empírico línguas como o catalão do Baleares e o picardo nas quais coexistem dois artigos definidos, um com força dêitica e outro usado com entidades intrinsecamente referenciais como os países e os nomes próprios. Note-se que no português clássico, esses são prototipicamente desprovidos de determinante (cf. Macedo Costa, 2016, e Galves *et al.*, em andamento).
Desse ponto de vista, a coexistência de dois padrões não corresponde a uma competição de gramáticas, mas a uma competição entre dois possíveis valores atribuídos a um item lexical em curso de gramaticalização.

5 Conclusão

Na introdução do seu livro, Naro e Scherre (2007, p. 17) afirmam:

> O conteúdo deste livro tem uma linha mestra clara: apresentar evidências de que características morfossintáticas e fonológicas do português brasileiro, atualmente envoltas em estigma e preconceito social, são heranças românicas e portuguesas arcaicas e clássicas, e não modificações mais recentes advindas das línguas africanas, que vieram para o Brasil com seus povos escravizados e subjugados, ou das línguas dos povos ameríndios, que aqui já se encontravam quando vieram os colonizadores europeus.... queremos identificar as raízes lingüísticas românicas e lusitanas que insistem em permanecer em nossas bocas e em nossas falas.

As raízes certamente existem. E identificamos um grupo delas neste capítulo. O português que chegou ao Brasil, da mesma maneira que continha muitas construções de redobro, continha sintagmas nominais nus (ou reduzidos, nos termos de Peres, 2013), singulares e plurais, bem como sintagmas possessivos sem artigo definido, que ficaram no português brasileiro. Aliás, mais do que ficar, frutificaram. A questão é o que os fez frutificar.

Naro e Scherre (2007, p. 47) afirmam que foi a deriva lusitana e românica, acelerada pelo contato:

> 1) A língua portuguesa falada em Portugal antes da colonização do Brasil já possuía uma deriva secular que a impulsionava ao longo de um vetor de desenvolvimento
> 2) No Brasil, este vetor se encontrou com outras forças que reforçavam e expandiam a direção original
> 3) No início uma dessas forças era a pidginização, que exercia uma influência sobre o português através da língua geral tupi e da "língua de preto" européia, revivificada no Brasil originalmente para uso com os ameríndios.
> 4) Ao longo de toda a história do Brasil, o processo de aprendizado do português como segunda língua teve seus efeitos documentados parcialmente.

Nisso, eles retomam a argumentação de Silva Neto (1977 [1950], p. 115) que coloca a questão nesses termos:

> Na constituição do português brasileiro há desde o século XVI duas *derivas*:
> a) Uma *deriva* bastante conservadora, que se desenvolve portanto muito lentamente e
> b) Uma *deriva* a que condições sociais próprias imprimem velocidade inesperada.

Silva Neto afirma em seguida que o papel dos aloglotas consiste "aqui como alhures, em realizar de imediato o que, em condições normais, levaria talvez séculos para completar-se." (1977 [1950], p. 116). [13] Além de aspectos fonéticos, ele cita o caso muito enfatizado na argumentação de Naro e Scherre, da redução das flexões:

> Na morfologia há um exemplo igualmente expressivo: trata-se da redução das flexões, característica dos falres rurais brasileiros, assim como dos crioulos em geral. Já sabemos que a tendência geral de toda a evolução morfológica das línguas indo-europeias consiste no prevalecimento do carater analítico sobre o sintético: mas no caso dos crioulos isso se verificou num repente, de improviso.

Ora, o que quis mostrar aqui é que o contato nem sempre acelera a deriva portuguesa e românica, como proposto por Serafim da Silva Neto e por Naro e Scherre. Ao contrário, bloqueou-a em muitos casos. É, em particular, o caso discutido aqui da deriva não só portuguesa como também "pan-românica" que diz respeito ao uso crescente do artigo definido.

Esta discussão recoloca a relação da deriva e do contato na dinâmica das línguas cuja história passa pela aquisição de segunda língua por aloglotas, mencionada no item 4) da citação de Naro e Scherre acima. Força é de constatar que contrariamente ao que eles afirmam, seguindo Silva Neto, o contato não precipita forçosamente a deriva. Na realidade, o contato tem um papel decisivo e não só de coadjuvante. No caso dos determinantes, Avelar e Galves (2014) sugerem, com base em dados de escrita de descendentes de africanos no Brasil e de escreventes africanos em regiões de fala bantu (cf. 18 e 19), que a tendência em usar menos artigos do que na língua-alvo se deve a dificuldades de processamento dos mesmos devidas tanto ao próprio português, onde os artigos definidos carecem de consoante inicial, sendo muito facilmente confundidos com vogais iniciais de palavras, quanto às línguas maternas dos aprendizes em que essa categoria não existe.

(18) a. "a este respeito nada disse aSemblea"
b. "depois de lido levou O Senhor Socio Thezoreiro ao Connhecimento da Senblea Os trabalhos..." (Oliveira, 2006, docs 17 e 20 de Antonio José Bracete)

[13] Essa afirmação contrasta com a postura de Naro e Scherre (2007) que procuraram a redução das flexões no português rural do século 20. Para uma discussão dessa análise, cf. Galves (2012) e Lucchesi (2012)

(19) a. “e quando o mesmo Dembo mostre o missão”
b. “ouvidas a legação do Aucthor pello Capitão e Chefe...” (Tavares e Madeira Santos, 2002, p. 140 - 151.)

Além disso, os dados apresentados na Seção 3 mostram que naquilo que os aprendentes do português como segunda língua ouviam, também havia uma forte variação no uso dos artigos e muitos nomes nus, o que deveria conturbar ainda mais o seu aprendizado do português nesse quesito. É na confluência de todos esses fatores que se dá a dinâmica de mudança. A herança lusitana tem a sua parte uma vez que ela proporciona os dados linguísticos primários da aprendizagem. A herança das outras línguas tem também a sua parte ao funcionar como filtro, tornando opacas certas sequências e levando a reinterpretar outras. Obviamente as condições sócio-históricas em que se dá a aprendizagem fazem com que o resultado possa variar de uma região à outra, o que se vê claramente no caso que nos interessa aqui (cf. Callou e Silva, 1997; entre outros). Desconfio que, afinal das contas, a deriva que vem de fora, despida das suas próprias razões históricas, não tenha muito mais a ver com o resultado, a não ser naquilo que ela tem de mais geral, ou seja nos processos universais cíclicos cujo motor são as duas forças antagônicas de economia e de expressividade.

Referências

AVELAR, J.; GALVES, C. O papel das línguas africanas na emergência da gramática do português brasileiro, *Linguística – Revista da ALFAL*, Montevidéu, n. 30, v. 2, p. 239-286, 2014.

BERNSTEIN, J.; ORDÓÑEZ, F; ROCA, F. Emphatic elements and the development of definite articles: evidence for a layered DP in early Romance. *Journal of Historical Syntax*, Volume 5, Article 22, 2021, p. 1-32.

CALLOU, D.; SILVA, M. Uso do artigo definido em contextos específicos. In: HORA, D. (Org.). *Diversidade Lingüística no Brasil*. João Pessoa: Idéia, 1997, p. 11-27.

CASTILHO, C. M. M. Seria quatrocentista o português implantado no Brasil ?. In: MATTOS E SILVA, R. V. (org.). *Para a História do Português Brasileiro*. São Paulo: Humanitas, v. 2, 2001, p. 57-89.

CORPUS TYCHO BRAHE. Corpus Anotado Tycho Brahe do Português Histórico. Unicamp. 2017. https://www.tycho.iel.unicamp.br/corpus/

CORPUS P.S. POST SCRIPTUM. *P.S. Post Scriptum. Arquivo Digital de Escrita Quotidiana em Portugal e Espanha na Época Moderna*. CLUL. 2014. http://ps.clul.ul.pt.

CYRINO, S.; ESPINAL, M. T. Bare Nominals in Brazilian Portuguese: more on the DP/NP analysis. *Natural Language & Linguistic Theory*. 33, p. 471-521, 2015.

FLORIPI, S. *Estudo da variação do determinante em sintagmas nominais possessivos na história do português*. 2008. 254 f. Tese (Doutorado em Linguística), Universidade Estadual de Campinas, 2008.

GALVES, C. Concordância e origens do português brasileiro. In: SEDRINS, A. P; CASTILHO, A. T.; SIBALDO, M. A.; LIMA, R. B (eds.), *Por amor à Linguística:* miscelânea de estudos linguísticos dedicados à Maria Denilda Moura, Maceió, Edufal, 2012, p.123-149.

GALVES, C.; NAMIUTI, C.; PAIXÃO DE SOUSA, M. C. Novas perspectivas para antigas questões: revisitando a periodização da língua portuguesa. In: ENDRUSCHAT, A.; KEMMLER, R.; SCHAFER-PRIEβ, B. (eds.) *Grammatische Strukturen des Europaischen Portugiesisch*, Tubingen: Calepinus Verlag, 2006, p. 45-75.

GALVES, C.; PAIXÃO DE SOUSA, M. C. The position of the verb in the history of Portuguese: Subject position, Clitic placement and Prosody, *Language*, 93, 3, p. 152-180, 2017.

GALVES, C.; DE ANDRADE, A.; NAMIUTI, C.; PAIXÃO DE SOUSA, M. C. *Classical Portuguese, grammar and history*. A ser publicado na coleção Studies in Diachronic and Historical Linguistics, Oxford University Press, em andamento.

INVERNO, L. C. *Contact-induced restructuring of Portuguese morpho-syntax in interior Angola*, 2011. 475 f. Tese (Doutorado em Letras), Universidade de Coimbra, 2011.

LOPES, C. R.; PAGOTTO, E. Nominal constructions in (non-) standard Brazilian Portuguese: a case of contact? But with which languages? Comunicação apresentada no 53º *Linguistic Symposium on Romance Languages*, Paris, junho de 2023.

LUCCHESI, D. A deriva secular na formação do português brasileiro: uma visão crítica. In: LOBO, T.; CARNEIRO, Z.; SOLEDADE, J.; ALMEIDA, A.; RIBEIRO, S. (eds.). *ROSAE: linguística histórica, história das línguas e outras histórias*. Salvador: EDUFBA, 2012, p. 249-274.

MACEDO COSTA T. *Determinantes definidos: um estudo sobre a estrutura dos DPs na história do português*. 2016. 294 f. Tese (Doutorado em Linguística), Universidade Estadual de Campinas, 2016.

MARQUILHAS, R. *A faculdade das letras*. Lisboa: INCM, 2001.

MARTINS, A. M. From unity to diversity in Romance syntax: Portuguese and Spanish. In: BRAUNMULLER, K.; FERRARESI, G. (eds) *Aspects of Multilingualism in European Language History***.** John Benjamins. 2003, p. 201-233.

MEDEIROS, C. *A ordem de palavras no português medieval*. 2018. 288 f. Tese (Doutorado em Linguística), Universidade Estadual de Campinas, 2018.

MULLER, A.; OLIVEIRA, F. Bare Nominals and Number in Brazilian and European Portuguese. *Journal of Portuguese Linguistics*, 3, p. 11-36, 2004.

NAMIUTI, C. *Aspectos da história gramatical do português: interpolação, negação e mudança*. 2008. 309f. Tese (Doutorado em Linguística), Universidade Estadual de Campinas, 2008.

NARO, A.; SCHERRE, M. M. Origens do português brasileiro. São Paulo: Parábola Editorial, 2007.

OLIVEIRA, K. Negros e escrita no Brasil do século XIX: sócio-história, edição filológica de documentos e estudo linguístico. 2006. 1198 f. Tese (Doutorado em Letras). Universidade Federal da Bahia, 2006.

PAIXÃO DE SOUSA, M. C. Língua barroca: Sintaxe e história do português nos 1600. 2004. 450 f. Tese (Doutorado em Linguística), Universidade Estadual de Campinas, 2004.

PERES, J. Semântica do sintagma nominal. In: RAPOSO, E.; BACELAR DO NASCIMENTO, M. F.; COELHO DA MOTA, M.A.; SEGURA, L.; MENDES, A. (orgs.) *Gramática do Português*, vol. 1, Lisboa: Fundação Calouste Gulbenkian, 2013, p. 735-818.

PIRES DE OLIVEIRA, R.; ROTHSTEIN, S. Bare singular noun phrases are mass in Brazilian Portuguese. *Lingua* 121, p. 2153-2175, 2011.

RAPOSO, E.; MIGUEL, M. Introdução ao sintagma nominal. In: RAPOSO, E.; BACELAR DO NASCIMENTO, M. F.; COELHO DA MOTA, M. A.; SEGURA, L.; MENDES, A. (orgs.) *Gramática do Português*, Lisboa: Fundação Calouste Gulbenkian, vol. 1, 2013, p. 703-734.

RINKE, E. A combinação de artigo definido e pronome possessivo na história do português. *Estud. lingüíst. galega* 2, p. 121-139, 2010.

ROCA, F. The left periphery of nominal constructions and the evolution of definite determiners in Romance. *Diachronic Linguistics.* p. 495–551, 2009.

SCHMITT C.; GALVES, C. Bare nominals in Brazilian and European Portuguese: a historical perspective. *Nominals II Getebra International Workshop*. Maceio, 2016.

SCHMITT, C.; MUNN, A. The syntax and semantics of bare arguments in Brazilian Portuguese. *Linguistic Variation Yearbook*, vol. 2, number 1, p. 185-216, 2003.

SILVA NETO, S. *Introdução ao estudo da língua portuguesa*, Rio de Janeiro, Presença, 1977 [1950].

SIMONENKO, A.; CARLIER, A. Between demonstrative and definite: A grammar competition model of the evolution of French l-determiners, *Canadian Journal of Linguistics* 65(3), p. 393–437, 2020.

SIMONENKO, A.; CARLIER, A. Evolution of possessive phrases and the rise of DP in French, Spanish, and Portuguese. *Glossa a Journal of General Linguistics*, abril 2022.

SPATTI, I. *Sujeito nulo e sujeito pronominal: estudo diacrônico em cartas brasileiras e portuguesas do século 17 a 20*. 2014. 202 f. Dissertação (Mestrado em Linguística). Universidade Estadual de Campinas, 2024.

TAVARES, A. P.; MADEIRA SANTOS, C. *Africae Monumenta - A apropriação da escrita pelos africanos, Arquivo Caculo Cacahenda*, Volume I, Lisboa, Instituto de Investigação Científica Tropical, 2002.

WALL, A. The distribution of specific and definite bare nominals in Brazilian Portuguese. In: KABATEK, J.; WALL, A. (orgs.) New *Perspectives on Bare Noun Phrases in Romance and Beyond*. Amsterdam/Philadelphia: John Benjamins. 2018, p. 223-254.

http://doi.org/10.5007/1984-8420.2025.e101224

A CONSTITUIÇÃO DE 1946 E A IDENTIFICAÇÃO DA LÍNGUA NACIONAL

THE 1946 CONSTITUTION AND THE IDENTIFICATION OF THE NATIONAL LANGUAGE

Maria Carlota Rosa | Lattes | carlota@acd.ufrj.br
Universidade Federal do Rio de Janeiro

Resumo: Em linguística a definição de língua é um ponto controverso, porque operacionalmente, para definir uma língua pode-se lançar mão de critérios não linguísticos. Este texto retoma o tema, ao focalizar a questão incluída na Constituição brasileira de 1946 sobre o nome da língua nacional do Brasil. Tem por objetivo demonstrar que os critérios que embasavam ambos os lados da disputa podiam chamar para si a fundamentação na linguística, a nova ciência que começava a gozar de prestígio no país. A discussão sobre o nome da língua nacional brasileira ficou registrada nos jornais da época, e oito jornais cariocas forneceram o material de base deste trabalho. A restrição a jornais do Rio de Janeiro se deveu ao fato de ser essa cidade, à época, o Distrito Federal, onde estavam instalados o Governo e a Assembleia Constituinte.

Palavras-chave: Língua; Língua brasileira; Língua portuguesa; Constituição dos Estados Unidos do Brasil; 1946.

Abstract: The definition of language is a controversial point in linguistics, because defining a language, in general, does not rely solely on linguistic criteria. This text revisits the discussion about the name of the national language, which took place in the Brazilian Constitution of 1946, with the aim of demonstrating that the criteria that supported both sides of the dispute could find support in linguistics. The discussion was recorded in the newspapers of the time, and eight newspapers from Rio de Janeiro provided the source material for this work. The restriction to newspapers from Rio de Janeiro was due to the fact that this city, at the time, was the Federal District, where the Government and the Constituent Assembly were installed.

Keywords: Language; Brazilian Portuguese; Constitution of the United States of Brazil; 1946.

http://doi.org/10.5007/1984-8420.2025e101224

> JÁ EXISTE UMA LÍNGUA BRASILEIRA, QUE REPRESENTA UMA EVOLUÇÃO BIOLÓGICA DO PORTUGUÊS EUROPEU? A hipótese evolucionista foi muito debatida a partir de 1820, dado o prestígio da Biologia Evolutiva, e também por influência do nacionalismo desencadeado pelo Romantismo. Afirmava-se que, assim como do latim surgira na Europa o português, também deste surgiria na América o brasileiro.
>
> (Castilho, 2010: 186)

1. Introdução

O conceito de *língua* é um dos problemas da linguística, e nem sempre os linguistas concordam se estão diante de línguas diferentes ou de variedades de uma língua. Identificar uma língua como o meio de comunicação que ocupa uma área de um mapa não é simples e, por vezes, pode não ser possível (Grace, 1992). Neste texto vamos focalizar *língua* como um construto criado "em resposta a certas demandas históricas" (Rajagopalan, 2003, p. 28).

No Brasil, uma demanda histórica com efeitos sobre o conceito de *língua* — ou melhor, *idioma*, porque "só se refere à língua nacional, propriamente dita, e pressupõe a existência de um estado político, do qual seja a expressão linguística" (Camara Jr., [1965]1973, p. 222) — decorreu da Independência do país em relação a Portugal. Que idioma era a expressão da nova nação? Era necessário responder com uma construção que simbolizasse a nova geografia política. Como símbolo de poder político, *língua* pode ser tema de discussão para os poderes constituídos, como o parlamento, por exemplo.

Este texto focaliza a *questão da língua* no contexto da Constituição de 1946. Com recurso à Biblioteca Digital da Fundação Biblioteca Nacional/ BNDigital, partimos do que ficou registrado em oito jornais cariocas, a saber: *Correio da Manhã, Diário Carioca, Diário de Notícias, O Jornal, Jornal do Brasil, Jornal do Commercio, A Manhã* e *A Noite*[1]. Justificamos o recorte, restrito a jornais cariocas: o Rio de Janeiro era então a sede administrativa do país, com a Assembleia Nacional Constituinte funcionando no Palácio Tiradentes e o chefe do poder executivo despachando no Palácio do Catete. Nosso objetivo é focalizar o conceito de língua que emerge das discussões em torno do texto legal. Organizamos o texto procurando inicialmente situar o problema, por vezes designado *a questão da língua*, mostrando que a preocupação com a língua nacional esteve presente já na primeira Assembleia Geral Legislativa, ainda no primeiro reinado, e chegou ao século XX, dando origem a uma lei no Distrito Federal em 1935. As seções 3 e 4 focalizam, respectivamente, a discussão na Assembleia Constituinte, em 1946, e o parecer da comissão

[1] *O Globo* não compõe a base da BNDigital.

nomeada para cumprir o disposto na nova Constituição. Por fim, centramos a última parte no conceito de *língua*.

2. A *questão da língua*: o nome da língua nacional do Brasil

Os jornais do Distrito Federal de 16 de outubro de 1946 tinham seu espaço grandemente ocupado pelas notícias ligadas às execuções, naquela madrugada, dos criminosos de guerra que haviam sido condenados à morte pelo tribunal de Nuremberg e ao fato de que o primeiro a ser executado, Hermann Goering (1893-1946), conseguira suicidar-se, fazendo uso de cianureto ainda na cela. Com a pauta ocupada em grande parte por biografias dos 10 executados, por relatos sobre os últimos momentos de cada um e manifestações de líderes mundiais, a notícia da decisão a que chegara na véspera a comissão que deveria dizer ao Brasil que nome tinha seu idioma era gota d'água no oceano, e nesse dia 16 de outubro o resultado dos trabalhos da comissão ganhou espaço na capa apenas no *Diário Carioca* e n'*A Noite* — não como manchete. Não encontrou lugar na primeira página de outros jornais; às vezes, nem mesmo nas páginas interiores[2].

Constituída em atendimento ao Art. 35 das Disposições Transitórias da *Constituição dos Estados Unidos do Brasil,* promulgada havia menos de um mês, a comissão tinha por tarefa opinar sobre o nome da língua nacional do Brasil, trazendo um final constitucional ao debate que sacudira por décadas a agenda política e cultural do país. E que debate era esse?O português era a língua de Portugal, propriedade demonstrada na relação morfológica entre as designações para o país e a língua. Então, qual era a língua do Brasil, nação independente?

O nome de uma língua "deve ser ou ter relação imediata com o nome do país"[3], defendia na Assembleia Legislativa, em 1826, José Clemente Pereira (1787-1854). Deveria ser *língua brasileira.* E, de fato, dez anos após a Independência, o Decreto de 12 de abril de 1832[4] (Faraco, 2016, p. 162) mencionava a *língua brasileira* em seu Art. 9º:

> Art. 9º.Constando ao Intendente Geral da Policia, ou a qualquer Juiz de Paz, ou Criminal, que alguem comprou ou vendeu preto boçal, o mandará vir à sua presença, examinará se entende **a língua brazileira;** se está no Braz.il antes de ter cessado o trafico da escravatura, procurando por meo de interprete certificar-se de quando veio d'Africa, em que barco, onde

[2] Foi notícia de capa em dois jornais: *Diário Carioca* e *A Noite*. Nos demais jornais: *Correio da Manhã* – p. 2; *Diário de Notícias* – p. 3; *Jornal do Commercio* – p. 3; *A Manhã* – p. 4. Em *O Jornal* e *Jornal do Brasil* – notícia não encontrada nessa edição.

[3] In Prazeres. A Constituinte e a Constituição. *Jornal do Brasil*, 18/07/1946, p. 5.

[4] A ementa do Decreto era "Dá regulamento para a execução da Lei de 7 de Novembro de 1831, sobre o trafico de escravos".

> desembarcou, por que lugares passou, em poder de quantas pessoas tem estado, etc. Verificando-se ter vindo depois da cessação do trafico, o fará depositar, e procederá na fórma da Lei, e em todos os casos, serão ouvidas summariamente, sem delongas superfluas as partes interessadas.
>
> (Brasil, 1874, II: 101 – ênfase adicionada)

Mas essa designação não foi consistentemente empregada em outros documentos. Por exemplo: no mesmo ano de 1832, nos *Estatutos para a Academia Militar, e de Marinha da Côrte do Imperio do Brazil, na conformidade do decreto acima*, reformulados em atendimento a Decreto do Poder Executivo de 9 de março, surge *língua vulgar* entre os requisitos de admissão: "Art. 25. É indispensavel a idade de quinze annos, ***o conhecimento da grammatica da lingua vulgar,*** e das quatro opperações da arithemetica, e saber traduzir a lingua franceza" (Brasil, 1874, p. 69 – ênfase adicionada), expressão novamente encontrada, por exemplo, no Decreto do Poder Legislativo de 25 de agosto de 1832[5] (Brasil, 1874, II, p. 64).

A questão do nome da língua adentrou o século XX, ganhando um novo ângulo: "[o] nosso Brasil está colocado entre os países sem língua própria", afirmava Antenor Nascentes ([1937] 1960, p. 5); um "dialeto ultramarino do português", "em que pese ao orgulho nacional" (Nascentes, [1937] 1960, p. 253). Reconhecia, no entanto, com Leite de Vasconcelos, "que a denominação de dialeto [...] ia ferir os nossos melindres patrióticos" (Nascentes, [1937] 1960, p. 254).

Camara Jr. (1935) recorria a M. E. De Witt para argumentar que no Brasil se falava português, não brasileiro, mas isso não tornava o Brasil menos brasileiro:

> A Suíça não usa uma língua suíça, nem o Brasil, uma brasileira; nações distintas têm o espanhol como língua nacional. Tais fatos não fazem a Suíça tornar-se não-suíça, nem o Brasil tornar-se não-brasileiro; como não fazem de uma dada nação que fale espanhol uma nação espanhola.
> (M. E. De Witt, [1924], p. 8, na tradução de Camara Jr., 1935 – In Rosa, 2024)

Mas se não havia denominação consensual, empregar *português* ou *brasileiro* levava a uma tomada de posição, que podia ser contornada com o emprego de *língua nacional* e *língua pátria* (Faraco, 2016, p.167-168), denominações que ganhariam as capas dos manuais escolares. Podiam evitar uma tomada de posição, mas seu emprego não equivalia à rejeição de que no Brasil o idioma era a *língua portuguesa,* argumentava Mattoso Camara Jr.:

[5] "Art. 4º - As lições serão feitas em tres dias da semana, interpoladamente á hora, que o Presidente em Conselho julgar mais propria, e ***pela collecção das lições vertidas em vulgar,*** e demonstradas pelas figuras em madeira do sobredito autor, as quaes foram offerecidas ao Conselho do Governo desta Província para este fim."

> Alguns autores brasileiros, é certo, têm usado o adjetivo *nacional* em título de livro de gramática ou de qualquer estudo de linguagem. Nunca houve porém, o intuito de ser dado desta sorte um sucedâneo ao vocábulo *português*, até porque *nacional* é termo essencialmente relativo no sentido, em função que está na nacionalidade de quem o emprega e, se designa o nosso idioma na boca de um brasileiro, designará outro muito diverso na boca de um francês, de um italiano, ou de um espanhol.
> (Camara Jr., 1935 – In Rosa, 2024)

O nacionalismo porejava por toda a discussão[6]. Os defensores da *língua brasileira* buscaram dar respaldo científico à sua posição recorrendo à metáfora da língua como organismo vivo, que, como tal, tem como destino se reproduzir (Elia, 1961, p. 86). Sim, já havia uma evolução do português europeu no Brasil. Do mesmo modo que o português saíra do latim, a língua do Brasil saíra do português. Davam resposta afirmativa à questão de Ataliba de Castilho que abre este texto.

Na década anterior, Câmara Jr. já se opusera a essa visão:

> O que dizem alguns filólogos nossos, filiados à escola linguística (hoje envelhecida) dos chamados neogramáticos da Alemanha, é que, com o correr dos séculos, acabará formando-se no Brasil uma língua saída da língua portuguesa, como esta saiu do latim. É, portanto, a projeção em futuro remoto de um fato inconcusso do passado. Mas nem isto sequer é possível assegurar, porque as condições da civilização moderna são muito diversas daquelas em que se debatia a Europa ao alvorecer das línguas românicas. Para que estas línguas surgissem e vingassem "*foi necessário um colapso da cultura material e espiritual*" como disse recentemente em lapidar expressão Amado Alonso, de Buenos Aires, ao discutir *El problema argentino de la lengua*, magnífico trabalho onde é rejeitada, com lógica e nitidez, a probabilidade da diferenciação linguística, fatal e progressiva, dos países hispano-americanos com a Espanha e entre si.
> (Camara Jr., 1935 – In Rosa, 2024)

O debate, que tivera início após a Independência, ganhara mais repercussão ao entrar nas preocupações do recém-criado legislativo carioca, sob a forma do Projeto 62/1935. Em sua primeira legislatura, a Câmara Municipal do Distrito Federal, por quase unanimidade, decidia que *brasileiro,* não *português*, estaria nos livros didáticos e nos currículos escolares no Rio de Janeiro (Rosa, 2024). Nos jornais, nomes como Mattoso

[6] Para um panorama das visões do problema entre os estudiosos brasileiros Virgílio de Lemos (1863-1926), Renato de Mendonça (1912-1990), Antenor Nascentes (1886-1972), João Ribeiro (1860-1934), Herbert Parentes Fortes (1897-1953), Gladstone Chaves de Melo (1917-2001) e Serafim da Silva Neto (1917-1960), ver Elia ([1940] 1961).

Camara Jr.[7], Júlio Nogueira e Barbosa Lima Sobrinho escreviam contra a lei; o *Correio da Manhã* a endossava.

Em 1940, na primeira das sete conferências sobre "*A boa linguagem na reconstrução nacional*", promovidas pelo Departamento de Imprensa e Propaganda, o DIP, o diretor do Colégio Pedro II, Clóvis Monteiro, tratando do tema "*A exaltação da nacionalidade através da boa linguagem*", defendeu que não se podia dissociar a língua do Brasil daquela de Portugal[8]. O tema "palpitante" chegava ao rádio[9], era matéria de conferências[10], de artigos nos jornais[11].

A maioria dos defensores de uma língua própria nacional encontrava em *língua brasileira* a melhor designação. Uma outra possibilidade, porém, foi aventada por um professor suplementar de Português do Colégio Pedro II[12], mais tarde revisor do *Diário Oficial*, publicação da Imprensa Nacional[13]: era Alcides d'Arcanchy.

Afamado por corrigir o vocabulário empregado no Brasil[14], o que lhe garantiu a re-

[7] Foi veemente a discordância de Joaquim Mattoso Camara Jr. ao Projeto 62, expressa no artigo *Língua Brasileira do Distrito Federal*, publicado no jornal carioca *Diário de Notícias* de 14 de julho de 1935 e lido em plenário na sessão do legislativo carioca que aprovou o referido Projeto. Cabe lembrar também as palavras de um dos vereadores, Heitor Beltrão (1889-1955), após a leitura do artigo de Mattoso Camara: "não se póde tratar desta materia do ponto de vista meramente do raciocinio porque se o tratamos, dariamos, em grande parte, razão á bellissima carta, eminentemente erudita, que acaba de ser lida, neste recinto, pelo nosso collega, Sr. Vereador Romero Zander".

[8] *Jornal do Brasil*, 21/02/1940, p. 11.

[9] Ainda na véspera da promulgação da nova *Constituição,* a rádio Roquette Pinto anunciava os *Debates sobre a Língua Brasileira*: "A Rádio Roquette Pinto, prosseguindo nos debates que está levando a efeito sobre a palpitante questão do idioma nacional, irradiará a partir de quinta-feira próxima, uma série de palestras do filólogo Cândido Jucá Filho, que abordará o interessante tema da inexistência da língua brasileira. A primeira audição dessa série de palestras tem o seu início marcado para as 21 horas" (*Jornal do Brasil*, 17 de setembro de 1946, p. 9).

[10] Para 15/07/1946 anunciava-se a conferência do Coronel Altamirano Nunes Pereira no Liceu Literário Português com o "atualíssimo tema: "Haverá uma língua brasileira?"" (*O Jornal*, 13/07/1946 – 2ª. seção - p 2 – coluna Vida Portuguesa). Altamirano Nunes Pereira era catedrático da Escola de Intendência do Exército (Neme, 1941: 115).

[11] Como exemplo, "Haverá língua brasileira?", de José Oiticica (*O Jornal,* 16/06/1946 – Revista p. 1-2). No mesmo jornal, as réplicas a Oiticica: "Língua brasileira", por Edgard Sanches, em três partes (*O Jornal,* 30/06/1946; 21/07/1946; 18/08/1946) e também "Língua-padrão e língua fonte", por Herbert Parentes Fortes (*O Jornal*, 18/08/1946).

[12] Soares & Silva (2018: 156).

[13] Informação extraída da coluna de Augusto Aguiar "Tolice em vernáculo: brasilista ou brasileiro", em *A Noite,* 26/12/1952, p. 3. Aguiar defendia punição para d'Arcanchy por corrigir no *Diário Oficial* todas as ocorrências de *brasileiro*, alteradas por Alcides d'Arcanchy para *brasilista*. As correções ao *Diário Oficial* já haviam repercutido no Senado, pela voz de Aloysio de Carvalho (1901-1970), que estranhava a introdução do neologismo naquela publicação (*Jornal do Brasil*, 10/09/1952, p. 6). As correções desse tipo introduzidas em outra publicação da Imprensa Nacional, o *Diário do Congresso*, seriam alvo de críticas de outro senador, Dário Cardoso (1899-1987): a preocupação com a criação de neologismo estava desviando os revisores de sua verdadeira função (*Jornal do Brasil,* 19/05/1953, p. 6).

[14] Afora *brasilista*, talvez a correção mais famosa por ele proposta tenha sido *balípodo,* para substituir *futebol*, e seus derivados: *balipodoca* ('craque de futebol'), *balipodista* ('jogador de futebol'), *balipodizar, balipódio, balipodístico, balipodística, balipodizante, balipodizador, balipodização*, *imbalipodizável*. Outros esportes mereceram sua atenção: *balipismo* ('polo'), *balineco/balineca* ('polo aquático; jogador ou jogadora de polo

putação de filólogo e espaço continuado na imprensa por pelo menos duas décadas, d'Arcanchy contestava a denominação *brasileiro*, que substituiu por *brasilista*[15]. Numa longa entrevista ao *Jornal do Brasil* mostrava porque preferia *brasilista*:

> Falamos Português ou Brasileiro? Qual a sua opinião a respeito? — assim responde o prof. d'Arcanchy:
> Nem Português nem Brasileiro. O que falamos é **Brasiliano**, **Brasiliense**[16] ou, melhor dizendo, **BRASILISTA.**
> Discorda, então, dos que se batem pela língua Brasileira?
> Não. Não tenho em mira discordar dos dignos patriotas que, na imprensa e na Assembleia Nacional Constituinte, aplaudidos por milhões de brasilistas, conscientemente e sem tergiversações, combatem pelo nome do Brasil na denominação do nosso idioma. [...]. O essencial, no momento, é que não se persevere no inveterado erro de dizer que o idioma do Brasil é o idioma de Portugal. Acabemos com essa lenda forjada pelos brasilatas[17]. [...]
> E por que afirma que não falamos Português?
> Porque Português não é falar brasilígena, mas alienígena.[...] Conclusão: **Português** não é o idioma do Brasil, mas conforme se lê nos léxicos lusitanos: "a língua falada pelos portugueses". O que falamos, queiram ou não os brasilatas, é o **Brasilista**, ou, como dizem os Franceses: le **Brésilien**. (*Jornal do Brasil,* Idioma brasileiro, brasiliano, brasiliense ou brasilista? – 11/08/1946, 2ª. seção, p. 1-2 – ênfase no original)

A questão chegaria à pauta da Assembleia Constituinte de 1946 e se tornaria o Art. 35 das Disposições Transitórias da *Constituição dos Estados Unidos do Brasil,* promulgada em 18 de setembro de 1946. O nome a dar ao idioma nacional do Brasil não seria decisão a ser tomada pelos constituintes, que a deixavam para o Executivo, como assinalou Barbosa Lima Sobrinho[18] no *Jornal do Brasil.* A comissão de professores, escritores e jornalistas foi nomeada pelo Presidente da República Eurico Gaspar Dutra (1883-1974) por indicação do Ministro da Educação e Saúde Ernesto de Sousa Campos (1882-1970).

---

aquático'), *cestibolio* ('basquete'), *volibólio* ('vôlei'), cada um acompanhado de derivados (*Vide* d'Arcanchy. Brasilismos indispensáveis. *A Noite*, 11/01/1944, p. 7).

[15] *Brasilista* chegou a ter algum uso, como, por exemplo, em "Língua-padrão e língua fonte", de Herbert Parentes Fortes (*O Jornal*, 18/08/1946) e se tornaria também sobrenome. O filólogo José de Sá Nunes (1889-1955), segundo relato de João Paraguassú no *Correio da Manhã* de 19/02/1949, p. 12, elogiava o neologismo e se declarou o responsável pelo registro do termo (e também de *brasilismo* e *brasilizar*) no *Vocabulário Ortográfico Resumido* de 1947.

[16] Nessa mesma entrevista rejeitaria *brasiliense* e *brasiliano* porque tinham cinco sílabas, não quatro, como *brasilista*, este "mais agradável ao ouvido" e que daria lugar a derivados "mais suaves".

[17] Outra criação de d'Arcanchy: "BRASILATAS, do Grego láthas (esquecido de), são os que não pensam no Brasil, os que se esqueceram de que devem agir em prol do Brasil. Os que não pensam no povo e na democracia são DEMOLATAS" (d'Arcanchy, Língua brasileira ou idioma brasilista?. *A Manhã*, 30/06/1946, p. 3)

[18] Barbosa Lima Sobrinho. O rótulo da lingua nacional. *Jornal do Brasil,* 15/09/1946, p. 5.

Menos de um mês depois da promulgação, na tarde de 15 de outubro de 1946, a resposta da comissão era entregue ao Ministro da Educação e Saúde, que anunciava o resultado a jornalistas, conforme o noticiário do dia seguinte.

3. Os antecedentes do parecer da comissão

Na sessão da Assembleia Constituinte de 12 de setembro de 1946, o deputado Soares Filho (1874-1952) propunha a seguinte emenda aditiva às Disposições Transitórias do texto constitucional:

> Art. – O Governo nomeará uma Comissão de professores, jornalistas e escritores que, estudando as diferenças entre o Português de Portugal e o Português do Brasil, opinará sobre a questão da denominação de Língua Brasileira, tão insistentemente reclamada por muitos intelectuais e já adotada por uma resolução aprovada pela quase unanimidade do Conselho Municipal do Distrito Federal."

Se Soares Filho mencionava os muitos intelectuais favoráveis à designação por ele defendida, não é menos certo que também havia forte rejeição à mudança, expressa pela Academia Brasileira de Letras, por professores como Júlio Nogueira, Cândido Jucá Filho, Otoniel Mota e José Oiticica. A emenda seria aprovada por 124 votos a 91[19] . Entre os que se opuseram à emenda estava o deputado Hermes Lima (1902-1978), que ocuparia o plenário para dizer que seria engraçado se alguém pedisse a tradução para brasileiro do texto que a Constituinte estava redigindo em português[20].

A redação proposta inicialmente para esse artigo seria modificada: ficava de fora a referência ao que dispusera a Câmara Municipal do Distrito Federal. *Língua Brasileira*, por sua vez, passava no texto a *língua nacional*. A redação final seria ainda um pouco diferente: "O Governo nomeará Comissão de professores, escritores e jornalistas, que opine sobre a denominação do idioma nacional".

Promulgada a *Constituição* em 18 de setembro, o Ministro Ernesto de Sousa Campos agilizaria a formação da comissão. Seus critérios de escolha de nomes foram assim justificados:

> resolvi propor ao Chefe da Nação que [...] fosse integrada dos presidentes das grandes instituições tradicionais do país, dos reitores das duas Universidades com sede nesta capital, do inspetor geral de ensino

[19] Sessão da Assembleia Constituinte de 12 de setembro de 1946. No *Jornal do Brasil*, 13/09/1946, p. 9, foram 90 votos contrários, não 91.

[20] *O Jornal,* 13/09/1946, p. 4.

> do Exército, de alguns escritores e jornalistas de intensa projeção e especializados no assunto de modo a formar um grupo de quinze pessoas que pudessem representar o pensamento brasileiro.
> (*Diário Carioca*, 16 de outubro de 1946, capa)

Os nomes escolhidos foram:

1. **Claudio de Sousa** (1876-1954), presidente da Academia Brasileira de Letras;
2. **Herbert Moses** (1884-1972), presidente da Academia Brasileira de Imprensa;
3. **Álvaro Ferdinando de Sousa da Silveira** (1883-1967), catedrático da Faculdade Nacional de Filosofia da Universidade do Brasil;
4. **José Carlos de Macedo Soares** (1883-1968), presidente do Instituto Brasileiro de Geografia e Estatística e do Instituto Histórico e Geográfico Brasileiro;
5. Gal. **Francisco Borges Fortes de Oliveira** (1886-?), inspetor de ensino do Exército;
6. **Levi Carneiro** (1882-1971), presidente do Instituto Brasileiro de Educação, Ciência e Cultura;
7. **Inácio Manuel Azevedo do Amaral** (1883-1950), reitor da Universidade do Brasil;
8. **Pedro Calmon** (1902-1985), diretor da Faculdade Nacional de Direito da Universidade do Brasil;
9. Pe. **Augusto Magne**, S. J. (1887-1966), catedrático de Filologia Românica da Universidade do Brasil;
10. **Júlio Nogueira** (1873-?), autor do manual *A linguagem usual e a composição,* premiado pela Academia Brasileira de Letras com o Prêmio Alves;
11. Pe. **Leonel Franca**, S. J. (1893–1948), reitor da Universidade Católica[21];
12. **Clóvis do Rego Monteiro** (1898-1961), diretor do Internato do Colégio Pedro II;
13. **Gustavo Capanema** (1900-1985), deputado;
14. **Gilberto Freyre** (1900-1987), deputado;
15. **Afonso d'Escragnolle Taunay** (1876-1958), escritor e historiador.

Em 8 de outubro Ernesto de Sousa Campos instalava a comissão e dava posse a seus membros. Ausentes ao ato os três últimos nomes da lista acima. Para a presidência da comissão foi eleito por unanimidade José Carlos de Macedo Soares, mas, em razão de ser,

[21] Somente no ano seguinte, 1947, a Universidade Católica receberia do Vaticano o título de *Universidade Pontifícia*.

naquele momento, interventor no estado de São Paulo, Pedro Calmon sugeriu a escolha também de um vice-presidente. Por aclamação, o escolhido foi o presidente da Academia Brasileira de Letras[22]. Para a relatoria, Sousa da Silveira.

A segunda e última reunião ocorreria uma semana mais tarde, em 15 de outubro. Além das ausências de Capanema, Freyre e Taunay, contou-se também a de Macedo Soares[23], o que levou a presidência da sessão para Claudio de Sousa. O relator, Sousa da Silveira, leu seu voto, na forma "do projeto de relatório da comissão"[24]. Seria esse o texto-resposta da comissão, assinado pelos 11 membros presentes. Por unanimidade a comissão decidia que a denominação do idioma nacional era *língua portuguesa*.

Em 18 de outubro, em carta aberta no *Jornal do Brasil,* Oto Prazeres — a quem se atribuiu a ideia da emenda aditiva proposta por Soares Filho[25] e que escrevera diversos artigos no *Jornal do Brasil* em defesa da língua brasileira — atacava o Ministro da Educação e Saúde, alegando a irregularidade do ato de Sousa Campos e o quanto de "monstruoso", "sob o ponto de vista da boa ética", havia "na nomeação da Comissão e no seu parecer, de antemão conhecido e preparado". Guardou espaço para o ataque à moral dos membros da Comissão, "que deveria estar na altura sob o ponto de vista da insuspeição e capaz de um estudo imparcial e demorado", mas que "tão altos personagens", se precisaram de lançar mão de "procedimentos semelhantes", é que reconheciam a força da questão, e que lhes era "preciso, no combate, não escolher meios":

> Há alguma coisa a dizer contra a moral e contra a competencia dos nomes escolhidos? Absolutamente não. Todos são brasileiros merecedores do mais alto respeito.
> Mas á luz da sã moral, não poderiam ser nomeados e não deveriam ter aceitado a indicação.
> (Prazeres, 1946b)

E Prazeres completava:

> Há, na Camara dos Deputados nada menos de 140 Deputados que são partidarios da Lingua Brasileira e quais foram os unicos indicados? Os Srs. Gustavo Capanema e Gilberto Freire, que ali combateram a ideia. Sem

[22] *O Jornal*, Qual o idioma falado no Brasil? Instalada a Comissão nomeada? De acordo com a Constituição, para opinar a respeito. 3ª. feira próxima a segunda reunião. 9/10/1946, p. 2.

[23] Os quatro ausentes enviaram o voto por escrito.

[24] *Diario de Notícias*, O idioma nacional é língua portuguesa, 16/10/1946. p. 3.

[25] "O autor da ideia, [...] não é deputado nem senador, mas um funcionário categorizado da casa, que permanece atrás da Mesa durante as sessões da Constituinte" apontava *O Jornal* (26/06/1946, p. 4). Esse funcionário era Oto Prazeres, secretário da Presidência da Assembleia Constituinte (*O Jornal,* 13/09/1946, p. 5).

> duvida, são ambos altos expoentes da cultura brasileira, mas não podiam, em boa ética, fazer parte da assembleia.
> Na Academia Brasileira de Letras foram escolhidos a dedo os contrários e assim nos demais setores.
> (Prazeres, 1946b)

4. O que dizia o parecer

O parecer, aprovado por unanimidade, assumia a forma de opinião: "opino que a denominação do idioma nacional do Brasil continue a ser: Lingua Portuguesa". Para sustentar essa opinião, o texto foi construído em três partes: *breve retrospecto histórico*, *considerações linguísticas* e *conclusão*.

O breve histórico inicial estabelecia a origem do idioma nacional: a língua portuguesa trazida no século XVI para esta parte do mundo. Aqui havia o tupi, "a língua dos índios". Depois do triunfo de uma "civilização muito superior", o monolinguismo tupi fora substituído pelo monolinguismo português:

> É a lingua portuguesa aquela em que nós, brasileiros, pensamos; em que monologamos, em que conversamos, que usamos no lar, na rua; na escola, no teatro, na imprensa, na tribuna, com que nos interpela, na praça pública, o transeunte desconhecido que nos pede uma informação; é, por assim dizer, a nossa língua de todos os momentos e de todo os lugares.

A língua "de todos os momentos e lugares" tem o "padrão ideal da lingua literária", exemplificada nas gramáticas. O parecer não parecia se afastar da ideia de um "padrão patriótico de unidade nacional" civilizado, discutido poucos anos antes, em 1937, no Congresso da Língua Nacional Cantada (ver Monteiro, 2021).

> Desde os primeiros tempos da nossa história já apareciam, escritas em português, obras relativas ao Brasil; e tôda a nossa literatura, de então para cá, tem sido vazada em lingua portuguêsa. Os nossos mais altos escritores, uns com maior, outros com menor apuro estilístico, êstes aproximando-se mais, aquêles menos, do padrão ideal da lingua literária, todos escreveram em português. Assim o fizeram José Bonifácio, João Francisco Lisboa, Odorico Mendes, Gonçalves Dias, Alvares de Azevedo, Casimiro de Abreu, Castro Alves, Fagundes Varela, Gonçalves de Magalhães, Pôrto Alegre, Manuel Antônio de Almeda, Alencar, Macedo, Machado de Assis, Aluizio Azevedo, Joaquim Nabuco, Eduardo Prado, Rui Barbosa, Taunay, Afonso Arinos, Euclides da Cunha; Raul Pompéia, João Ribeiro, Olavo Bilac, Alberto de Oliveira, Raimundo Correia, Vicente de Carvalho, etc., etc. .
> A própria literatura nossa regional exprime-se numa lingua que, apesar de tudo, não deixa de ser a portuguêsa; e o falar dialetal da nossa gente inculta é, na essência, lingua portuguêsa.

> Alguns dos grandes escritores brasileiros, como Rui Barbosa, João Ribeiro e Raimundo Correia, que no principio da sua carreira literária, embora escrevessem em português, se afastavam um pouco do bom tipo linguístico, esforçaram-se depois por acompanhá-lo de mais perto, e conseguiram· tornar-se modelos da mais formosa vernaculidade.

As *considerações linguísticas* rebatiam o argumento da evolução linguística inexorável:

> **Considerações linguísticas:** E' inteiramente falso dizer-se que; assim como do latim vulgar transplantado para o ocidente da Peninsula Ibérica resultou o idioma português, assim do português trazido· para o Brasil resultou a lingua brasileira.
> Proceder dêsse modo é comparar fatos diversos, e a conclusão a que se chega percorrendo semelhante caminho será, forçosamente, errada.

Em cena, o padrão literário:

> Com o português transplantado para o Brasil outros, bem outros são os fatos. Nunca ficou em abandono igual ao do latim vulgar na Península Ibérica; ao contrário, estêve sempre em contacto com o da metrópole, onde a ·literatura atingiu alto cume no século XVI e continuou no seu desenvolvimento florescente até os nossos dias. Frei Vicente do Salvador, nascido no Brasil, escrevia em português a sua História do Brasil; o Padre Antônio Vieira pregava no Brasil muitos dos seus sermões; Morais, nascido no Brasil, compunha o seu Dicionário da Lingua Portuguêsa; brasileiros iam a Portugal e formavam-se na Universidade de Coimbra; D. João VI, com a sua corte, veio para o Rio. de Janeiro e aqui permaneceu por mais de uma década. Os nossos grandes poetas épicos Santa Rita Durão e Basilio da Gama.; outros ilustres poetas nossos, como Cláudio Manuel, Alvarenga Peixoto, etc., escreviam em excelente lingua portuguêsa, com os olhos sempre voitádos para os monumentos literários de Portugal.

A homogeneidade no território brasileiro refletia aquela no outro lado do Atlântico:

> As palavras brasileiras são iguais às portuguêsas na sua composição fonética, apenas diferindo na pronúncia; os nomes de números são os mesmos em Portugal e no Brasil; as conjugações são as mesmas, num e noutro país; as mesmas são também as palavras gramaticais, os pronomes (pessoais, possessivos, demonstrativos, relativos, interrogativos, indefinidos), os artigos, os advérbios (de tempo, modo, quantidade e, lugar, afirmação, negação), as preposições e as conjunções. Em geral é o mesmo o gênero gramatical, cá e lá; são as mesmas as regras de formação do plural; o mesmo o sistema de graus de substantivos e adjetivos; os mesmos os preceitos de concordância nominal e verbal; quase na totalidade dos casos é a mesma a regência dos complementos dos nomes e dos verbos; o mesmo o emprêgo de modos e tempos, e a mesma a estrutura geral do periodo quanto à sucessão das orações e à ligação de umas com outras.

As listas de formas que poderiam ser consideradas *brasileirismos* vinham sendo reduzidas, porque "alguns dêles existem em dialetos portuguêses [...] e que, se outros podem ser admitidos como inovações nossas, podem também considerar-se relíquias brasileiras de arcaismos portugueses".

Decidia-se que o nome da língua oficial do Brasil deveria ser língua portuguesa[26]:

> **Conclusão:** À vista do que fica exposto, a Comissão reconhece e proclama esta verdade: o idioma nacional do Brasil é a Língua Portuguesa. E, em consequência, opino que a denominação do idioma nacional do Brasil continue a ser: Língua Portuguesa. Essa denominação, além de corresponder à verdade dos fatos, tem a vantagem de lembrar, em duas palavras - Língua Portuguesa - história da nossa origem e a base fundamental da nossa formação de povo civilizado.

5. Uma língua ou mais de uma?

"Então o que é uma língua? Dizem por brincadeira que uma língua é aquilo que tem um exército e uma marinha. Não é um conceito linguístico, nem uma definição linguística" — ironizava Chomsky (1977, p. 175). Mas as críticas não vieram apenas de Chomsky. Chambers & Trudgill (1980, p. 5), por exemplo, deixaram claro que *língua* não era de modo algum uma noção linguística, mas um termo "relativamente não técnico". Apesar das críticas, o termo é corrente na área: falamos de línguas minorizadas, de famílias de línguas; de políticas linguísticas; organizamos currículos e departamentos com base nas línguas focalizadas. Os linguistas em trabalho campo indicam a língua ou a variedade em estudo.

Na década de 1970, Joseph E. Grimes, então editor do *Ethnologue* — um catálogo de línguas do mundo —, começou a desenvolver uma padronização na identificação de línguas que está na origem da ISO 693-3 e nos critérios em uso no *Ethnologue*.

O primeiro critério para distinguir línguas é a inteligibilidade: "Duas variedades relacionadas são normalmente consideradas variedades da mesma língua se os falantes de uma variedade compreendem a outra sem necessidade de aprendizado" (Critérios ISO 639-3 para identificação de línguas).

E se os falantes se entendem mais ou menos? "Se, na fala, a inteligibilidade entre as

[26] A nova Constituição não encerraria de vez a questão da língua fora do parlamento. Ainda em 1952, o presidente da Academia Brasileira de Letras Aníbal Freire (1884-1970) informava aos demais membros que o Clube Tertúlia de São Paulo comunicara à ABL que rejeitava *brasileiro* e *brasilista*, tendo adotado *brasilês* (*Jornal do Brasil*, 31/10/1952, p. 6). Depois de lembrar que Varnhagen preferia *brasiliense* e Roquette Pinto, *brasiliano*, o colunista M.C.F. pedia ao Presidente Juscelino Kubitschek, que à época construía Brasília, que retomasse o gentílico *brasilista* e o aplicasse ao que dissesse respeito à nova capital (*Jornal do Brasil*, 29/01/1957, 1º cad. p. 5).

variedades é marginal, a existência de literatura comum ou de identidade etnolinguística com uma variedade central que ambas as variedades compreendem pode ser um forte indicador de que deveriam ser consideradas variedades da mesma língua" (Critérios ISO 639-3 para identificação de línguas).

Há, porém, um outro critério: "Se a inteligibilidade é suficiente para permitir a comunicação, pode ser um forte indicador de que deveriam ser consideradas línguas diferentes a existência de identidades etnolinguísticas distintas bem estabelecidas" (Critérios ISO 639-3 para identificação de línguas).

Se os dois primeiros critérios acima permitiam argumentar em favor de uma mesma língua para Brasil e Portugal, este último coloca em cena o imaginário binacional: duas nações, duas línguas.

Simons & Gordon Jr, (2006, p. 250), num verbete sobre o *Ethnologue,* defendiam que uma língua deveria ser definida operacionalmente: "That is, the definition of language one chooses depends on the purpose one has in identifying a language. Some base their definition on purely linguistic grounds. Others recognize that social, cultural, or political factors must also be taken into account". E a inteligibilidade, que parecia ser o fator mais importante?

Vimos que em 1946 um dos constituintes brincou que seria engraçado traduzir a Constituição para brasileiro. Essa situação encontrou paralelo no noticiário internacional mais recente, num episódio referido em Anderson (2012, p. 79). Dentre os países surgidos da desintegração da antiga Iugoslávia está a Macedônia do Norte. Kiro Gligorov (1917-2012), o primeiro presidente da Macedônia do Norte, visitou a Bulgária em 1995. Para os búlgaros o macedônio é um dialeto do búlgaro e perfeitamente compreensível. Para o novo país não era um dialeto de uma língua de um outro país, mas uma língua distinta, identificada com o novo país. "Melindres patrióticos"? Gligorov se fez acompanhar de um intérprete. Os acordos que motivaram a visita presidencial não foram assinados, mas o lado macedônio, entre outras exigências, havia incluído uma cláusula que dispunha que os documentos teriam de ser escritos na língua oficial da Macedônia e na língua oficial da Bulgária.

Como o conceito de língua responde a "a certas demandas históricas", a confusão se estabelece, na dependência de relações políticas como desabafava Moore acerca do trabalho com línguas brasileiras:

> Um problema é a confusão frequente entre línguas, dialetos e grupos étnicos. Por exemplo, na família Mondé do tronco tupi, a fala dos Gavião

> de Rondônia e a fala dos Zoró são geralmente listadas como línguas distintas, enquanto, de fato, são dialetos tão próximos quanto o português de Salvador e o português de São Paulo. Os falantes desses dialetos podem, dependendo das relações políticas do momento, afirmar que os dois dialetos são idênticos ou que são bem diferentes. Qualquer critério técnico para distinguir entre dialetos de uma mesma língua e línguas distintas tem os seus limites; todavia, linguistas geralmente utilizam o critério de inteligibilidade mútua. Sem critérios desse tipo, esforços para registrar línguas podem se complicar com um número indefinido de formas de fala consideradas como línguas distintas. Nossa sugestão seria utilizar um critério técnico para agrupar dialetos da mesma língua, mas também obter informações sobre o sentimento dos falantes em relação às outras variedades e sobre a situação sociolinguística dos grupos envolvidos.
>
> (Moore, 2007)

A leitura dos jornais da época aponta para o prestígio crescente que a linguística começava a ter no Brasil. Como ciência, dava respaldo aos argumentos empregados de ambos os lados da disputa. Mas havia um problema, e voltamos ao texto de Rajagopalan (2003: 25), citando Neil Smith: "a linguística não versa sobre a linguagem, nem sobre as línguas [...]; ela versa sobre as gramáticas".

Referências

A MANHÃ. O idioma nacional é língua portuguêsa. *A Manhã,* 16/10/1946, p. 4. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=116408&pasta=ano%20194&pesq=&pagfis=31668

A NOITE. O idioma nacional é língua portuguesa. *A Noite,* 16/10/1946, capa. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=348970_04&pasta=ano%20194&pesq=&pagfis=42724

ANDERSON, Stephen R.. *Languages: A Very Short Introduction*. Oxford: Oxford University Press, 2012.

BN DIGITAL BRASIL/ Hemeroteca Digital. Disponível em: https://bndigital.bn.gov.br/hemeroteca-digital/

BRASIL. *Collecção das leis do Imperio do Brazil. 1832 parte primeira.* Rio de Janeiro: Typographia Nacional, 1874. Disponível em: https://bd.camara.leg.br/bd/bitstream/handle/bdcamara/18461/colleccao_leis_1832_parte1.pdf?sequence=1

BRASIL. *Collecção das leis do Imperio do Brazil. 1832 parte segunda.* Rio de Janeiro: Typographia Nacional, 1874. Disponível em: https://bd.camara.leg.br/bd/bitstream/handle/bdcamara/18461/colleccao_leis_1832_parte2.pdf?sequence=3

BRASIL. Constituição dos Estados Unidos do Brasil (de 18 de setembro de 1946). https://www.planalto.gov.br/ccivil_03/constituicao/constituicao46.htm

BRASIL/Ministério da Justiça e Negócios Interiores. Parecer da Comissão instituída nos termos do art. 35 do Ato das Disposições Constitucionais Transitórias. *Revista de Direito Administrativo,* 8, p. 330-333, 1947. Disponível em: https://periodicos.fgv.br/rda/article/view/9853

CAMARA Jr., Joaquim Mattoso. *Dicionário de filologia e gramática referente à língua portuguesa*. 5. ed. Rio de Janeiro: J. Ozon, [1965] 1973.

CASTILHO, Ataliba T. de. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2010.

CHAMBERS, J. K. & TRUDGILL, Peter. *Dialectology*. Cambridge: Cambridge University Press, 1980.

CHOMSKY, Noam. *Diálogos com Mitsou Ronat.* Trad. de Álvaro Lorencini e Sandra Margarida Nitrini. São Paulo: Cultrix, [1977] s.d.

CORREIO DA MANHÃ. O idioma nacional é língua portuguesa. *Correio da Manhã,* 16 de outubro de 1946, p. 2. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=089842_05&pasta=ano%20194&pesq=&pagfis=33574

D'ARCANCHY, Alcides Carlos. Brasilismos indispensáveis. *A Noite,* 11/01/1944, p. 7. Disponível em: https://memoria.bn.gov.br/docreader/DocReader.aspx?bib=348970_04&pagfis=24808

DIÁRIO CARIOCA. Constitucionalmente: nossa língua é mesmo o português. *Diário Carioca,* 16/10/1946 – capa. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=093092_03&pasta=ano%20194&pesq=&pagfis=26349

DIÁRIO DE NOTÍCIAS. O idioma nacional é língua portuguesa, *Diário de Notícias,* 16/10/1946. p. 3. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=093718_02&pasta=ano%20194&pesq=&pagfis=30127

ELIA, Sílvio. *O problema da língua brasileira*. Rio de Janeiro: Instituto Nacional do Livro/Ministério da Educação e Cultura, [1940] 1961.

FARACO, Carlos Alberto. 2016. *História sociopolítica da língua portuguesa*. São Paulo: Parábola.

FORTES, Herbert Parentes. Língua-padrão e língua fonte, *O Jornal*, 18/08/1946, Revista, p. 1, 2, 4. Disponível em: https://memoria.bn.gov.br/DocReader/DocReader.aspx?bib=110523_04&Pesq=%22l%c3%adngua%20portuguesa%22&pagfis=34579

GRACE, George W. How Do Languages Change? (More on "Aberrant" Languages). Oceanic Linguistics, 31(1), p. 115-130 (Summer, 1992). Disponível em: http://www.jstor.org/stable/3622968?origin=JSTOR-pdf

ISO/ INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 639. Language codes, s.d. Disponível em: https://www.iso.org/iso-639-language-codes.html

JORNAL DO BRASIL. Acta da 49ª Sessão Ordinária da Câmara Municipal do Distrito Federal - de 18 de julho de 1935. *Jornal do Brasil*, 20/07/1935, p. 23-26. Disponível em: https://memoria.bn.br/DocReader/docreader.aspx?bib=030015_05&pasta=ano%20193&pesq=&pagfis=55565

JORNAL DO BRASIL. Idioma brasileiro, brasiliano, brasiliense ou brasilista?, 11/08/1946, 2ª. seção, p. 1-2. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_06&pasta=ano%20194&pesq=&pagfis=41233 (cont) p. 2. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_06&pasta=ano%20194&pesq=&pagfis=41234

JORNAL DO BRASIL. A sessão da Assembleia Constituinte. 13/09/1946, p. 9. Disponível em: https://memoria.bn.gov.br/DocReader/DocReader.aspx?bib=030015_06&Pesq=disposi%c3%a7%c3%b5es%20transitorias&pagfis=41923

JORNAL DO BRASIL. Brasilista em lugar de brasileiro. *Jornal do Brasil,* 10/09/1952. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_07&pasta=ano%20195&pesq=brasilista&pagfis=22876

JORNAL DO BRASIL. Na Academia Brasileira. *Jornal do Brasil,* 31/10/1952, p. 6. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_07&pasta=ano%20195&pesq=brasilista&pagfis=24228

JORNAL DO BRASIL. Com a Imprensa Nacional. *Jornal do Brasil,* 19/05/1953, p. 6. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_07&pasta=ano%20195&pesq=brasilista&pagfis=29650

JORNAL DO BRASIL. Em respeito ao disposto no artigo 35, das disposições transitórias da nova Constituição. *Jornal do Brasil*, 9 de outubro de 1946, p. 7. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_06&pasta=ano%20194&pesq=%22lingua%20portuguesa%22&pagfis=42451

JORNAL DO COMMERCIO. Lingua portuguesa. *Jornal do Commercio*, 16/10/1946, p. 3. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=364568_13&pasta=ano%20194&pesq=&pagfis=32795

LIMA SOBRINHO, Barbosa. O rótulo da lingua nacional. *Jornal do Brasil*, 15/09/1946, p. 5. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_06&pasta=ano%20194&pesq=&pagfis=41959

M.C.F. Neologismo. *Jornal do Brasil,* 29/01/1957, 1º caderno, p. 5. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_07&pasta=ano%20195&pesq=brasilista&pagfis=70031

MONTEIRO, Luciano. *Modernismo como política de língua: o Congresso da Língua Nacional Cantada (1937)*. Niterói: Universidade Federal Fluminense/ Programa de Pós-Graduação em Estudos da Linguagem. Tese de Doutorado. 2021. Disponível em: https://app.uff.br/riuff/handle/1/24834

MOORE, Denny. Línguas indígenas: situação atual, levantamento e registro. *Patrimônio, Revista Eletrônica do IPHAN.* Línguas do Brasil. Nº 6 — Jan. /Fev. de 2007. Disponível em: https://acervo.socioambiental.org/index.php/acervo/noticias/linguas-indigenas-situacao-atual-levantamento-e-registro

NASCENTES, Antenor. *O idioma nacional*. 3ª. ed. de acordo com a Nomenclatura Gramatical Brasileira. Rio de Janeiro: Livraria Acadêmica, [1937] 1960.

NEME, Mário. A acentuação na ortografia simplificada. *Revista do Arquivo Municipal,* 73, p. 101-141, 1941.

O JORNAL. A língua brasileira. *O Jornal,* 26/06/1946, p. 4. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=110523_04&pasta=ano%20194&pesq=&pagfis=33707

O JORNAL. Vida Portuguesa. *O Jornal,* 13/07/1946 – 2ª. seção - p 2.Disponível em: https://memoria.bn.gov.br/DocReader/DocReader.aspx?bib=110523_04&Pesq=%22l%c3%adngua%20portuguesa%22&pagfis=33985

O JORNAL. Comissão para estudar a denominação da língua nacional. *O Jornal,*

13/09/1946, p. 5. Disponível em: https://memoria.bn.gov.br/DocReader/DocReader.aspx?bib=110523_04&Pesq=%22l%c3%adngua%20portuguesa%22&pagfis=34973

O JORNAL. Qual o idioma falado no Brasil? Instalada a Comissão nomeada? De acordo com a Constituição, para opinar a respeito. 3ª. feira próxima a segunda reunião. *O Jornal,* 9/10/1946, p. 2. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=110523_04&pasta=ano%20194&pesq=&pagfis=35400

PARAGUASSÚ, João. Brasilistas. *Correio da Manhã*, 19/02/1949, p. 12. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=089842_05&pasta=ano%20194&pesq=brasilista&pagfis=45904

PRAZERES, Oto. A Constituinte e a Constituição. *Jornal do Brasil*, 18/07/1946, p. 5, 1946a. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_06&pasta=ano%20194&pesq=&pagfis=40701

PRAZERES, Oto. Carta ao Sr. Ministro da Educação. *Jornal do Brasil*, 18/10/1946, p. 5, 1946b. Disponível em: https://memoria.bn.gov.br/DocReader/docreader.aspx?bib=030015_06&pasta=ano%20194&pesq=&pagfis=42633

RAJAGOPALAN, Kanavillil. *Por uma linguística crítica: linguagem, identidade e a questão ética*. São Paulo: Parábola, 2003.

ROSA, Maria Carlota. Mattoso Camara num episódio da política carioca: a "Língua Brasileira do Distrito Federal" . *Confluência*, 67, p. 42-81, . 2024. Disponível em: https://www.revistaconfluencia.org.br/rc/issue/view/71

SIMONS, Gary F. & GORDON Jr., Raymond G.. Ethnologue. In: In. BROWN, Keith (ed.). Encyclopedia of language and linguistics. Boston: Elsevier. 2nd ed. 14vv. v.4. p. 250-253, . 2006.

SOARES, Jefferson da Costa & SILVA, Gustavo da Motta. Dentre a Reforma Rocha Vaz e o Estado Novo: os professores suplementares do Colégio Pedro II. *História da Educação,*22 (56), p. 146-164. set./dez, . 2018. Disponível em: https://www.scielo.br/j/heduc/a/cvxTf7Jw5YGKMyJLH99rDPB/abstract/?lang=pt

http://doi.org/10.5007/1984-8420.2025.e108152

HISTÓRIA SOCIAL, CONTATO E A AQUISIÇÃO DO PORTUGUÊS BRASILEIRO COMO L2

SOCIAL HISTORY, CONTACT AND THE ACQUISITION OF BRAZILIAN PORTUGUESE AS L2

Maria Eugenia Lammoglia | Lattes | eugenia@letras.ufrj.br
Universidade Federal do Rio de Janeiro/CNPq

Resumo: Este artigo reinterpreta resultados de análises contrastivas de peças de teatro portuguesas e brasileiras escritas ao longo dos séculos XIX e XX, levando em conta informações sobre os quase quatrocentos anos de colonização do Brasil, em que a presença de escravizados africanos e seus descendentes, sempre superior à de imigrantes portugueses, e o contato entre eles foi intenso. A grande mobilidade populacional propiciada pelas atividades nos engenhos, no garimpo, na cultura cafeeira, entre outras, contribuiria para a miscigenação e a propagação do português por esse grande contingente, formado majoritariamente por analfabetos, tanto os escravos e trabalhadores portugueses quanto portugueses os donos de terras e suas famílias. Os letrados constituíam uma parte ínfima da população, que aprendia o português em Coimbra ou em escolas só acessíveis a uma elite privilegiada. Além da deriva do português arcaico, um fruto desse contato é a emergência de uma variedade do português com significativa contribuição africana, muitas vezes ignorada, porque a escrita produzida por uma minoria não dava espaço ao português geral aqui aprendido como L2. O lento processo de acesso à escola, só iniciado a partir da virada do século XIX, seguindo o modelo lusitano, ainda hoje presente nos livros didáticos, produziria uma diglossia no Brasil entre o português oral e o escrito. Evidências empíricas dessas duas gramáticas são observadas na semelhança entre a escrita portuguesa e brasileira dos anos 1800 e quase metade dos anos 1900, quando finalmente, a gramática brasileira começa a mostrar sua face. Só com o desenvolvimento dos estudos linguísticos com base em gravações a partir dos anos 1970, vemos a confirmação de que os mesmos traços que caracterizam a sintaxe do português brasileiro urbano, "culto" ou "popular", são atestados nas peças do último quartel do século XX.

Palavras-chave: História social. Contato linguístico. Aquisição de L2. Português brasileiro. Português europeu.

Abstract: This article re-interprets results of contrastive analyses of European and Brazilian theatre plays written across the 19th and the 20th centuries taking into account information about almost four centuries of colonization in Brazil, when the presence of slaved Africans and their descendants, always outnumbering the Portuguese, and the contact among them was intense. As a consequence of the activities in the sugar mills, in the mines of gold and diamonds, and in the coffee farming, among others, there was great mobility from the north-east to the south-east and to the center of the country. This would contribute to miscigenation and the propagation of the Portuguese acquired by the large contingent, constituted mostly by illiterate speakers, not only slaves and Portuguese workers but landlords and their families as well. Literate people constituted a tiny part of the population, who studied in Coimbra or had accesss to a few schools available to a privileged elite. Besides the drift of Old Portuguese, a consequence of such contact was the emergence of a variety of Portuguese with significant African contribution, many times ignored, because the writing practiced in Brazil by few would not give space to the grammar acquired here as L2. The slow process of access to school, from the turn of the 19th century on, following the European model, still present in school books and grammars, would give rise to a diglossy between speech and writing in Brazil. Empirical evidence of these two grammars are shown in the similarity between the Portuguese and the Brazilian writing attested in the 1800s and almost half of the 1900s, when finally Brazilian grammar begins to show its face. Only since the development of linguistic investigations of recorded speech from the 1970s on could we confirm that the same features characteristic of urban Brazilian syntax, regardless of the level of formal education, are the same attested in the plays of the last quarter of the 20th century.

Keywords: Social history. Language contact. L2 Acquisition. Brazilian Portuguese. European Portuguese.

1 Introdução

Este artigo tem como objetivo reinterpretar uma série de análises contrastivas de peças de caráter popular, escritas ao longo dos séculos XIX e XX, principalmente por portugueses nascidos, em sua maioria, em Lisboa, e por brasileiros, nascidos no Rio de Janeiro. Essa reinterpretação é feita à luz de dados da história social do Brasil-Colônia, quando se deu o desenvolvimento da gramática do português brasileiro a partir, especialmente, do intenso contato entre as sucessivas levas de africanos escravizados e seus des-

cendentes e as de imigrantes portugueses, na sua quase totalidade analfabetos, vindos de diferentes regiões de Portugal, sempre em quantidade inferior à de escravizados durante esse período. Graças a esse contato, os africanos aprenderam o português, sem deixar de imprimir suas marcas na aquisição da nova gramática, marcas que acabaram por se espalhar pelo país, caracterizando o português brasileiro (PB) em aspectos que o distinguem do português europeu (PE), embora traços decorrentes da deriva lusitana constituam importante face da gramática brasileira. Sem considerar o acesso à escolarização tardia, que viria a ensinar traços de uma gramática não adquirida em solo brasileiro, é inegável que os que chegaram a aprender traços dessa segunda gramática da escola só utilizam boa parte deles na escrita e na fala muito monitorada. Assim, dadas as pequenas diferenças sintáticas que distinguem a fala espontânea dos mais e a dos menos escolarizados, não se justifica uma rígida divisão do português brasileiro urbano em "culto" e "popular". Temos evidências empíricas que nos permitem fazer essa afirmação.

A Seção 2 apresenta um pouco da história social do Brasil, desde o início da colonização, que coincidiu com a chegada da primeira leva de africanos escravizados em 1538, e traz informações demográficas sobre os mais de trezentos e cinquenta anos de tráfico, antes comandado por Portugal, que teve a participação ativa de reis e chefes africanos, associados ao "empreendimento", e que teria mais tarde a associação de navios partindo do Brasil, com o mesmo propósito. A seção ainda contempla as primeiras discussões, já no Brasil-Império, sobre a necessidade de alfabetizar a população, analfabeta em sua quase totalidade, um processo que se estenderia até a virada do século XIX, acabando por levar à adoção do modelo de gramática do colonizador. Essa gramática do português europeu moderno (PE) fixada em fins do século XVIII, trazia inovações em relação à gramática do Português Clássico (PC), atestadas a partir do século XVI, e implementadas em Portugal (também eles tiveram seus contatos!), que acabaram por afetar a fonologia e a sintaxe lusitanas, o que sinalizava mudanças de rumo aqui e além-mar, distanciando ainda mais o português do Brasil e de Portugal. Assim, na ocasião em que se decide elaborar um modelo de gramática no Brasil, o conhecimento de tais mudanças por uma elite intelectual, que consistia numa minoria, e a "superioridade cultural" do europeu, viriam a inspirar os intelectuais brasileiros, que, independentes politicamente, se tornavam subservientes a uma gramática com tantos traços diferentes - tanto em relação ao português clássico (PC) quanto ao português adquirido como L2 no Brasil ao longo da colonização. As discussões em torno das questões gramaticais e a lenta implantação de escolas nos prin-

cipais centros populacionais, Salvador, Rio de Janeiro e São Paulo, a partir do século XIX, e as sucessivas e fracassadas tentativas de universalização do acesso à escola ao longo do século XX serão aqui referidas.

A Seção 3 mostra o quanto a escrita dos poucos letrados do século XIX e de boa parte do século XX "engana" o analista das peças de teatro popular (e certamente de outros gêneros), que acaba por achar que se encontra diante de material que representa o português brasileiro daquela época. Como se verá, só a partir da segunda metade do século XX a face do PB começa a se revelar nas peças de teatro, que trarão evidências empíricas ao que este artigo deseja mostrar: que as análises contrastivas PE-PB que sugerem a presumível mudança na sintaxe brasileira revelada nas curvas das figuras é, na verdade, a revelação da verdadeira gramática do PB, até então escondida sob uma gramática lusitana aprendida por uma minoria com acesso ao seu ensino. As evidências vêm de análises dos padrões de interrogativas-Q, das realizações do sujeito pronominal, do uso de *haver* existencial e dos clíticos não dêiticos - o indefinido *se* para a representação do sujeito indeterminado (de referência arbitrária e genérica), e o acusativo *o/a* e dativo *lhe* para referência à 3ª pessoa e suas formas variantes.

A Seção 4 apresenta traços da gramática do PB a partir do momento em que tivemos acesso a amostras da língua oral, gravadas desde o projeto Norma Urbana Culta (NURC), desenvolvido a partir dos anos 1970, ao mesmo tempo em que foram feitas gravações no Rio de Janeiro da fala de alunos adultos do Projeto Movimento Brasileiro de Alfabetização (MOBRAL), e, nos anos 1980, do Projeto Programa de Estudos sobre o Uso da Língua (PEUL), com falantes do Rio de Janeiro com ensino fundamental e médio. Este último serviria como modelo para a formação de outros bancos de dados pelo nordeste (ValPB) e sul (VarSul), espalhando-se por tantos outros estados do Brasil. Veremos que os resultados revelados nas análises mostram o que é, de fato, o PB, resultante do contato intenso, revelado na fala "não culta" e "culta", com muito mais semelhanças do que diferenças.

O êxito relativo dos lentos programas de alfabetização conseguiu ensinar a concordância verbal e nominal, que passaria a fazer parte do *input* que uma parcela já escolarizada da população brasileira urbana oferecia a seus filhos, já como L1. No mais, nossa sintaxe tem muitos traços comuns a todas as camadas sociais, como mostraremos. Não há razão para supor que os falantes cultos das cinco capitais mais populosas do Brasil nos anos

1970, gravados pelo Projeto NURC, perderiam traços de uma gramática presumivelmente adquirida após o desenvolvimento do PB depois de quase 400 anos da colonização, justamente quando começa a aumentar o acesso à escola. Concluímos o capítulo com a certeza de que a distância entre a fala e a escrita vem diminuindo em favor da gramática do PB, porque o modelo anacrônico adotado não mais se sustenta com a universalização do ensino, quando novas gerações de docentes já não aprendem uma gramática com regras exógenas, tão distantes da sua L1.

2 Um pouco de história social – o Brasil colonial, o contato e o eterno atraso na escolarização

A dissertação de mestrado de Alberto Mussa (1991), defendida no Programa de Pós-Graduação em Letras Vernáculas da UFRJ, sob orientação de Dinah Callou, servirá de ponto de partida para esta seção. Com base em informações de especialistas sobre o tráfico negreiro (historiadores ingleses citados em sua extensa bibliografia, entre outros), o trabalho de Mussa nos fornece informações sobre a distribuição e a classificação dos africanos nas áreas do tráfico bem como a evolução demográfica da população a partir da sua chegada ao solo brasileiro. Embora o começo da ocupação da costa africana tenha se iniciado no século XV, com a distribuição de escravizados para Cabo Verde, São Tomé e Caribe, só a partir de 1538, exatamente quando Portugal se interessa pela terra descoberta em 1500, o Brasil começaria a receber, além de imigrantes portugueses, os africanos escravizados trazidos regularmente de seus portos de embarcação, localizados inicialmente na costa oeste, especialmente na Guiné, Costa do Ouro (atual Gana) e Benin, na zona Equatorial, e em Angola. A "contra-costa", como o autor se refere à região que corresponde a Moçambique, a leste do continenrte africano, seria explorada mais tarde. Com base no material a que Mussa teve acesso,[1] os percentuais de africanos escravizados trazidos ao Brasil ao longo de quase quatro séculos são apresentados na Tabela 1.

Tabela 1. Percentual de escravizados africanos chegados ao Brasil, segundo as zonas de origem e século.

Zona	Séc. XVI 1538-1600	Séc. XVII	Séc. XVIII	Séc. XIX 1801-1850

[1] A tese de Mussa, disponível em www.letras.ufrj-nurc, contou com documentos trazidos da Inglaterra. É possível que estudos posteriores contenham dados adicionais, que possam discordar dos de Mussa, graças ao acesso a mais informações. Entretanto, a consulta a Laurentino Gomes (2019) mostra muitos pontos em comum.

Zone da Guiné (Senegal – Costa do Marfin)	55%	20%	3%	10%
Costa de Ouro – Benin (Gana - Nigeria)	5%	10%	30%	30%
Zona Equatorial (Camarões – Gabão)	5%	5%	3%	10%
Congo – Zona Angolana (Congo – Angola)	**35%**	**60%**	**60%**	**40%**
Costa Sudeste (Moçambique)	--	**5%**	**4%**	**10%**
TOTAL (%)	100%	100%	100%	100%

Fonte: Mussa (1991, p. 141 – Tabela sem título)

Os percentuais estimados (destacados em negrito) revelam que, com raras exceções, os africanos eram originários principalmente de áreas em que as línguas faladas pertenciam ao grande grupo Níger-Congo. Segundo o autor, seria razoável supor que os captores de escravos não os levariam a portos de embarque excessivamente distantes dos pontos onde eram capturados, fosse pelos próprios africanos ou por captores de outras origens. Assim, a Tabela 1, oferece informações preciosas do ponto de vista linguístico. As línguas faladas na zona Congo-Angola e na zona leste eram línguas Bantu, que correspondem a 35%, 65%, 64% e 50%, por século. O autor chama a atenção para o fato de que o predomínio dos escravizados pertença ao grupo das línguas Bantu ajuda a compreender "a anterioridade do registro de itens lexicais dessa família bem como sua maior integração morfológica, graças ao padrão CV, além do seu alcance a um maior número de campos semânticos" (Mussa, 1991, p.146-147). Mussa traz um amplo capítulo que trata de fatores internos, – particularmente os processos fonológicos e algumas implicações sintáticas, em geral relacionadas à concordância nominal e verbal.

Creissels (2000) traz evidências a favor da hipótese de Mussa. Segundo o autor, a maioria das línguas do grupo Níger-Congo, entre as quais estão as línguas Bantu, não distinguem morfologicamente as funções de sujeito e objeto; não têm afixos e adposições. Essas informações, juntamente com aquelas relativas ao contingente de imigrantes portugueses, que também chegariam ao Brasil em sucessivas levas, serão importantes para as interpretações aqui levantadas sobre a aquisição do português no Brasil colonial.

A estimativa do contingente demográfico no Brasil entre os séculos XVI e XIX é apresentada na Tabela 2.

Tabela 2. Contingente populacional no Brasil, estimado entre os séculos XV e XIX

	1583-1600	1601-1700	1701-1800	1801-1850	1851-1890
Africanos	20%	**30%**	**20%**	**12%**	**2%**
Brasileiros afro-descendentes	--	**20%**	**21%**	**19%**	**13%**
Mestiços	--	**10%**	**19%**	**34%**	**42%**
Brasileiros brancos	---	5%	10%	17%	24%
Europeus	30%	25%	22%	14%	17%
Indígenas integrados	50%	10%	8%	4%	2%

Fonte: Adapt. de Mussa (1991, p.163, Tabela sem título)

No século XVI, os indígenas integrados constitutíam a maior parte da população no Brasil (50%), enquanto os africanos respondiam por 20%. A partir de então, a população originária começa a descrescer significativamente à medida que cresce a população constituída por africanos, seus descendentes e mestiços, sempre com índices superiores à população branca de imigrantes portugueses. Segundo Gomes (2019, p.255-256), o número de escravos traficados para o Brasil, 4,9 milhões de cativos (47% do total desembarcado em todo o continente americano), só comparável ao número recebido pelo Caribe, é "infinitamente maior do que o número de europeus que aqui chegaram nesses 350 anos. Estima-se que, até a metade do século XIX, cerca de 750 mil imigrantes portugueses tenham entrado no território brasileiro", o que nos leva a 86% de africanos escravizados e 14% de portugueses (uma estimativa que não se distancia da de Mussa).

Ainda segundo Gomes (2019, p. 228), "o tráfico era um negócio que exigia, principalmente, um cuidadoso trabalho de relacionamento dentro do continente africano com reis e chefes locais, que lutavam e controlavam o fornecimento de cativos em suas respectivas áreas". Lovejoy, citado por Gomes (2019, p. 229), diz que "os participantes africanos do tráfico de escravos incluíam os príncipes e os mercadores mais ricos e poderosos do continente. A elite africana estava profundamente envolvida na venda de escravos". Assim, eram os chefes africanos que controlavam as ofertas, faziam alianças, fechavam negócios, tratavam com diferentes interlocutors europeus, a fim de evitar que um ou outro grupo assumisse o controle do território. A partir do século XVIII, brasileiros acabaram por se juntar aos portugueses no "empreendimento", tornando-se os maiores traficantes de escravos.

Durante três séculos e meio, o crescimento do tráfico e da imigração portuguesa foi movido pelo número crescente de engenhos de acúcar, seguido pela descoberta de minas de ouro e diamantes e o desenvolvimento da agricultura cafeeira, um movimento do Nordeste ao Sudeste, facilitando o contato. Não trazemos aqui informações sobre os escravizados no Sul, mas sabemos da existência de quilombos nos três estados da Região. As tranformações na distribuição populacional têm implicações linguísticas, mas, segundo Mussa, não se pode pensar na existência de variedades de uma *língua franca* africana. O que se pode assumir é que a maior parte das populações africanas acabou aprendendo o português como L2, em diferentes graus, segundo sua posição na esfera social. Boa parte dos escravos já chegava com experiência e algum conhecimento do português (porque, uma vez capturados, tinham de aguardar a chegada dos navios em locais onde se falava o português), o que lhes rendeu posições como capatazes, especialistas nas operações nos engenhos (atividade já exercida em Cabo Verde, na Madeira), na mineração e nos trabalhos domésticos. As sucessivas atividades econômicas acima mencionadas sugerem sua mobilidade interna para diferentes regiões do país, juntamente com a mão de obra portuguesa. Esta é a razão pela qual Mattos e Silva (2002, p. 456) não tem dúvida em afirmar que os africanos e seus descendentes "foram o principal elemento de difusão do português brasileiro", que eles adotaram, não sem introduzir novos traços à língua vernácula, uma contribuição só reconhecida em pesquisas recentes. E, tendo em conta que o Brasil era um país rural, situação que só se inverteria a partir da segunda metade do século XX, podemos supor que o contato entre africanos, afrodescendentes e portugueses tenha sido intenso, a ponto de a língua assumir muitos traços comuns.

Essa hipótese já tinha sustentação em Houaiss (1988 [1985]), para quem o resultado desse duradouro contato foi que a língua falada no Brasil nunca chegou ao estado de uma outra língua, isto é, nunca foi um crioulo. Nosso argumento, segundo o qual o português brasileiro recebeu significativa contribuição africana desde o início da colonização encontra respaldo na observação de Houaiss, de que houve a expansão *de uma variedade vernacular compartilhada pela força de trabalho formada por escravos e portugueses, tanto trabalhadores quanto senhores de terra*:

> Crioulizado ou não (mas sem jamais ter chegado ao estádio de "outra língua", o português emergente no Brasil, já diversificado regionalmente, passa a ser a língua geral e logo a língua comum de uma coletividade que guardava resíduos linguageiros mais ou menos vivos e uma estranha assimetria social; *nos albores do século XIX, afinal de contas, para 3,5 a 4 milhões de habitantes, não haveria mais de 18-20 mil (0,5%) letrados,* que constituíam parte da classe dirigente e dominante, e parte dos "assessores"

> intelectuais dessa classe – como escritores, agregados, escrivães, escreventes, escriturários, correspondentes, comerciantes e o clero. *O analfabetismo ainda grassava entre senhores de terra e seus familiares*." (Houaiss 1988 [1985], p.131) [ênfase acrescida]

A citação acaba por mostrar que durante o período colonial nunca houve preocupação com uma norma brasileira e a alfabetização nunca foi sequer um projeto (cf. Duarte, Gomes e Paiva, 2016). A luta contra o analfabetismo, mesmo depois da independência, quando o Brasil passa de Colônia a Império, em 1822, não mostrou qualquer agilidade. Pagotto (2013) nos diz que a gramática que aparece no texto da Constituição do Império muda no texto da Constituição da República. Eram mudanças ocorridas no português europeu em sua fase moderna, aquela que se mostrava no segundo texto constitucional. A situação se agrava na virada do século XIX, quando a aprovação do primeiro Código Civil brasileiro cai nas mãos da Câmara, tendo como figura proeminente nas discussões o senador Rui Barbosa, que levou alguns anos discutindo não o conteúdo dos artigos, mas sua gramática, como nos conta Pagotto (2011). O alvo das polêmicas eram principalmente questões de regência verbal e de colocação pronominal diferentes das que se praticavam aqui, mudanças que tinham ocorrido na história do PE e não na do PB, o que nos levou à virada do século ainda sem uma decisão sobre que norma adotar no ensino da língua. Como era de esperar, a escolha recaiu sobre a norma do PE moderno. Assim, as gramáticas que viriam a surgir adotavam uma atitude exonormativa, inspiradas na escrita dos grandes escritores portugueses do século XIX, ignorando não só os traços conservadores do português escrito pelos poucos letrados no Brasil, que conservava traços do português arcaico e do português clássico, mas também traços inovadores já atestados na língua falada no Brasil (Nascentes, 1939; Duarte, Gomes. Paiva, 2021).

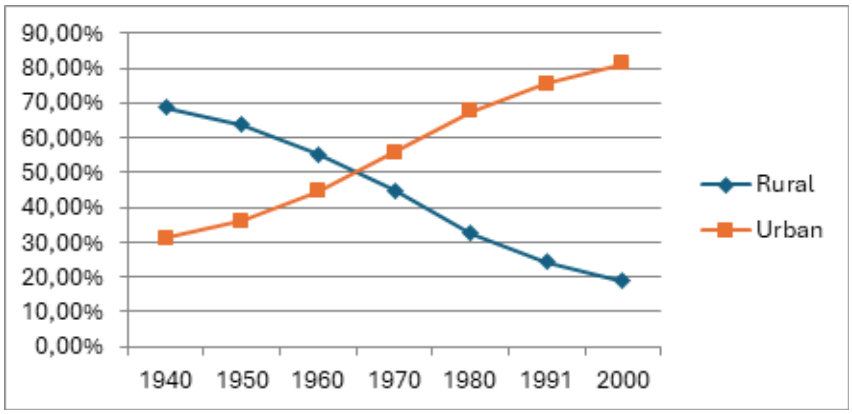
Quanto ao ensino público, segundo Callou e Barbosa (2009), até 1870, não se dispunha de dados mais sitemáticos sobre o estado da instrução no país. As palavras de Calllou e Barbosa são bem ilustrativas da situação paradoxal que resulta das polêmicas em torno da norma no Brasil referidas acima.

> De todo modo [...] é a língua de Portugal que vai gerar um efeito de legitimidade e unidade, pois são os textos escritos nessa língua que criam discursivamente o Brasil. É desse efeito que parece surgir uma relação paradoxal no século XIX quanto ao ensino da língua materna, entre norma e nacionalidade: o princípio de quanto mais próximo da lusofonia e mais distante da alofonia (em competição), mais brasileiro." (Callou e Barbosa, 2009, p. 633)

Concluímos esta seção com fatos sociais que caraterizaram o século XX e interessam à questão da língua do Brasil: os fluxos migratórios internos, que viriam a contribuir ainda mais para o contato entre grupos de diferentes regiões do Brasil em décadas recentes. Com base em dados do IBGE (acesso em 2012), até a primeira metade do século XX, o Brasil era um país predominatemente rural. Os primeiros fluxos migratórios foram particulamente expressivos a partir da década de 1950, das áras rurais dos estados nordestinos para o Sudeste, especialmente para São Paulo e Rio de Janeiro. Migrantes, impulsionados pelo desemprego, seca e latifúndios, e atraídos pela industrialização, partiram em busca de melhores condições de trabalho e subsistência. Um outro fluxo migratório, no final dos anos 1950, dessa vez do Nordeste e Sudeste (especialmente Minas Gerais), ocorreu em direção ao Centro-Oeste, onde a nova capital – Brasília – estava sendo construída, com migrantes atraídos por programas para ocupar a região como força de trabalho na construção de estradas e ferrovias, além de atividades na agricultura. Um terceiro fluxo migratório, do Nordeste em direção ao Norte, especialmente à Amazônia, atrairia migrantes para o trabalho na mineração, na extração de madeira e borracha, ou mesmo para participar de projetos de criação de gado ou ocupando áreas rurais como pequenos fazendeiros. Finalmente, entre 1960 e 1970, ocorreu um fluxo migratório devido à expansão da agricultura, do Sul e Sudeste (São Paulo e Minas Gerais) para os estados do Centro-Oeste e Norte. A maior motivação para essa diáspora de migrantes sulistas estava relacionada a dificuldades na divisão de propriedades herdadas entre membros de fazendeiros de segunda e terceira gerações, com pequenos fazendeiros, que se estabeleciam em novas propriedades, e trabalhadores rurais em busca de emprego na agricultura.

O resultado dessas migrações provoca uma inversão na distribuição da população no Brasil – do campo para as cidades. O êxodo rural, a partir de dados do Instituto Brasileiro de Geografia e Estatística (IBGE), pode ser visto na Figura 1:

Figura 1. Êxodo rural no Brasil em 70 anos



Fonte: IBGE – acesso em 2012 [IBGE | Estatíticas do Século XX (https://seculoXX.ibge.gov.br)

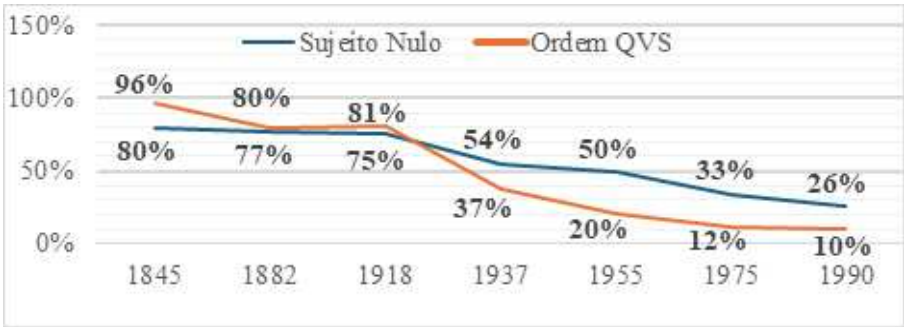
Não é propósito deste trabalho discutir as consequências desse rapído processo de urbanização, como o desemprego, a ampliação das diferenças sociais, a emergência de favelas, a gritante falta de infraestrutura, de assistência social e de acesso à educação. Mas essa sequência de fluxos migratórios representa, sem dúvida, um intenso contato entre todas as variedades linguísticas no país, e pode constituir, até certo ponto, um componente capaz de contribuir para a regularidade atestada em fenômenos morfossintáticos do português brasileiro. Na verdade, as descrições que temos do PB (cf. por exemplo, Martins e Abraçado, 2015, entre muitas outras) não nos permitem traçar fronteiras sintáticas rígidas separando essas variedades. Se a fonologia segmental e suprassegmental nos distinguem, bem como o léxico, nossa morfossintaxe nos aproxima, excetuando alguns poucos traços, como a variação no uso mais ou menos frequente de marcas de concordância verbal e nominal e no uso de tu/você e no sistema de reflexivos, os dois últimos fenômenos de natureza diatópica. Tentativas de proclamar grandes diferenças na sintaxe brasileira – como componente da clivagem social - são baseadas numa visão impressionística ou ideológica da língua, nos termos de Milroy (2000), como mostraremos na Seção 4. As circunstâncias sócio-culturais relacionadas à formação do português brasileiro não nos permitem continuar a tomar o português europeu, que evoluiu desde o descobrimento do Brasil, como modelo para o ensino do português brasileiro na escola na sua modalidade escrita, seguindo a advertência de Faraco (2008). Tal atitude só tem prejudicado o sucesso no ensino de uma gramática da escrita mais próxima da fala do que distante dela. E isso é o que autores brasileiros altamente prestigiados, como Carla Madeira e Itamar Vieira Junior, têm nos mostrado numa produção lietrária primorosa, já trazendo finalmente a seus textos a gramática do Brasil, não na fala do pobre, do "pouco-letrado", mas no texto do próprio narrador. A maioria quase absoluta dos leitores sequer se dá conta de que não está diante daquela gramática que fazia Drummond sofrer, mas diante de uma gramática que usa na sua própria fala.

3. Como o desconhecimento da história social pode "enganar" o pesquisador acerca da gramática de uma língua

Como dito na Introdução, as primeiras análises de Duarte (1992; 1993) sobre a ordem VS/SV nas interrogativas-Q e sobre a realização do sujeito pronominal, respectivamente, foram realizadas com base em um *corpus* de peças de teatro brasileiras, escritas ao longo dos séculos XIX e XX, por autores nascidos no Rio de Janeiro. Os resultados mostravam uma semelhança com a gramática portuguesa, tal como descrita nas gramáti-

cas brasileiras. A partir dos anos 1937, uma mudança na curva das ocorrências sugeria a entrada de traços da gramática do PB oral. Os resultados para os dois fenômenos citados aparecem na Figura 2:

Figura 2. Interrogativas Q-VS (vs Q-SV) e Sujeitos Nulos (vs Sujeitos Expressos) no PB



Fontes: Duarte (1992, p. 41) e Duarte (1993, p.112)

Os altos índices e sua grande proximidade nos três primeiros períodos seguidos por expressiva curva descrescente deu a impressão de que o PB das peças, presumivelmente mais próximas da fala espontânea, tinha uma gramática [+Sujeito Nulo] e [+QVS], que teria mudado a partir do segundo quartel do século XX para uma gramática [-Sujeito Nulo] e [-QVS]. Duarte atribuiu a mudança à entrada de você no paradigma pronominal a partir das peças dos anos 1930 e, mais tarde, à entrada de a gente, reduzindo ainda mais o paradigma flexional. O que os dados sugeriam era que, em um espaço de tempo relativamente muito curto, se considerarmos o período de 1930-40 a 1990, o PB teria passado por duas mudanças drásticas. Uma delas, a ordem QVS, que ocorria com qualquer tipo de verbo (cf. os transitivos em (1a,b)), passando a ficar restrita a verbos inacusativos, entre os quais se inclui a cópula (cf. (1c,d), desde que o argumento interno, o sujeito posposto, seja um DP lexical (2a,b); a ordem VS com um pronome não é atestada. Em (3a-d) vemos a "nova" ordem QSV:

(1) a. Mas que tenho **eu** a temer? (1845)
b. Mas o que tens **tu** ? (1882)
c. Quando é que vem **filme da Kim Novak**? (1992)
d. Onde andará **a Neiva**? (1992)

(2) a. E por que persistem **os erros**?
b. *E por que persistem **eles**?

(3) a. Mas que emprego **eu** posso arranjar? (1955)
b. Onde **você** andou? (1955)
c. Do que **tu** tá falando? (1992)
d. E por que **os erros** persistem? (1975)

Os resultados relativos aos sujeitos pronominais sugerem que teríamos passado igualmente por uma drástica mudança, que nos levaria dos sujeitos nulos como em (4) a sujeitos preferencialmente expressos a partir de meados do século XX (5):

(4) a. Quando **Ø**$_{1ps}$ te **vi** pela primeira vez, **Ø**$_{1ps}$ não **sabia** que **Ø**$_{2ps}$ **eras** viúva e rica. **Ø**$_{1ps}$ Amei-te por simpatia. (1945)
b. **Ø**$_{2ps}$ Terá o cavalo que **Ø**$_{2ps}$ deseja (1918)
c. Já ontem comprei-**lhe**$_i$ o hábito com que **Ø**$_i$ andará vestido. Assim **Ø**$_i$ não estranhará. **Ø**$_i$ Será frade feliz. (1945)

(5) a. Se **eu** ficasse aqui **eu** ia querer ser a madrinha. (1992)
b. **Você** não entende meu coração porque **você** 'tá sempre olhando pro céu e procurando chuva. (1992)
c. T**u** engravidou de novo? (1992)
d. Agora **ele** não vai mais poder dizer as coisas que **ele** queria dizer. (1992)

Curiosamente, o pronome *você*, tal como ocorre no PE, não é atestado nas primeiras sincronias e o pronome *tu*, ora nulo ora expresso, aparece sistematicamente nas peças dos três primeiros períodos. Sua manutenção em peças dos anos 1930 e a entrada de *você* nas peças analisadas já começa a revelar a mistura de tratamento apontada por Duarte (2018 [1993]) e Lopes e Cavalcante (2011, entre outros)[2]. Muitos autores cariocas abandonam o pronome *tu*, usando apenas o pronome *você*, geralmente expresso; quando *tu* retorna às peças cariocas, já aparece em variação com *você*, sem a flexão distintiva, como em (3c) e (5c) acima. Em 2012, Pinheiro e Marins retomaram a pesquisa de Duarte (1992) incluindo os padrões QV (com o sujeito nulo), com resultados que reforçam a impressão de que, de fato, o padrão QV é predominante nas peças do PB dos três primeiros períodos,

[2] Essa mistura de tratamento é um bom argumento a favor da hipótese de Faraco (1996: p. 64,65) , segundo o qual a razão para um "uso tão amplo de *você* no Brasil deve ser encontrada, certamente, na história da formação do País" [...] "podemos pressupor que, desde o início da ocupação européia do Brasil, as formas predominantemente de tratamento do interlocutor eram as diferentes variantes de Vossa Mercê". Como veremos adiante, um paradigma com muitas flexões não seria adquirido pela população em contato no Brasil colonial e o uso das variantes que viram a originar o pronome *você* deveriam estar presentes na formação do PB.

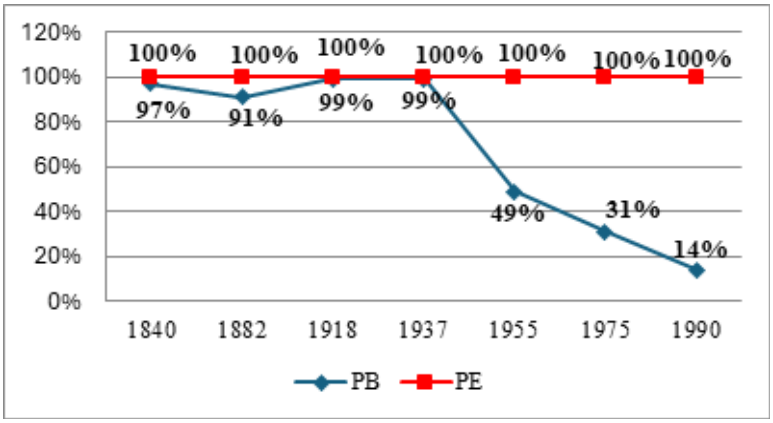
seguido por QVS e QSV, este em índices baixíssimos; ainda sem uma comparação com o PE, a conclusão leva a considerar o PB uma gramática de sujeitos QV ou QVS, para, a partir de então, mostrar a perda dos dois padrões, chegando a ordem QSV a 76% no último período analisado, como veremos adiante, uma mudança drástica.

Foram essas análises que levaram Duarte a julgar que os resultados para o século XIX e primeira metade do XX exibiam a gramática do PB, idêntica à do PE. Na ocasião, Duarte não levou em conta o analfabetismo reinante no Brasil nem o fato de que a escrita era privilégio de poucos, aquela minoria que aprendia a gramática lusitana, em Coimbra, ou poucas instituições privadas no Brasil (ver Houaiss na Seção 2). E, mesmo depois de iniciadas as análises contrastivas, com peças portuguesas (comédias de costumes, farsas e textos leves que seguiriam os costumes de seu tempo), escritas principalmente por lisboetas e distribuídas pelos mesmos períodos, levaria algum tempo para que essa interpretação equivocada quanto aos primeiros períodos fosse revista.

Entre as primeiras análises contrastivas (as fontes históricas são mencionadas antes das referências) estão a tese de Marins (2013), sobre os verbos *haver* e *ter* existenciais, a dissertação de Mourão (2015), que comparou os resultados para o sujeito pronominal no PE com a análise de Duarte (1993) para o PB, a tese de Nicolau de Paula (2016), que comparou as interrogativas-Q com as análises para as peças brasileiras, e, finalmente as dissertações de Cardoso (2017), Marques de Sousa (2017) e Fernandes (2018), descrevendo, respectivamente, o uso do clítico indefinido *se*, e dos clíticos acusativo *os(as)* e dativo *lhe(s)* para referência à 3a pessoa e suas formas variantes.

A continuação desta seção, embora pareça excessiva, precisa da paciência do leitor para que observe a semelhança nos resultados entre PE e a PB nas primeiras sincronias e a mesma curva descendente do PB observada na Figura 2 acima. Comecemos pela análise de Marins (2013) para manter a cronologia dos trabalhos contrastivos.

Figura 3. Uso de *haver* (vs *ter* existencial) no PE e PB



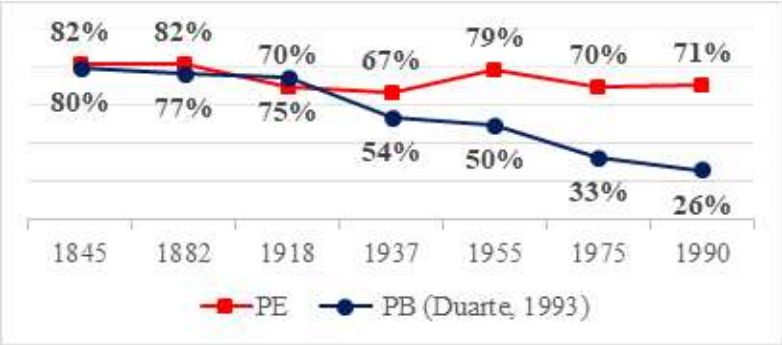
Fonte: Adapt. de Marins (2013, Fig. 9, p. 79)[3]

[3] Os anos que aparecem no eixo horizontal representam o período em torno do qual as peças foram escritas.

A Figura adaptada de Marins, exclui os raros usos do verbo "substantivo" *existir* em oposição ao verbo "funcional" *haver* para o PE e *ter* para o PB (cf. Avelar, 2016). Não é necessário perguntar por que autores brasileiros tão consistentes no uso de *haver* nos quatro primeiros períodos, quase conseguindo chegar ao uso categórico do PE, acabam mostrando nas gerações seguintes, quando a escolarização se expande, seguindo as normas portuguesas, uma expressiva curva, que significa a opção pelo verbo *ter* existencial. A percepção de que esses primeiros períodos *não eram instâncias da gramática do PB* só aconteceria exatamente em 2023,[4] quando Duarte, ao preparar uma conferência em homenagem a Mary Kato, publicada em 2024, reuniu as três análises sobre os clíticos, mencionadas acima, que serão apresentadas nas figuras 7 e 8.

A análise seguinte, de Mourão (2015), compara os resultados de Duarte (1993) para o sujeito pronominal com a amostra portuguesa, como se vê na Figura 4.

Figura 4. Distribuição geral dos sujeitos nulos (vs. plenos) ao longo do tempo no PE e PB



Fonte: Mourão (2015, Fig. 5, p. 69)

Mais uma vez, vemos a proximidade de PE e PB, considerando as três pessoas do discurso, ilustradas em (4) e (5) acima para o PB. Não é necessário reiterar o que foi dito em relação à figura anterior. Só a partir dos anos 1930, o PB começa a se mostrar, com a curva descendente, revelando a sua preferência pelos sujeitos expressos. A análise de cada pessoa do discurso coloca a 1ª e a 2ª pessoas mais avançadas nas taxas de sujeito

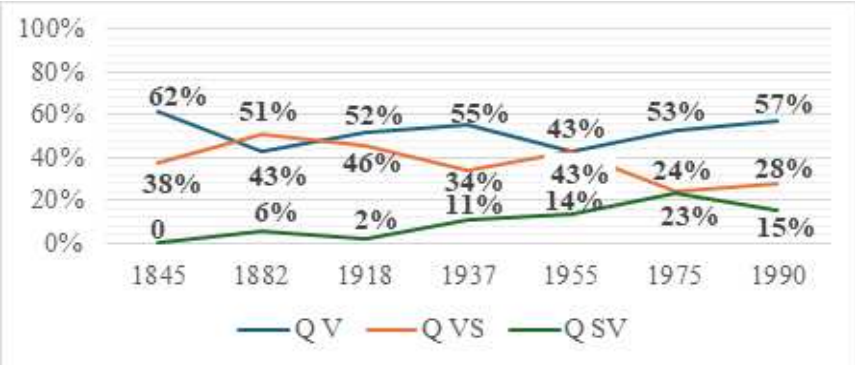
Cada trabalho aqui citado mostra as datas que correspondem aos limites dos períodos que incluem as peças analisadas.

[4] Exatamente em 2023, uma aluna de Marins, ao fazer sua monografia de conclusão de curso de graduação, com base na tese de sua orientadora (Marins, 2013), se perguntou: "Há quanto tempo o português é brasileiro?" (Mariante, 2023). A graduanda enxergou o óbvio, de maneira muito rápida. Seria porque se trata de uma jovem, que adquiriu o verbo *ter* existencial, e não se deixou convencer por resultados tão eloquentes?

preenchido em relação ao PE, que mantém sujeitos nulos em todas as pessoas independentemente do padrão sentencial (a correferência entre sujeito e seu antecedente, com ou sem c-comando, o antecedente em sentença adjacente com a mesma função, ou mesmo um antecedente com função diferente e/ou distante). Por limite de espaço, encaminho o leitor para a análise contrastiva de Mourão (2015) e para a análise contrastiva dos sujeitos de 3ª pessoa no PE e PB orais de Duarte (2019), que mostra as taxas de sujeitos expressos no PB sempre superiores e os contextos em que um sujeito nulo já não exibe índices significativos; quanto ao PE, fica evidente seu estatuto de língua de sujeito nulo consistente.

Acompanhando as análises pelo ano de defesa dos trabalhos, vemos a seguir os resultados encontrados por Nicolau de Paula (2006), que expandiu a amostra e comparou os três padrões de interrogativas-Q do PB (seguindo Pinheiro e Marins, 2012, referidos nesta seção) com as peças portuguesas. A Figura 5 mostra os resultados para o PE:

Figura 5. Padrões de interrogativas Q ao longo de sete períodos no PE

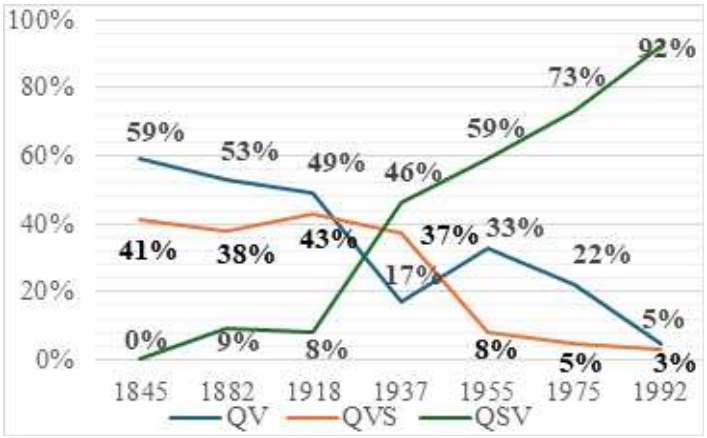


Fonte: Nicolau de Paula (2016, Fig. 4.1, p. 72)

A preferência do PE pelo padrão QV, mantendo forte o sujeito nulo, é seguida de perto pelo padrão QVS. Quanto à ordem QSV no PE, esta aparece como uma estratégia tímida, que só revela pequeno crescimento nos dois períodos finais,[5] desde que a interrogativa tenha a presença da clivagem. A comparação dos padrões lusitanos com os brasileiros na Figura 6 não surpreenderá o leitor:

[5] Nicolau de Paula (2016) atesta que no PE oral as interrogativas-Q preferem o sujeito nulo em todas as pessoas do discurso, sendo muito frequente a clivagem em todos os padrões, que é, porém, obrigatória quando a ordem QSV é usada. A autora usou como corpus para a segunda pessoa as perguntas dos entrevistadores portugueses e brasileiros em diversas amostras de entrevistas.

Figura 6. Padrões de interrogativas-Q ao longo de sete períodos do PB



Fonte: Nicolau de Paula (2016, Fig. 44, p. 93)

A concorrência entre os padrões QV e QVS nos três períodos iniciais, quando comparamos as Figura 5 e 6, não deixa dúvida de que os letrados desses períodos seguiam um modelo quase com perfeição. O que se vê a partir dos anos 1930 é o decréscimo no PB dos padrões preferenciais lusitanos, que se mantêm estáveis, e a subida expressiva e consistente da ordem QSV, que vai alcançar 92% nos anos 1990 da amostra analisada para o PB. Chegamos aos anos 1990, com 5% de sujeitos nulos (QV) e 3% de QVS.

Passemos às análises que contemplam os clíticos não dêiticos: o indefinido e o acusativo e dativo anafóricos.

A análise de Cardoso (2017) leva em conta o fato de que o clítico indefinido, geralmente apontado pelas gramáticas como uma das estratégias de indeterminação do agente juntamente com o verbo na 3ª pessoa do plural e outras variantes ainda não incorporadas à descrição gramatical que orienta o ensino não constituem um conjunto com formas em variação; na verdade temos dois conjuntos: o que veicula referência arbitrária e o que veicula referência genérica (Egerland, 2003). A referência arbitrária exclui o falante e seu interlocutor, e, em geral, alcança um número bem restrito de pessoas, podendo ser representada pelo clítico *se* em variação com o verbo na 3ª pessoa do plural, com o sujeito nulo no PE e em variação com o sujeito nulo ou expresso no PB, como ilustramos em (6) e (7) para ambas as variedades, respectivamente:

(6) a. Se ao menos por cá houvesse a moda que há na bela França, aonde **Ø$_{arb}$ se faz** caldo para toda a semana [...] (PE – *Uma cena de nossos dias,* Paulo Midosi, 1843)

b. Pois se **Ø**$_{arb}$ o **fizerem** deputado, é capaz de vender a pátria como vendia a justiça. (PE - *A liberdade eleitoral*, Teixeira de Vasconcellos, 1870)

(7) a. Já em Nicéia Ø$_{arb}$ **reuniu-se** um concílio para decidir se a mulher devia ou não fazer parte do gênero humano. (PB – *As doutoras*, França Júnior, 1889)
b. Imagine a senhora que Ø$_{arb}$ **confundiram** o Adalberto, meu cunhado, com terrorista procuradíssimo. (PB – *Como encher um biquíni selvagem*, Miguel Falabella, 1992)

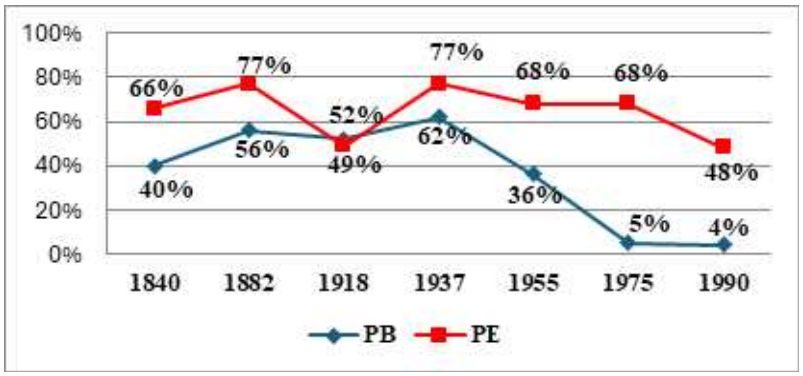
Por outro lado, temos a referência genérica (representada pelo mesmo clítico *se* e por formas nominativas, como *você/tu*, ou mesmo como *nós/a gente*), que pode incluir o falante e seu interlocutor, além de alcançar um número infinito de pessoas, ilustradas em (8) e (9) para PE e PB, respectivamente:

(8) a. Quando Ø$_{gen}$ **se gosta** de alguém como eu não é difícil ter coragem. (PE- *A menina feia*, Manuel Frederico Pressler, 1970)
b. Não diga mal dos doidos. Todos os homens que fizeram alguma coisa no mundo eram doidos. Ø$_{gen}$ **Devemos**-lhes a vida artificial. Na realidade Ø$_{gen}$ **devemos**-lhes tudo. Se não fossem eles ainda hoje Ø$_{gen\ i}$ **seríamos** bichos. (PE - *O doido e a morte*, Raul Brandão, 1923)
c. A vida não pára. E **a gente**$_{i}$ não **pode** também **parar**, **Ø**$_{gen\ i}$ não pode voltar-lhe as costas... **Ø**$_{gen\ i}$ tem de acompanhá-la, quer **Ø**$_{gen\ i}$ queira, quer não. (PE - É urgente o amor, Luiz Francisco Rebello, 1957)

(9) a. É verdade, Ø$_{gen}$ nunca **se soube** do paradeiro dele. (PB - *O colar de coral*, Antonio Callado, 1954)
b. O segredo é Ø$_{gen}$ não se importar. Se **você** não se importar, **você** acaba dormindo. Mas **você** precisa não se importar de verdade. Aí, **você** não se importa se matam crianças para vender órgãos para países mais ricos, **você** simplesmente não se importa. (PB – Como encher um biquíni selvagem, Miguel Falabella, 1992)
c. As coisas são como são e não como **nós queremos** que elas sejam. (PB - *O simpático Jeremias*, Gastão Tojeiro, 1918)
d. É essa a minha intenção. Mas nem tudo que **a gente deseja** pode ser alcançado (PB - *O hóspede do quarto nº 2*, Armando Gonzaga, 1937)

Por limitação de espaço, a Figura 7 reúne as ocorrências do clítico indefinido *se* (arbitrário e genérico) versus as formas nominativas variantes acima ilustradas (em Cardoso, 2017, as análises contemplam cada tipo de indeterminação e suas variantes à parte).

Figura 7. Clítico indefinido (arbitrário e genérico) (vs. formas variantes) no PE e no PB



Fonte: Cardoso, (2017, Fig. 4.7, p. 89)[6]

A linha vermelha, que representa o uso do clítico *se* no PE, mostra uma variação estável com dois picos, que se devem à concorrência de uma das variantes do *se* arbitrário: a 3ª pessoa do plural, variante igualmente padrão no PE e mais frequente do que o clítico *se* para a expressão da referência genérica. A linha azul, para o clítico no PB, segue de perto o clítico no PE nos quatro primeiros períodos. Só na segunda metade do século XX, vemos a linha descendente com 36%, 5% e 4%, com o *se* arbitrário sendo substituído ou pela terceira pessoa do plural com o pronome *eles* nulo ou expresso, e o *se* genérico, substituído pelas formas nominativas ilustradas em (9b,c,d), preferencialmente expressas no

[6] A análise não atestou o sujeito nulo arbitrário com o verbo na 3ª pessoa do singular, apontado por Lunguinho e Medeiros (2009), mas é encontrada no PB oral, em que seu uso fica bem distante da 3ª pessoa do plural, a estratégia preferida. Duarte e Marins (2021) atestam 376 casos de 3ªpp com o pronome nulo ou expresso (91,3%); 19 casos do nulo de 3ªps (4,6%) e 17 casos de *se* (4,1%), como mostra (i):
(i) Na televisão $[\mathbf{Ø}]_{arb}$ **fala** muito isso. (Marins e Duarte, 2021)
A mesma estrutura com a referência genérica, apontada por Galves (1987) – "$Ø_{gen}$ **Não usa** mais saia" - , apareceu numa única ocorrência, na última sincronia da amostra PB:
(ii) $Ø_{gen}$ **Tem que aprender** a mexer com aqueles ferrinhos, mas o salário parece que compensa. (PB - *No coração do Brasil*, Miguel Falabella, 1992)
No PB oral o nulo genérico alcança 3,7% (57 casos), o clítico 3,2% (50 casos) e você, 93,1%) em 1.455 dados (Duarte e Marins, 2021). Nota-se que ambas as estruturas com o nulo arbitrário e genérico aparecem em construções que veiculam aspecto (contínuo ou pontual) e modalidade (deôntica ou epistêmica) (como é o caso de (i) e (ii) acima, que podem exibir a categoria vazia em variação com o pronome de 3ª pp nulo ou expresso para referência arbitrária (i)' ou por qualquer uma das estratégias genéricas (ii)':
(i)' Na televisão $[\mathbf{Ø}]_{arb}$ /**eles falam** muito isso.
(ii)' **Você/ a gente tem que aprender** a mexer com aqueles ferrinhos.

PB. Note-se que o uso de *você* (ou *tu*), a depender de variação diatópica, se mostra como a forma preferencial para a indeterminação genérica no PB, tanto nas peças do século XX, como mostra Cardoso (2017), quanto na fala (cf. Duarte e Marins, 2021 e nota 6).

Passemos ao clítico acusativo[7] e suas variantes, o objeto nulo e o pronome nominativo pelo acusativo, atestado apenas no PB[8], ilustrados em (10) e (11) para PE e PB.

(10) a. Viste [**a figueira que plantaste atrás da casa**]? Vamos vê-[**la**]? (PE – *O rosto levantado*, Noberto Ávila, 1977)

b. Estás a gozar, como é evidente. **Essa canção** nunca existiu. **Isso** é **uma aberração**. Onde é que já se viu **uma cantiga assim**? Vocês conhecem [**Ø**]? (PE – *Um filho*, Luísa Costa Gomes, 1996)

c. Heitor – [**Os pianos Sakamura**]$_i$ já vão amestrados. O cliente compra [**Ø**]$_i$, chega a casa, e o piano toca. (PE – *Um filho,* Luisa Costa Gomes, 1996)

(11) a. – [**Os meirinhos**]$_i$ entrarão aqui e hão de levar por força alguma coisa – esse é o seu costume. O que é preciso é enganá-[**los**]$_i$. (PB – *O noviço,* Martins Pena, 1845)

b. – Enquanto eu fui comprar cigarros, [**Rosa**]$_i$ subiu pelo elevador e o elevador enguiçou com **ela**$_i$ lá dentro. Você quer ver se o Nanico tira [**ela**]$_i$ de lá? (PB – *Um elefante no caos,* Millôr Fernandes, 1955)

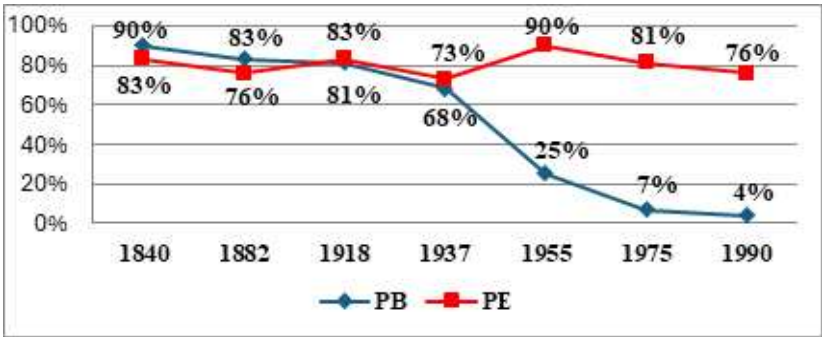
c. – Eu tenho pra mim que [**esse navio**]$_i$ já partiu há muito tempo, ó, e a tonta da Margareth tá esperando [**Ø**]$_i$ no cais. (PB – *No coração do Brasil,* Miguel Falabella, 1992)

Na Figura 8, temos os resultados para a análise de Marques de Sousa (2017), apontando apenas o clítico (versus suas formas variantes).

[7] Para uma análise mais detalhada dos clíticos acusativo e dativo, remeto o leitor a Duarte (2024).

[8] A análise de Marques de Sousa (2017) inclui o SN anafórico ou um epíteto que o retoma, ilustrados em (i) e (ii), mas, considerando que se trata de uma estratégia mais geral entre as línguas, o autor acaba por excluí-las na seção final da Dissertação:
(i)Então, temos aqui [**um bastardo**]$_i$**?** E ela vai de mandar **o** [**rapazola**]$_i$ para a grande cidade, a Casa dos Bons dos Rossetti? (PE – *Um filho*, Luísa Costa Gomes, 1996)
(ii)Enquanto aqui estiverem hospedados [**os americanos**]$_i$ ninguém quer ser criado desta pensão. A cidade inteira está farta de saber que não se pode aturar [**essa gente**]$_i$ (PB – *O simpático Jeremias*, Gastão Tojeiro, 1918)

Figura 8. Clítico acusativo de 3ª pessoa (vs. formas variantes) no PE e no PB



Fonte: Marques de Sousa (2017, Fig. 11, p. 100)

No PE, o clítico acusativo se mostra estável ao longo do tempo, com índices modestos do objeto nulo, já presente na história do português (Cyrino, 1994). O PB, por outro lado, se mostra praticamente idêntico ao PE no uso do clítico até os anos 1930, para iniciar uma curva descendente na segunda metade do século XX, chegando a 4% de ocorrências nos anos 1990, tal como atestamos para o clítico indefinido.

Finalmente, vejamos as ocorrências do clítico dativo anafórico e suas variantes nas peças portuguesas e brasileiras, ilustradas em (12) e (13):

(12) a. Pois esteve com [**o ministro**]$_i$? Falou-**lhe**$_i$ de mim? (PE – Casar *ou meter freira*, Antonio Pedro Lopes de Mendonça1848)

b. – Sabe que está a chegar [**o Pantaleão de Sá Repolho**]$_i$, Morgado das Barrocas, e promettido esposo de sua filha?

– Sei sim, e até [**para elle**]$_i$ mandei apromptar esses dois quartos, que dão sobre o cáes. (PE – *Os logros numa hospedaria*, Paulo Midosi, 1841)

c. Para mim foi a primeira vez. E para **ele**$_i$, olha que não sei. Tinha um facies que não enganava muito. E então eu disse [**Ø**]$_i$:"bem, apanhou-me. Mas agora vamos considerar o problema de outro ponto de vista", e ele disse: "Pois, o meu ponto de vista. Tome lá a multa." (PE – *Um filho,* Luísa Costa Gomes,1996)

(13) a. **Ela**$_i$ foi com suas parentas assistir a um filme e, na recepção, uma senhora muito delicada **lhe**$_i$ ofereceu um bombom. (*No coração do Brasil,* Miguel Falabella, 1992)

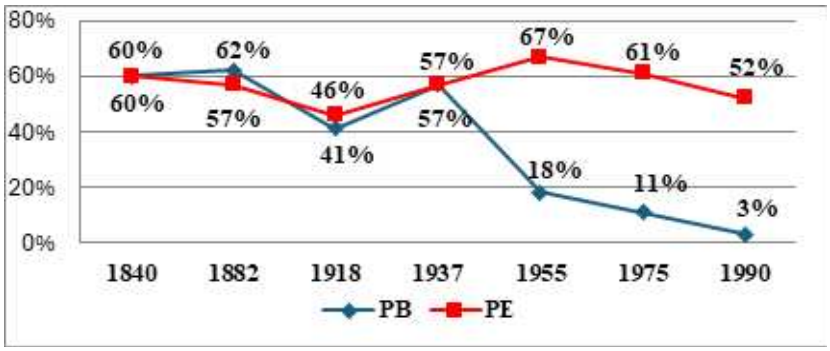
b. Quando você enfrenta [**um espelho**]$_i$, o que é que você pergunta [**a ele**]$_i$? (PB –*A mulher integral,* Carlos Eduardo Novaes, 1975)

c. É isso mesmo, Cristina, [**seu marido**]$_i$ gosta de ser servido por você, não pela sua empregada. Faz uma salada [**pra ele**]$_i$. (PB – *A mulher integral,* Carlos Eduardo Novaes, 1975)

d. Se for [**seu filho**]$_i$, abra a porta naturalmente e não **diga** nada **Ø**$_i$**.** (PB – *Um elefante no caos,* Millôr Fernandes, 1955)

e**. Ela**$_i$ me amou como se eu fosse o último homem sobre a Terra. Ou o primeiro. E só depois **eu perguntei Ø**$_i$ quem era. E ela me disse. (PB – *Um elefante no caos, Millôr Fernandes, 1955*)

Figura 9. Clítico dativo de 3ª pessoa (vs. formas variantes) no PE e no PB



Fonte: Fernandes (2018, Fig. 5, p.78).

Mais uma vez, o uso do clítico dativo comparado com o acusativo revela a mesma semelhança entre PE e PB, até os anos 1930. Da mesma forma, o clítico se mostra estável no PB ao longo de todos os períodos enquanto no PB ele segue a mesma queda vertiginosa em direção à sua quase completa extinção no final do século XX.

Temos evidências de que o comportamento muito próximo do PE e do PB quanto aos fenômenos aqui analisados até a primeira metade do século XX não deixa dúvidas de que eles não foram adquiridos no PB como L1, mas aprendidos por uma minoria letrada no início do século XIX, como menciona Houaiss. Não faria sentido acreditar que esses traços seriam perdidos ao longo da segunda metade do século XX, se adquiridos como L1, justamente quando começa a escolarização no Brasil seguindo o modelo lusitano. Não seria plausível que um processo de escolarização tardio, que mal conseguiu ser posto em prática ao longo do século XIX, como discutiremos na seção seguinte, ainda que inspirado no modelo lusitano, levaria à perda dos traços aqui mostrados. Pelo contrário, só os reforçaria, se tivessem sido adquiridos como L1.

Considerando que os clíticos não dêiticos são marcas de concordância (Galves,

2019; Nunes, 2019; 2022), podemos considerar que sua não aquisição em muito se aproxima da não aquisição das flexões verbais número-pessoais e das flexões nominais de número (e, em raros, casos de gênero). Tal como se pode concluir em relação aos clíticos, podemos presumir que o sujeito nulo nunca deve ter sido predominante num país de analfabetos, com um paradigma flexional possivelmente bastante simplificado.

4. Desvendando o português brasileiro – as primeiras entrevistas sociolinguísticas

Os resultados da história social apresentados na Seção 3 e os dados das análises contrativas entre o PE e o PB nos levam a fatos incontestáveis sobre o lento processo de alfabetização no Brasil. Como vimos na Seção 2, o Brasil entra no século XIX com 05% de sua população letrada. É certo que havia uma escrita, privilégio de poucos, que continuava o curso do português arcaico na fase que chamamos português clássico a partir dos anos 1500. Ao longo do século XIX, tivemos dois momentos históricos importantes – a passagem da colônia a Brasil-Império em 1882 e a Brasil-República em 1889. É então que a questão da alfabetização começa a preocupar. Mas, como vimos, pouco se fez ao longo do século nesse sentido e, na virada do século, ainda se discutia que norma seguir.

Enfim, o Brasil começa o acesso à alfabetização com um programa que contraria a gramática desenvolvida nos 350 anos de colonização e em mais cem anos de absoluta falta de iniciativa. E a questão que nunca se colocou é: qual seria a gramática da fala dos 95% de analfabetos do início do século XIX, mencionados por Houaiss, ao final desse mesmo século? É certo que os grupos escolares tiveram importante papel no processo de alfabetização. Mas a escola pública era restrita ao curso primário e o acesso não era universal. Segundo Rocha e Cunha Júnior (2022), o ensino secundário, que incluía o antigo ginasial (hoje Fundamental II) e o Científico (hoje Ensino Médio), teve uma expansão entre 1931-1961. No entanto, o ensino se concentrava na iniciativa privada, com alguma subvenção do Estado, através de bolsas de estudo. Os autores ressaltam, entretanto, que o ensino secundário não era gratuito, mas obrigatório naquele momento histórico; havia ainda a exigência do exame de admissão (para ingresso ao Ginasial), o que dificultava o acesso da classe trabalhadora à educação.

O fato é que os primeiros registros da fala urbana vieram de dois projetos pioneiros, com material gravado de forma sistemática e criados com propósitos diferentes, nos permitindo conhecer a face do português brasileiro, aquele já adquirido como L1. O Projeto Movimento Brasileiro de Alfabetização (Mobral) foi formalizado em 1968, mas iniciado em 1971, no contexto do regime militar no Brasil, quando o governo passa a controlar os

programas de alfabetização de forma centralizada (Menezes e Santos, 2001). Planejado duas décadas antes e derivado da reflexão e do debate em torno do analfabetismo no país, seus objetivos convergiam no sentido da consolidação de um novo modelo pedagógico, que via no analfabetismo o resultado de uma situação de pobreza gerada por uma estrutura social desigual. O plano, segundo Menezes e Santos (2001), que consistia em examinar a realidade dos educandos, identificar seus problemas e superá-los, foi interrompido pelo regime e esvaziado do senso crítico subjacente. Não alterou as bases do analfabetismo, calcadas fundamentalmente na estrutura organizacional da educação no país, com um modelo discutível em relação a uma proposta pedagógica cuja preocupação principal era ensinar a ler e a escrever, sem nenhuma relação com a formação do homem.

Uma importante contribuição advinda do Mobral se deve a Lemle e Naro (1987), que desenvolveram o projeto Competências Básicas do Português, com a Fundação Mobral e a Fundação Ford. Com a participação de grande equipe, foram entrevistados 20 falantes nos postos do Mobral, situados em Nova Iguaçu e Caxias, no Rio de Janeiro. O projeto traz contribuição para a descrição e análise de um dos traços que mais discriminam o falante, sem dúvida, um traço desenvolvido desde o início do contato no período colonial, que se estendeu por cerca de 350 anos.

Com base na Amostra Mobral, foram defendidas dissertações de mestrado na PUC-RJ, sob a orientação de Naro, sobre a ausência de relativas "padrão" em funções preposicionadas (Mollica, 1977), sobre variação no uso das marcas de concordância nominal (Scherre, 1978 e Braga, 1978), e a ausência do clítico acusativo (Omena, 1978). A dissertação de Omena, que viria a ser o ponto de partida para o estudo de um fenômeno que caracterizava a não aquisição deste e de outros clíticos não dêiticos no PB, em inúmeras investigações que se seguiriam, mostrando a semelhança entre a fala de letrados e não letrados em todo o território nacional.

Ao mesmo tempo em que se desenvolviam análises com base na Amostra Mobral, já com a ida de Anthony Naro para a UFRJ, tinha início no Brasil o primeiro Projeto que viria a nos fornecer um retrato da fala culta brasileira – Norma Urbana Culta (NURC), fundado em 1969 e iniciado na década de 1970, com gravações de falantes das cinco capitais mais populosas de então: Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre. A amostra compreendia três tipos de inquéritos das cinco capitais: fala entre o entrevistador e o entrevistado, fala entre dois entrevistados e elocuções formais, que são, em geral, aulas gravadas em universidades. O objetivo consistia em coletar sistematicamente material que permitisse a análise do português brasileiro "culto", ou seja, com formação

universitária, em seus diversos níveis: fonologia, morfologia, sintaxe, léxico e texto.

Um dos volumes que consolidam anos de trabalho desenvolvido pelo Projeto NURC, com base num recorte da amostra com os três tipos de inquéritos recolhidos em cada capital, o volume sobre sintaxe (Kato e Nascimento, 2015 [2007]) mostra que o PB culto **não** é uma língua de sujeito nulo (todas as pessoas apresentam índices abaixo de 20% de sujeitos nulos), **não** prefere o verbo *haver* existencial (que oscila entre 10% e 35%, só alcançando 59% nas "elocuções formais", ou seja, nas aulas gravadas), **não** usa os clíticos não dêiticos (indefinido, acusativo e dativo) com consistência: o indefinido *se* alcança 13,5%, o acusativo, 12%, o dativo está completamente ausente da amostra compartilhada que foi analisada. Para os sujeitos indeterminados (arbitrários e genéricos) são preferidos os pronomes nominativos, e para os complementos, lideram os objetos diretos nulos, os SPs anafóricos e complementos nulos para o objeto indireto. Quanto aos complementos oblíquos, excluídos os objetos indiretos (dativos), em 67 ocorrências de verbos transitivos como *gostar, precisar,* 40% têm seu complemento anafórico nulo; o mesmo ocorre com os complementos oblíquos circunstanciais de verbos como *ir, pôr, colocar*: dos 27 dados computados, em 21 casos, o complemento já mencionado no contexto precedente não se manifesta foneticamente.

O Projeto NURC acabaria por revelar que, excetuando um uso mais consistente de marcas de concordância verbal e nominal (Graciosa, 1991), não havia diferenças significativas entre os resultados para a fala não culta e a culta no que se referia à morfossintaxe. É o que mostram os resultados obtidos para pesquisas que viriam a ser desenvolvidas em todo o país, por outros projetos, a começar pelo Projeto PEUL (Programa de Estudos sobre o Uso da Língua), coordenado por Anthony Naro e seus orientandos, na UFRJ. A amostra PEUL gravada nos anos 1980, conhecida como Amostra Censo, com falantes da comunidade de fala carioca, com os níveis Fundamental 1 e 2 e Ensino Médio, serviria de inspiração para outros projetos que se estenderiam pelo país, como o ValPB, no Nordeste, e VarSul, na região Sul, entre outros tantos. Essas amostras de fala viriam a cobrir, com o Projeto NURC, todos os níveis de escolaridade no Brasil, nos trazendo uma fotografia sociolinguística do português brasileiro. Em Paiva e Duarte (2003), Martins e Abraçado (2015) temos resultados que confirmam a regularidade na distribuição dos fenômenos morfossintáticos no Brasil, lembrando sempre que apenas três fenômenos podem nos distinguir – a frequência maior ou menor do uso de marcas de concordância verbal e nominal, a diferença no tratamento tu/você e o sistema de clíticos reflexivos, esses dois últimos de distribuição diatópica.

As diferenças que os resultados efetivamente mostram se situam no sistema fonético-fonológico, tanto no nível segmental quanto suprassegmental, e no nível lexical, que nos identificam e nos localizam diatopicamente, sem dúvida, um campo de estudos que tem merecido destaque na pesquisa acadêmica brasileira.

Com base em tais resultados, não se pode sustentar a ideologia de que a morfossintaxe brasileira exibe uma profunda clivagem social. Uma curiosidade a esse respeito: duas pesquisas realizadas em Florianópolis, comunidade descendente de açorianos, obtêm os percentuais mais baixos de clítico acusativo nas análises de Luíze (1997), 1,6%, e Vieira Pinto (2015), 0,3%. Os resultados para esse fenômeno, na verdade, nunca ultrapassam os 5%. O clítico dativo anafórico não apareceu nas pesquisas realizadas em diferentes pontos do Brasil, relatadas igualmente em Martins e Abraçado (2015).

Na verdade, costumamos confundir a escrita do letrado com sua fala espontânea. A esse respeito, vale a pena lembrar como Kato (2005) lança mão da proposta de Kroch (2001), acerca do conceito de "competição de gramáticas", embora o utilize para tratar não da mudança linguística mas da gramática do letrado brasileiro. Segundo ela, a criança chega à escola com sua gramática adquirida, sua L1, e entra em contato com uma outra gramática (é bom levar em conta os casos em que essa outra gramática da escrita pode ser mais próxima ou mais distante da L1 da criança – a depender do contexto social que a cerca, da sua maior ou menor exposição à leitura pelos que com ela convivem, entre outros fatos; e aqui, sim, podemos dizer que no Brasil, há uma clivagem entre a gramática adquirida e a da escola). O resultado, segundo Kato é que a gramática do letrado brasileiro contém traços da sua L1, traços de uma L2, que vêm através principalmente do contato com uma gramática de gerações passadas, que podem ser aprendidos apenas parcialmente; além disso, essa gramática pode conter "arbitrariedades", ou seja, traços que não estão presentes nem numa nem noutra gramática, mas frutos exatamente desse descompasso entre a L1 e gramática da escola. Basta ler em matérias de autores letradíssimos, em jornais de ampla circulação, casos de ênclises (para não fugir ao tema mais caro aos gramáticos tradicionais), que não são nem da gramática do PE nem do PB, pela simples razão de que este é ponto de honra no ensino e a falta de intuição leva o falante brasileiro, que tem na próclise a posição natural do clítico, a fazer uma ênclise com atrator, contrariando sua L1, que é proclítica e a gramática do PE, que usaria a próclise por conta dos "atratores" (como em *Também falou-se, Até deu-me*), com abundante ocorrência nos jornais dirigidos ao público A, ou ainda num contexto em que a ênclise não ocorre em qualquer língua que tenha clíticos (*o aluno me disse que tinha entregado-**o** no prazo*).

De todo modo, o "modelo anacrônico" que ainda orienta os livros didáticos e só prejudica o ensino, como sempre insistiu Faraco (2008, p. 47) vai, aos poucos, se modificando porque já não é dominado nem pelas gerações mais novas de docentes. O mesmo autor (Faraco, 2018, p. 47) chama a atenção para o fato de que com a internet "acervos quase infinitos de texto estão ao alcance direto dos internautas. Ao mesmo tempo, tem-se a impressão de que nunca antes tantos escreveram tanto como nos meios eletrônicos (textos, porém, nem sempre padronizados). [...] Com isso as mais diferentes variedades linguísticas se tornaram audíveis, o que não ocorre com outros meios de comunicação". O autor termina seu artigo perguntando: "Que resultados trará a ampla circulação da escrita não padronizada? Que resultados decorrerão da abertura da internet às mais diversas variedades da língua? Só o futuro dirá".

Eu ouso dizer que as legendas de filmes constituem uma evidência de que o português adquirido no Brasil já vai ganhando o espaço antes reservado ao português da escola. O espectador atento verá que, em sua maioria, muitos fenômenos da L1 do brasileiro, como os que foram ilustrados neste artigo, já são mais frequentes do que os que ainda caracterizam a escrita mais padronizada. E o mais importante é o fato de que o público em geral nem se dá conta, porque as legendas não se diferenciam da sua gramática.

5. Algumas considerações finais

Com este artigo, espero ter trazido algumas reflexões sobre a importância de levar em conta a história social numa análise sobre mudança linguística. Com base no que foi aqui relatado, e emergência de uma gramática "brasileira" era inevitável. As gerações mais jovens não aprenderam as regras do PE que contradizem os traços nacionais resultantes do intenso contato durante a colonização. Por outro lado, muitos traços do PE foram implementados pelos aloglotas, que, ao mesmo tempo, implementaram traços do PB, que seguiriam seu caminho por aqui, muitos dos quais viriam a mudar além-mar ao longo dos séculos que se seguiram ao descobrimento.

Tarallo (2018 [1993, p.19) já suspeitava da introdução de inovações sintáticas no PB antes do século XIX. Segundo o saudoso linguista, "circunstâncias sociais antes da virada do século XIX podem não ter sido suficientes o bastante para encorajar os brasileiros a deixar escorrer sua própria tinta". Acredito que as questões mostradas por Tarallo (e as que aqui são acrescentadas) tenham ocorrido muito antes da interpretação dos analistas, escondidas pelo duradouro analfabetismo no país e pela ideia de que o que se falava aqui pela quase totalidade da população não era uma "gramática nacional".

A verdadeira clivagem que existe no português brasileiro é a ignorância entre distância do modelo de gramática adotado e do português que se desenvolveu no Brasil, continuando o curso do português que aqui chegou mas incorporando outros traços, de difícil aprendizagem pelo grande contingente de africanos escravizados e seus descendentes, como as flexões, entre as quais podemos incluir os clíticos não-dêiticos, que não foram adquiridos, mas substituídos principalmente pelos complementos nulos (já presentes em pequena quantidade no português europeu e se expandiriam com variantes perfeitas para clíticos de difícil aquisição/aprendizagem), além das inovações, que incluem o uso de pronomes nominativos em todas as funções no PB. É urgente uma revisão mais realista da ideologia em torno da "unidade na diversidade", para que a diversidade seja conhecida e reconhecida. Essa norma da escrita, em que já se implementa tanta inovação (vinda da gramática da fala) exige muita pesquisa e uma mudança no preconceito que cega os que ainda acham que a fala, esse patrimônio que é adquirido naturalmente, seja apenas uma coisa de menor importância. Toda tentativa de sistematização da escrita brasileira, fruto do descompasso entre o que "se pode falar" e o que "se deve escrever", uma pregação de inúmeros filólogos do século XX, exige conhecimento de gramática e do legado que temos de uma extensa agenda de pesquisas acadêmicas aqui desenvolvidas. Olhar para esses resultados à luz do que o Projeto *Para a História do Português Brasileiro*, criado por Ataliba de Castilho, desde 1997 (ver os frutos dessa fundação registrados em Castilho, 1998), certos de que a história linguística não pode ser apartada da história social certamente nos iluminará o caminho de uma pesquisa isenta de ideologias.

Fontes Históricas

As peças portuguesas se encontram publicadas em sua maioria em Rebello, L. F. (org.) *Teatro Português em um Acto (1800-1899)*; *Teatro Português em um Acto (1900-1950)*; e *Todo o Teatro*, publicados em Lisboa pela Imprensa Nacional, Casa da Moeda, respectivamente em 1997, 2003 e 2006. Outras peças foram recolhidas na Biblioteca da Faculdade de Letras da Universidade de Lisboa. As peças brasileiras de Martins Pena, França Júnior e Millôr Fernandes estão todas publicadas. As demais se encontram disponíveis na Biblioteca Nacional do Rio de Janeiro, que atualmente distribui sob encomenda, as peças que se encontram cadastradas na SBAT – Sociedade Brasileira de Autores Teatrais. As peças de Miguel Falabella são encontradas na internet com uma marca d'água "para fins didáticos". Todas as dissertações e teses contendo as análises reunidas e reinterpretadas neste artigo trazem as peças utilizadas e as respectivas referências relativas à sua publicação.

Referências

AVELAR, J. De verbo funcional a verbo substantivo: uma hipótese para a supressão de HAVER no português brasileiro. *Letras de Hoje*, Porto Alegre: PUC-RS, v. 143, p. 49-74, 2006.

BRAGA, M. L. *Concordância de número do sintagma nominal no Triângulo Mineiro*. 1978. 103f. Dissertação (Mestrado em Língua Portuguesa), Rio de Janeiro: PUC-RJ, 1978.

CALLOU, D; BARBOSA, A. Normatização e ensino no final dos oitocentos
In: AGUILERA, V. (ed.) *Para a História do Português Brasileiro*, v. II, Vozes Veredas Voragens, Londrina: Eduel, 2009, p. 631-653.

CARDOSO, M. M. C. *Estratégias de indeterminação em peças portuguesas e brasileiras:* uma análise contrastiva. 2017. 115f. Dissertação (Mestrado em Letras Vernáculas), Rio de Janeiro: UFRJ, 2017.

CASTILHO, A. T. *Para a História do Português Brasileiro*, v. I: Primeiras ideias (Org.). São Paulo: Humanitas / FAPESP, 1998.

CREISSELS, D. Typology. *In*: HEINE, B.; NURSE, D. (eds.) *African languages: An introduction.* Cambridge: Cambridge University Press, 2000, p. 232-258.

CYRINO, S M. L. *O objeto nulo no Português do Brasil:* um estudo sintático-diacrônico. 1994. 217f. Tese (Doutorado em Linguística), Campinas: UNICAMP, 1994.

DUARTE, M. E. L. A perda ordem V(erbo) S(ujeito) em interrogativas Qu no português do Brasil. *DELTA*, v. 8, n. especial, p. 37-52, 1992.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. *In*: ROBERTS, I; KATO, M. A. (Orgs.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Ed. da UNICAMP. 1993, p. 107-128. [ed. São Paulo: Contexto, 2018, p. 83-103].

DUARTE, M. E. L. *A perda do princípio "Evite Pronome" no português brasileiro*. 1995. 151f. Tese (Doutorado em Linguística), Campinas: UNICAMP, 1995.

DUARTE, M. E. L. (ed.). *O sujeito em peças de teatro (1833-1992):* Estudos diacrônicos. São Paulo: Parábola Editorial, 2012.

DUARTE, M. E L. O sujeito nulo referencial no português brasileiro e no português europeu. *In*: GALVES, C.; KATO, M. A.; ROBERTS, I. (orgs.) *Português Brasileiro: uma segunda viagem diacrônica*. Campinas, Ed. da Unicamp, 2019. p. 93-126.

DUARTE, M. E. L. Nondeictic accusative and dative clitics and their variant forms in European and Brazilian Portuguese. *PROBUS*, v. 36, n. 2, p. 318-248, 2024.

DUARTE, M. E. L; GOMES, C.; PAIVA, M. C. Codification and standardization in Brazilian Portuguese. *In*: MUHR, R, (ed.), *Pluricentric languages and nondominant varieties worldwide*. Frankfurt am Main: Peter Lang. 2016, p. 51-65.

DUARTE, M. E. L; GOMES, C.; PAIVA, M. C. Beyond the dichotomy Dominantand/ Non-Dominant varieties of Pluricentric Languages: the case of Brazilian Portuguese; *In*: MUHR, R.; DUARTE, M. E. L.; RODRIGUES, C.; THOMAS, J. (eds.) *Pluricentric Languages in the Americas*: Series 2: Researching Pluricentric Languages, 2021, p. 157-152.

DUARTE, M. E. L.; MARINS, J. Brazilian Portuguese: a null subject language? *Cadernos de Estudos Linguísticos*, v. 63, p. 1-21, 2021.

EGERLARD, V. Impersonal pronouns in Scandinavian and Romance. *Working Papers in Scandinavian Syntax*, v. 71, p. 75-102, 2003.

FARACO, C. A. O Tratamento *você* em Portugues: Uma Abordagem Historica. *FRAGMENTA*, v. 13, p. 51-82, 1996.

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

FARACO, C. A. Aspectos da história socioeconômica e linguística do Brasil. *Diadorim* (Rio de Janeiro), v. 20, n. especial, p. 23-52, 2018.

FERNANDES, U. S. B. *Realizações do dativo de 3a pessoa em peças brasileiras e portuguesas*: uma análise diacrônica. 2018. 97f. Dissertação (Mestrado em Letras Vernáculas), Rio de Janeiro: UFRJ, 2018.

GALVES, C. A sintaxe do português brasileiro. *Ensaios de Lingüística*, v. 13, p. 31-49, 1987.

GALVES, C. Revisitando a concordância no português brasileiro. *In*: GALVES, C.; KATO, M. A.; ROBERTS, I. (eds.), *Português brasileiro: uma segunda viagem diacrônica*, Campinas: Editora da UNICAMP. p. 127-150, 2019.

GRACIOSA, D. M. *Concordância Verbal Na Fala Culta Carioca*. 1991. 99f. Dissertação (Mestrado em Letras Vernáculas), Rio de Janeiro: UFRJ. 1991.

GOMES, L. *Escravidão,* v. 1. Rio de Janeiro: Ed. Globo. 2019.

HOUAISS, A. *O Português no Brasil* - Pequena Enciclopédia da Cultura Brasileira, 2ª. ed. Rio de Janeiro: Unibrade, 1988 [1985].

KATO, M. A. A Gramática do Letrado: questões para a teoria gramatical. *In*: MARQUES, M. A.; KOLLER, E.; TEIXEIRA. L. J.; LEMOS, A. (eds.), *Ciências da Linguagem*: trinta anos de investigação e ensino, Braga: CEHUM, U. do Minho, 2005, p. 131-145.

KATO, M. A.; NASCIMENTO, M. do. (Orgs.) *A Construção da Sentença*. São Paulo Contexto. 2015. [1ª ed. 2007 Campinas, Editora da Unicamp]

KROCH, A. Syntactic Change. *In:* BALTIN, M.; COLLINS, C. (eds.) *The Handbook of Contemporary Syntactic Theory*. Malden: Blackwell. 2002, p. 699-729.

LEMLE, M.; NARO, A. *Competências básicas do Português*. Relatório de Pesquisa às instituições patrocinadoras: Fundação MOBRAL e Fundação FORD, maio de 1987.

LOPES, C. R.; CAVALCANTE, S. R. A cronologia do voceamento no português brasileiro: expansão de você-sujeito e retenção do clítico-te. *Revista Linguística* (Madrid), v. 25, p. 32-65, 2011.

LUÍZE, T. B. *Entre o PE e o PB:* o falar açoriano de Florianópolis. 1997. 115f. Dissertação (Mestrado em Linguística), Florianópolis: UFSC. 1997.

LUNGUINHO, M. V.; MEDEIROS JR. P. Inventou um novo tipo de sujeito: características sintáticas e semânticas de uma estratégia de indeterminação do sujeito no português brasileiro. *Revista de Estudos em Língua e Literatura*, v. 23, p. 147-191, 2009.

MARIANTE, B. M. S. *Há quanto tempo o português é brasileiro?* uma reanálise de trabalhos passados sobre a evolução do *ter* e *haver* existencial no contexto do PB. 2023. 36f. Monografia de Conclusão de Curso. Faculdade de Letras, Rio de Janeiro: UFRJ, 2023.

MARINS, J. *As repercussõesda remarcação do Parâmetro do Sujeito Nulo: um estudo diacrônico das sentenças existenciais com ter e haver no PB e no PE.* 2013. 154f. Tese (Doutorado em Letras Vernáculas), Rio de Janeiro: UFRJ. 2013.

MARQUES DE SOUSA, A. A. As realizações do acusativo anafórico no português europeu e brasileiro: um estudo diacrônico. 2017. 114 f. Dissertação (Mestrado em Letras Vernáculas), Rio de Janeiro: UFRJ, 2017.

MARTINS, M. A.; ABRAÇADO, J. (eds.). *Mapeamento sociolinguístico do português brasileiro.* São Paulo: Contexto, 2015.

MATTOS E SILVA, R. V. Para a história do português culto e popular brasileiro. *In:* ALKMIM, T. (org.) *Para a história do português brasileiro*, vol 3. São Paulo: Humanitas, 2002, p. 443-464.

MENEZES, E. T; SANTOS, T. H. Verbete Mobral (Movimento Brasileiro de Alfabetização). *Dicionário Interativo da Educação Brasileira - EducaBrasil.* São Paulo: Midiamix Editora, 2001. Disponível em https://educabrasil.com.br/mobral-movimento-brasileiro-de-alfabetizacao/. Acesso em 04 dez. 2024.

MILROY, J. Language ideologies and the consequences of standartization. *Journal of Sociolinguistics,* n. 5, v. 4, 2000, p. 530-555.

MOLLICA, C. *Estudo da cópia nas construções relativas em Português*. 1977. 86f. Dissertação (Mestrado em Língua Portuguesa), Rio de Janeiro: PUC-RJ, 1977.

MUSSA, A. B. A. *O papel das línguas africanas na história do português do Brasil.* 1991. 242f. Dissertação (Mestrado em Letras Vernáculas), Rio de Janeiro: UFRJ, 1991.

NASCENTES, A. *Estudos filológicos*. Rio de Janeiro: Civilização Brasileira.1939.

NUNES, J. 2019. Clíticos acusativos de terceira pessoa em português brasileiro como concordância de objeto. *In*: GALVES, C.; KATO, M. A.; ROBERTS, I. (eds.), *Português brasileiro*: uma segunda viagem diacrônica, Campinas: Editora da Unicamp, 2019, p. 151-172.

NUNES, J. Generalized phasal defectivity in Brazilian Portuguese. *Isogloss* 8(4), p. 101-129, 2022.

NICOLAU DE PAULA, M. *A ordem VS/SV e as interrogativas-Q no PE e no PB: uma análise diacrônica*. 2016. 131f. Tese (Doutorado em Letras Vernáculas), Rio de Janeiro: UFRJ, 2016.

OMENA, N. *Pronome pessoal de terceira pessoa. Suas formas variantes em função acusativa*. 1978. 152f. Dissertação (Mestrado em Língua Portuguesa), Rio de Janeiro: PUC-RJ, 1978.

PAGOTTO, E. Rui Barbosa e a crise normativa brasileira. *In:* CALLOU, D.; BARBOSA, A. *A Norma Brasileira em Construção*: cartas a Rui Barbosa (1866-1899), Rio de Janeiro: Fundação Casa de Rui Barbosa. 2011, p. 105-166.

PAGOTTO, E. A norma das constituições e a constituição da norma no século XIX, *Revista Letra*, p. 31-50, 2013.

PAIVA, M.C.; DUARTE, M. E. L. *Mudança linguística em tempo real* (orgs.) Rio de Janeiro: ContraCapa. 2003.

PINHEIRO, D; MARINS, J. A trajetória das interrogativas-QU, clivadas e não clivadas no português brasileiro. *In*: DUARTE, M. E. L. (org.) *O sujeito em peças de teatro (18331992): Estudos diacrônicos.* São Paulo: Parábola, 2012, p. 161-180.

ROCHA, P. R.; CUNHA JR, M. O ensino secundário no Brasil. *In*: CASTRO, P. A. *et al.* (orgs.). *Escola em tempo de conexões*. Campina Grande: Realize Editora, 2022. p. 319-339.

SCHERRE, M. M. P. *A Regra de Concordância de Número no Sintagma Nominal em Português*. 1978. 174f. Dissertação (Mestrado em Língua Portuguesa), Rio de Janeiro: PUC-RJ. 1978.

TARALLO, F. Diagnosticando uma gramática brasileira: o Português d'aquém e d'além mar ao final do século XIX. *In*: ROBERTS, I; KATO, M. A. (orgs.) *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora Unicamp. 1993, p. 55-82. [2018, ed. São Paulo: Contexto].

VIEIRA PINTO, C. A. *Variação do objeto anafórico acusativo na fala de Florianópolis*. 2015. 164f. Dissertação (Mestrado em Linguística), Florianópolis: UFSC, 2015.

http://doi.org/10.5007/1984-8420.2025.e108152

O SUJEITO NULO EM CARTAS BRASILEIRAS DOS SÉCULOS XIX E XX: NOVOS DADOS PARA UMA VELHA MUDANÇA[1]

NULL SUBJECTS IN BRAZILIAN PRIVATE LETTERS FROM THE 19TH TO THE 20TH CENTURIES: NEW DATA FOR AN OLD CHANGE

Silvia Regina de Oliveira Cavalcante[2] | Lattes | silviare@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Resumo: A sintaxe do Português Brasileiro (PB), especificamente a remarcação do Parâmetro do Sujeito Nulo (PSN), tem sido estudada amplamente tanto no Brasil quanto no exterior, desde uma perspectiva estritamente teórica (com dados de intuição) a uma perspectiva mais empirista, com dados da fala e da escrita. O sujeito nulo no PB tem mais restrições para a sua interpretação e licenciamento do que no PE e em outras línguas românicas de sujeito nulo. Este trabalho traz uma análise diacrônica do sujeito nulo de referência definida numa amostra de cartas pessoais de brasileiros nascidos ao longo dos séculos XIX e XX. A metodologia de pesquisa alia os pressupostos teóricos da gramática gerativa com a metodologia de análise estatística de dados proveniente da Sociolinguística, considerando que a mudança linguística ocorre no período de aquisição da linguagem, e que a gradualidade que vemos nos dados se dá na substituição de uma gramática por outra. As frequências de uso, desse modo, vão refletir a competição entre gramáticas distintas até que a gramática nova substitua a mais antiga. Os resultados revelam que o índice geral de sujeito nulo sai de 70% e vai para 38%; além disso, os contextos sintáticos do sujeito nulo nas cartas dos missivistas nascidos na primeira metade do século XIX são mais amplos do que os contextos do sujeito nulo nas cartas dos missivistas da segunda metade do século XX.

Palavras-chave: Sujeito nulo; Mudança linguística; Português Brasileiro; Cartas pessoais.

[1] Este trabalho é um dos resultados do Projeto de Pesquisa "O sujeito nulo na história do Português Brasileiro: uma mudança encaixada", financiado parcialmente pelo CNPq (com bolsas de Iniciação Científica para os alunos que integram o projeto). Agradeço aos dois pareceristas anônimos, que enriqueceram o trabalho com as suas indagações e sugestões. A maioria das sugestões foram incorporadas ao trabalho e serão levadas em consideração em trabalhos futuros.

[2] https://orcid.org/0000-0003-3264-3572

http://doi.org/10.5007/1984-8420.2025.e108152

Abstract: The syntax of Brazilian Portuguese (BP), specifically as for the resetting of the Null Subject Parameter (NSP), has been studied extensively both in Brazil and abroad, from a strictly theoretical perspective (with intuition data) to a more empiricist perspective, with data from oral and written samples. The null subject in BP has more constraints for its interpretation and licensing that in EP or other Romance Null Subject Languages. This paper presents a diachronic analysis of the null subject of definite reference in a sample of private letters written by Brazilians born throughout the 19th and the 20th centuries. The methodology of this research combines the theoretical assumptions of generative grammar with the methodology of statistical data analysis from Sociolinguistics, considering that language change occurs during the period of language acquisition, and that the gradualness that we see in the texts occurs in the replacement of a new grammar for another. The frequency of usage, in this way, will reflect the competition between different grammars until the new grammar replaces the older one. The results reveal that the overall null subject rates go from 70% to 38%. Furthermore, the syntactic contexts of the null subject in letters written in the first half of the 19th century are broader than the contexts of the null subject in the letters of the second half of the 20th century.

Keywords: Null subject; Linguistic change; Brazilian Portuguese; Private letters.

Introdução

Desde Tarallo e Kato (1989) e Tarallo (1993), os estudos sobre mudança linguística no Brasil ganharam nova roupagem, com a associação do modelo teórico Gerativista com a análise quantitativa da Sociolinguística Variacionista. Nesta seara, os estudos sobre a sintaxe do Português Brasileiro (PB) têm mostrado um quadro de diferenças microparamétricas (Baker, 2008) em relação ao Português Europeu (PE) e a outras línguas românicas, principalmente no que tange ao Parâmetro do Sujeito Nulo. O capítulo seminal de Duarte (1993), que traz uma análise diacrônica sobre o sujeito nulo em peças brasileiras, foi o pontapé para os estudos sistemáticos sobre a remarcação do Parâmetro do Sujeito Nulo no PB. Podemos citar os trabalhos de Duarte (1995) sobre a mudança na fala de indivíduos cultos que foi seguido por inúmeros trabalhos tanto na fala quanto na escrita, em comparação com o PE e outras línguas românicas (Barbosa, Duarte e Kato, 2005; Duarte, 2000; 2007; 2019) entre outros trabalhos diacrônicos com base em amostras de cartas pessoais (Vieira-Pinto, 2020; Coelho et al. 2017, 2021; Coelho e Vieira-Pinto, 2021)

Resumidamente, o PB caracteriza-se não só por ter índices de sujeito nulo mais baixos do que os do PE, mas também apresenta contextos sintáticos mais restritos para o licenciamento e a interpretação do sujeito nulo de referência definida. Além disso, o sujeito nulo de terceira pessoa do singular recebe interpretação genérica ou arbitrária, o que tem levado a afirmar que o PB se caracteriza como sendo uma gramática de sujeito nulo parcial, se afastando das outras línguas românicas e se aproximando do Finlandês e do Hebraico (Kato, 2000; Rodrigues, 2004; Holmberg, 2005, 2010; Marins, Soares da Silva, Duarte, 2017, entre outros[3]).

Poderíamos questionar se um tema tão estudado na literatura ainda suscita campo de trabalho, depois de 30 anos de investigação. Felizmente, ainda há o que dizer sobre o sujeito nulo no PB, não só do ponto de vista teórico, mas também empírico. Desse modo, neste artigo, apresento uma análise do sujeito nulo de referência definida em cartas pessoais brasileiras, a fim de mostrar a implementação e o encaixamento da mudança paramétrica na escrita. Trata-se de uma escrita particular, pois a amostra analisada é constituída de cartas pessoais escritas por brasileiros nascidos do início do século XIX ao final do século XX, que compõem o Corpus de História do Português – Corpus HistLing, da UFRJ.

Por hipótese, os resultados das cartas pessoais refletirão os resultados já encontrados na literatura, qual seja, a diminuição nos índices de sujeito nulo ao longo do tempo, e a diminuição do sujeito nulo em determinados contextos sintáticos. A pergunta principal levantada aqui é: é possível detectar a mudança na interpretação e licenciamento desse sujeito nulo de terceira pessoa em uma amostra diacrônica de cartas pessoais? Para fins metodológicos, trago aqui somente a análise dos sujeitos de 3ª pessoa de referência definida em sentenças matrizes, completivas, adverbiais e relativas, ao longo de sete períodos de tempo, considerando a data (provável) de nascimento dos missivistas.

Este artigo está organizado da seguinte forma: na seção 1, apresento a discussão sobre a interpretação e o licenciamento do sujeito nulo no PB, considerando o quadro teórico gerativista e as hipóteses levantadas; na seção 2, o corpus e a metodologia de trabalho são apresentados, bem como os objetivos específicos; na seção 3, apresento os resultados e a discussão, procurando responder se as hipóteses foram confirmadas e, por fim, considerações finais e referências.

[3] Não é consenso na literatura a caracterização do PB como uma língua parcialmente *prodrop*: para além da interpretação indefinida (genérica e arbitrária) do sujeito nulo de 3ª. pessoa do singular, o PB apresenta algumas restrições sobre sujeitos expletivos que as línguas notadamente parcialmente *prodrop* apresentam. Como neste artigo, estou tratando exclusivamente dos sujeitos de referência definida, não entro em detalhes sobre a questão, apenas apresento como uma das características do PB, em face às línguas românicas de sujeito nulo.

1. O sujeito nulo no PB: interpretação e licenciamento

A discussão sobre a natureza do sujeito nulo no PB se inicia nos anos 1980 com Moreira da Silva (1983) e até hoje ainda traz desafios para a teoria linguística, não só pela frequência do sujeito nulo no PB, mas também pelos contextos sintáticos em que é licenciado. Apresento uma breve discussão sobre o fenômeno em que a análise aqui apresentada se baseia, cotejando alguns resultados de pesquisas empíricas e também alguma descrição teórica.

Partimos do trabalho seminal de Duarte (1993), que mostra a diminuição dos índices de sujeito nulo em peças brasileiras dos séculos XIX e XX. Nas peças de meados do século XIX o índice de sujeito nulo de terceira pessoa sai de 83% e chega a 55% na peça de 1992. A partir daí, diversos trabalhos empíricos com base em amostras de fala, procuraram descrever o PB em comparação a outras línguas românicas de sujeito nulo, tais como Duarte (2003), Marins (2009), sobre o Italiano; Soares da Silva (2006), sobre o Espanhol. De modo geral, os índices de sujeito nulo no PB são bem mais baixos do que nas outras línguas românicas de sujeito nulo, como mostrado na Tabela 1 a seguir:

Tabela 1: Percentual de Sujeito Nulo no PB e outras línguas românicas

	PB (EM) (Duarte, 2003)	PB 3o. Grau (Duarte, 2003)	PE (Duarte, 1995)	Esp. BA (Silva, 2006)	Esp. Madri (Silva, 2006)	Italiano (Marins, 2009)
Nulo	843	415	738	843	943	630
Total	4262	1424	1116	1221	1244	751
%Nulo	20%	29%	66%	69%	76%	84%

Fonte: Duarte (2003); Silva (2006); Marins (2009). Elaboração própria.

Podemos notar que os índices de sujeito nulo no PB falado, tanto por falantes com ensino médio, quanto ensino superior, são bem mais baixos do que os índices de sujeito nulo em outras línguas românicas, tanto da România Velha (Espanhol de Madri, Italiano e PE), quanto da România Nova (Espanhol de Buenos Aires), o que já pode evidenciar uma diferença paramétrica.

Além de trabalhos com base em amostras de fala, a mudança no sujeito nulo tem sido estudada também em amostras da escrita, tais como Duarte, Mourão e Santos (2012) e Duarte (2019), que comparam PE e PB em peças de teatro e também Vieira-Pinto (2020), Coelho et al. (2017, 2021), Coelho e Vieira-Pinto (2021), que trazem análises do sujeito nulo em cartas pessoais brasileiras. Destaco aqui o trabalho de Vieira-Pinto (2020), realizado com base em uma amostra de cartas catarinenses, que mostra uma di-

minuição nos índices de sujeito nulo entre os séculos XIX e XX: o índice geral de sujeitos nulos (excluídas as orações coordenadas) sai de 69% na segunda metade do século XIX e passa a 60% na segunda metade do século XX.

Segundo Kroch (1989), Tarallo e Kato (1989), os dados estatísticos são importantes, pois indicam características da *gramática* de uma língua. Entendemos aqui *gramática* no sentido gerativista da palavra, e *gramáticas* distintas produzem índices distintos de uma determinada construção ou forma. Desse modo, os dados estatísticos apontam para características sintáticas diferentes envolvendo o sujeito nulo no PB e nas outras línguas românicas de sujeito nulo. De fato, Galves (1998) afirma que:

"Duas Línguas-I serão consideradas diferentes se contêm na sua parametrização pelo menos um parâmetro fixado diferentemente. Quando isso ocorre, não só as duas gramáticas produzem enunciados diferentes, mas também atribuem a enunciados superficialmente idênticos (por exemplo no arranjo dos constituintes) estruturas diferentes". (Galves, 1998:80)

Do ponto de vista teórico, desde o modelo de Regência e Ligação até análises mais recentes dentro do Programa Minimalista, tem havido diversas análises da natureza do sujeito nulo no PB, em comparação, principalmente com o PE: as análises comparativas entre PB e PE apontam que o sujeito nulo no PE é de natureza pronominal, a "verdadeira" categoria vazia *pro*; ao passo que no PB, o sujeito nulo tem sido analisado como uma categoria vazia do tipo *pro*, mas que não é legitimada pela concordância, mas precisa de um antecedente numa posição A-barra, ou como vestígio de movimento (Ferreira, 2000; Rodrigues, 2004; Modesto, 2000; Galves, 2019). Não é objetivo deste artigo advogar a favor de uma das análises, mas cabe aqui mencionar os contextos sintáticos em que o sujeito nulo é licenciado, que são a base da nossa análise quantitativa dos dados. O sujeito nulo no PB pode corresponder a um expletivo nulo, como vemos nos exemplos em (1):

(1) a. [**0**] Parece que o João passou por aqui.
b. [0] Choveu a noite inteira.

O sujeito nulo da subordinada tem o seu referente controlado pelo sujeito da matriz, como vemos em (2b); isso justifica o referente do sujeito nulo da subordinada ter que ser controlado pelo antecedente na matriz, o que explica o julgamento em (2b). Quando há dois antecedentes possíveis, como em (2c), o referente do sujeito nulo é o

mais próximo e a referência disjunta do sujeito nulo é agramatical, como em (2d), como mostra Rodrigues (2004):

(2) a. O João$_1$ disse que e$_1$ vem amanhã de manhã.
b. [o pai do Paulo1]2 disse que e*1/2 vai ser promovido.
c. O Paulo1 me contou que o João2 disse que e*1/2 vai mudar para São Paulo.
d. * João1 disse que Maria2 acha que e1+2 vão viajar. (Cf. Rodrigues, 2004)

O sujeito nulo é excluído em contextos de ilha relativa, como vemos nos contrastes entre o espanhol (3a) e o PB (3b), mas pode ocorrer em sentenças adjuntas, que também são um contexto de ilha forte, como vemos no contraste em (4):

(3) a. Juan$_1$ vio a la chica que el/*pro*$_1$ besó anoche
Juan viu a moça que ele/pro beijou noite passada
b. O João1 encontrou a carteira que ele/*e perdeu. (Cf. Rodrigues, 2002)

(4) a. ?? João$_1$ disse que as meninas que e$_1$ encontrou na rua eram estrangeiras.
b. João1 comeu um pastel quando e1 foi na feira. (Cf. Ferreira, 2000:102-103)

Além disso, o PB permite sujeitos plenos para retomar antecedentes com traço [-animado], como vemos nos exemplos em (5):

(5) a. **A casa** virou filme quando **ela** teve de ir abaixo.
b. Nova Trento é do tamanho da rua São Clemente de Botafogo. Ela é desse tamanho. Ela não tem paralelas. (Duarte, 2000:22)

Como vemos, além das diferenças de frequência de uso, há diferenças na interpretação do sujeito nulo e do pronome pleno no PB, em relação ao PE e na possibilidade de o pronome pleno retomar um antecedente [-animado], e os contextos sintáticos que licenciam o sujeito nulo. À vista disso, neste trabalho, buscamos controlar os contextos sintáticos do sujeito nulo (ou pleno) em análise para saber como se dá a mudança nos textos históricos. Por hipótese, esperamos uma diminuição dos índices de sujeito nulo e uma restrição dos contextos sintáticos em que ele pode aparecer. Vejamos o *corpus* e a metodologia antes de passar aos resultados.

2. *Corpus* e Metodologia

A amostra desta pesquisa é constituída de cartas pessoais brasileiras que compõem o Corpus de História do Português – Corpus HistLing (https://histling.letras.ufrj.br/), que é organizado de acordo com o Fundo Documental das Famílias. O Corpus HistLing está se formando ao longo do tempo de cartas depositadas em arquivos públicos e também de acervos pessoais. Desse modo, dentre os missivistas estudados, encontram-se pessoas "ilustres", tais como o médico sanitarista Oswaldo Cruz e pessoas "não-ilustres", tais como o casal de namorados Jayme e Maria, cujas cartas foram encontradas no lixo e entregues à coordenação do Corpus.

Por se tratar de pessoas, de certo modo, conhecidas, foi possível determinar períodos de tempo relacionados com as datas de nascimento desses missivistas. Na amostra considerada, estabelecemos sete períodos de tempo para as datas de nascimento dos missivistas, considerando um intervalo de 25 anos para tratar uma geração: Período 1 (1801-1825); Período 2 (1826-1850), Período 3 (1851-1875), Período 4 (1876-1900); Período 5 (1901-1925); Período 6 (1926-1950); Período 7 (1951-1975).

Apresento a amostra, objeto de análise deste artigo, organizada por famílias, para que tenhamos uma ideia das informações sócio-históricas relevantes dos nossos missivistas. As famílias estudadas são as seguintes:

- **Avós Ottoni**: 41 cartas escritas entre 1879 e 1889 pelo casal Christiano Benedicto Ottoni (engenheiro, Senador do Império e depois Senador da República) e sua esposa Barbara Balbina de Araújo Maia Ottoni a seus netos Mizael e Christiano enquanto estes moravam em Paris. O casal é nascido no primeiro quartel do século XIX, 1801-1825.
- **Pedreira Ferraz Magalhães**: correspondência entre os pais e filhos da família. O pai, Jerônimo de Castro Abreu Magalhães (engenheiro civil, nascido em 1851), e a mãe, Zélia Pedreira Abreu Magalhães (esposa de Jerônimo, nascida em 1857), se correspondem com os filhos, que eram internos de conventos e mosteiros. Há três gerações de missivistas nessa família, nascidos entre 1826 e 1900.
- **Oswaldo Cruz**: correspondência entre Oswaldo Cruz, médico sanitarista brasileiro, sua esposa Emília e sua filha. Os missivistas nasceram entre 1851 e 1900.
- **Afonso Penna Jr.**: correspondência entre Affonso Penna Jr., filho de Affonso Augusto Moreira Penna (ex-presidente da República), e seus parentes. Os missivistas nasceram entre 1826 e 1900.

- **Casal Jayme-Maria**: cartas escritas entre 1936 e 1937 por um casal de namorados residentes no Rio de Janeiro, Jayme e Maria. O que se sabe deste casal são as informações retiradas das próprias cartas: Jayme morava em Ramos, subúrbio carioca; e Maria morava em Petrópolis, na Serra Fluminense. Estima-se, pelo conteúdo das cartas, que o casal tenha nascido entre 1901 e 1925.
- **Frazão Braga**: cartas escritas na 2ª metade do século XX, entre 1956 e 1994, por R. F. B. e seus familiares: W. (filho), M. (filho), M. R. (filha), E. (nora), M. H. (nora), D. (neto), A. (neta), At. (neto) ,Wl. (neto) e Ax. (bisneto). Os missivistas desta família nasceram entre 1876 e 1950.
- **Salgado Lacerda**: cartas escritas entre 1977 e 1983, que fazem parte da correspondência da jovem MSL, que fazia intercâmbio no exterior, com seus pais, irmãos e amigo da família. Os missivistas nasceram entre 1926 e 1975.

Como se pode perceber, com exceção do casal Jayme e Maria, todos os missivistas poderiam ser enquadrados na classe média / média alta, ou até mesmo, na elite brasileira. O casal Ottoni, por exemplo, é composto por um político e sua esposa. Os missivistas da Família Pedreira Ferraz Magalhães foram estudar, na sua maioria, em conventos e monastérios e isso pode, de certo modo refletir na sua escrita. Os fundos Frazão Braga e Salgado Lacerda podem ser considerados pertencentes à classe média / média alta, pois os missivistas se correspondem com pessoas que foram morar no exterior, seja para trabalhar (que é o caso de A., da família Frazão Braga), seja para estudar (que é o caso da jovem MSL, da família Salgado Lacerda).

Antes de passar aos contextos sintáticos considerados para análise, cabe uma observação. Esta pesquisa parte da metodologia variacionista para o tratamento estatístico dos dados com o programa GoldVarb X (Sankoff, Tagliamonte, Smith, 2005). A escolha para uma análise estatística probabilística (para além dos percentuais), com o uso do GoldVarb X, se justifica pelo quadro teórico de mudança que subjaz este trabalho: considero, seguindo Kroch (1989), que a mudança linguística ocorre via competição de gramáticas, em que uma gramática substitui a outra ao longo do tempo. Além disso, por se tratar de uma análise diacrônica, em que a distribuição dos dados ao longo do tempo não é uniforme – fazemos usos dos dados que encontramos – a análise probabilística corrige essas distorções, mas não só. A análise probabilística, nas palavras de Tarallo e Kato (1989:14), é "um modelo de análise da linguagem que estatisticamente garante sua cientificidade ao projetar as probabilidades dos fatores que mais favorecem ou, ao contrário, inibem o comportamento de formas em variação e mudança".

Além disso, na determinação dos grupos de fatores utilizados na descrição dos dados, parto de análises gerativistas e por isso, como valor de aplicação foi escolhido o sujeito nulo, para saber quais contextos sintáticos favorecem o sujeito nulo, observando os contextos sintáticos de "resistência" do sujeito nulo (Duarte, 2019), diferentemente de considerar os contextos de implementação do sujeito pleno. Essa decisão metodológica procura responder uma das perguntas apresentada na introdução e retomada aqui: é possível detectar a mudança na interpretação e licenciamento desse sujeito nulo de terceira pessoa em uma amostra diacrônica de cartas pessoais?

Na tentativa de responder a essa pergunta, controlo aqui, além do período de nascimento e do gênero (masculino ou feminino) dos missivistas, os seguintes fatores estruturais: tipo de sentença, traço de animacidade e especificidade do antecedente, padrão estrutural da sentença com sujeito nulo, pessoa e número (3ª. Singular ou Plural) do sujeito nulo (vs. pleno) em análise.

Passemos às justificativas para tratar cada contexto sintático. Em primeiro lugar, consideramos o período de tempo como a provável data de nascimento do missivista, por se tratar de uma análise gerativista que parte do pressuposto que a mudança linguística ocorre no período de aquisição. Estabelecemos um intervalo de 25 anos entre uma geração e outra, conforme já mencionado anteriormente.

Com relação ao traço de animacidade e especificidade do antecedente, a pesquisa se baseia em estudos anteriores, principalmente o de Cyrino, Duarte e Kato (2000), que mostram que existe uma relação entre os traços de animacidade e especificidade para a pronominalização. As autoras propõem a escala da referencialidade, que sai de antecedentes menos referenciais, como os não-argumentais, até chegar aos referentes mais referenciais, como a 1ª. pessoa, por exemplo, tal como se vê abaixo:

não-argumento	proposição	[-humano]		[+humano]
		3 p.		3 p. 2 p. 1 p.
			-espec.	+espec.
[-ref] < ---------- ---------- ---------- > [+ref.]				

Quadro 1: Escala de Referencialidade segundo Cyrino, Duarte e Kato (2000)

Segundo as autoras, o estatuto referencial do antecedente (da forma nula ou plena) é um fator forte para a seleção de uma ou outra forma, qual seja, o uso do pronome

pleno é preferível em contextos mais referenciais, ao passo que os contextos menos referenciais preferem o pronome nulo (tanto na posição de sujeito quanto de objeto). Neste trabalho, como só analiso os sujeitos de 3ª pessoa, controlamos os traços de animacidade [+/-animado] em combinação com o traço de especificidade [+/- específico]. A mudança paramétrica que afeta o PB vai fazer aumentar os índices de sujeito pleno não só nos casos de antecedentes [+animados, +específicos], mas também nos casos de referentes [-animados, -específicos]. Os exemplos a seguir trazem os casos de antecedente [+animado, +específico] (6), [+animado, -específico] em (7), [-animado, +específico] em (8) e [-animado, -específico] em (9):

(6) a. Diga a **Marieta** q se [**0**] quizer alguma cousa escreva. (Período 3, Fundo Affonso Penna)
b. Vae tambem a sua esposa a eles desejo que [0] sejam felizes e que se acostumem no clima frio e que nada lhes acuntece de alteração em sua saude. (Período 4, Fundo Frazão Braga)

(7) Moral da estória, **todos os brasileiros** que pintam por aqui são assim, só querem que nós servimos de guia e na hora de nós perdimos um favor êles partem, somem, são uns imprestáveis (Período 6, Fundo Frazão Braga)

(8) Enfim êste **diploma** está em casa. Nada de ficar andando, pois tôda a burocracia a ser feita eu já fiz. Repito êle está em casa. Bom não se afobem. Se não acharem darei um jeito, pois tenho meu amigo (Período 6, Fundo Frazão Braga)

(9) a. E os seus **frangos**, [**0**] tem dado lucro? Estou doida para ver as fotografias do sítio. (Período 7, Fundo Salgado Lacerda).
b. e os sabados sempre faco Paõ doce e melembro de de que sevoce estivesse a qui avia meajudar equando elles vem açadinhos do forno eu digo logo se Tixe istivesse aqui como elle avia de gostar. (Período 1, Fundo Ottoni).

O tipo de sentença analisado está relacionado com as hipóteses sobre o sujeito nulo no PB ser licenciado em contextos de controle, que seriam os casos ilustrados em (2) e (3) acima. Em um contexto de mudança linguística, vamos esperar que um contexto de resistência ao sujeito nulo seriam as completivas, cujo referente está na matriz, como se vê em (10). Além das completivas, controlamos a matriz (11), subordinada adverbial (12), subordinada relativa (13) e segunda coordenada (14), como se vê a seguir.

(10) Lembro-me que, quando [teu maninho] disse que [**0**] se-chamaria- C. Ottoni Vieira. (Período 1, Fundo Ottoni)

(11) O passaporte do Libanês estava faltando visto de entrada em paris (protocolo do Líbano) e com isto as **autoridades** começaram a desconfiar do toque aventureiro dos 2 e de mim tam- bém . [**0**] Nos massacraram de perguntas. (Período 6, Fundo Frazão Braga)

(12) Podem avisando o**s irmãos preguiçosos** se [**0**] quizerem alguma coisa [rasurado] fazerem economia e mandarem o tutu pois não terei limite de peso e provàvelmente levarei 3 baús com antigos para casa e outras milongas mais (Período 6, Fundo Frazão Braga).

(13) Alexandre caza e segue no mesmo dia para aqui onde [**0**] vem passar três dias. (Período 3, Fundo Affonso Penna)

(14) **Papai** está de remedios do Dr F. de Castro e [**0**] tem melhorado graças a Deos. (Período 3, Fundo Affonso Penna)

O próximo contexto sintático controlado é o padrão estrutural, que diz respeito não só ao sujeito que está sendo analisado, mas também ao seu antecedente. Esse contexto é controlado, a fim de buscar responder sobre a natureza do sujeito nulo, se *prodrop* ou *topic drop*: Galves (1993; 2019) e Kato, Martins e Nunes (2023) propõem que o sujeito nulo no PB pode ser licenciado por um tópico, ou mesmo gerado via movimento. Além disso, o sujeito nulo do PB deve ser controlado pelo sujeito da oração matriz. Torregrossa, Andreou e Bongartz (2020)[4] apresentam outra maneira de controlar o antecedente, com a Estratégia da Posição do Antecedente (PAS, em inglês), em sua análise comparativa do sujeito nulo do grego e do italiano. Os autores consideram a Estratégia da Posição do Antecedente, que de modo geral, prevê que em situações intrassentenciais, o sujeito nulo tem o seu antecedente na posição de sujeito; ao passo que o sujeito pleno aparece quando o seu antecedente está numa posição sintática abaixo de IP, como a de objeto. Além disso, a Estratégia de Posição do Antecedente pode aparecer em contextos inter-sentenciais, uma vez que o antecedente do sujeito nulo pode estar no contexto anterior. Neste trabalho, optei por controlar os padrões sentenciais seguindo Duarte (2019), a fim de podermos comparar os resultados do PB e do PE.

Vejamos os cinco padrões estruturais, explicitados a seguir, a fim de verificar os contextos de resistência do sujeito nulo; i.e., os contextos que favorecem o sujeito nulo.

- Padrão A: O antecedente do sujeito nulo está na sentença matriz, com a mesma função e controla a referência do sujeito na sentença subordinada:

[4] Agradeço ao parecerista anônimo que me indicou essa referência bibliográfica e que poderá enriquecer futuros passos da pesquisa.

(15) **Bêbê** me diz que [**0**] fica - até o fim deste anno em Petropolis; (Período 4, Fundo Pedreira Ferraz Magalhães)

(16) Ontem é que **Atefeh** me falou que **ela** vai esperar p/ chegar no Iran para me mandar o tapete. (Período 7, Fundo Salgado Lacerda).

- Padrão B: O antecedente está na sentença adverbial precedente, com a mesma função e identifica o sujeito na sentença matriz:

(17) Samuel anda doente, e hontem eu o mandei ir pa caza tomar remedio e o Missias é quem está cozinhando. Hoje elle varreu a caza e fez o almoco e agora está as voltas com o jantar. Quando **elle** quer [**0**] é bem bomsinho. (Período 3, Fundo Affonso Penna)

(18) Leonor está bem melhor graças a Deus se **ella** fizesse os votos creio que [**0**] ficaria boa não achas?

- Padrão C: O referente está numa sentença adjacente, com a mesma função do sujeito. Entendemos como sentença adjacente o período imediatamente anterior ao sujeito em análise, que seja diferente do período em que se encontra o sujeito. Assim, em (19) a seguir, o referente do sujeito nulo é "Christiano", que é um sujeito de uma subordinada; entretanto, tal referente está no período imediatamente anterior ao período do sujeito nulo[5]:

(19) Observo que **Christiano** nada me diz de relaçoes adquiridas a bordo: [**0**] diz que so se occupa de ler, comer e dormir: mas isso naõ convem. (Período 1, Fundo Ottoni).

(20) Estimei saber que o Mario está melhor; mas não seja isso rasão para se descuidarem de tratamento definitivo . **A molestia** é incidiosa e [**0**] vae seguindo seo caminho de modo implacavel, se [**0**] não for atalhada a tempo. (Período 2, Fundo Affonso Penna)

- Padrão D: O antecedente está numa sentença adjacente, com função diferente da de sujeito:

(21) Estou ajudando **o Mestre Cel**. num livro que [**0**] lançará no Brasil. Excelente, atualíssimo e sempre escorregando no nosso atual govêrno. Como diz êle: é inevitável e é aí que entra o problema. Poderá ser editado ou não??!! (Período 6, Fundo Frazão Braga)

(22) Vou levar **Olga** ao Dor Couto pa ver se **ella** ain-da deve tomar um mez de banhos pois agora é que estam principi-ando as melhoras . (Período 3, Fundo Affonso Penna)

[5] Entendemos como "sentença adjacente" a sentença imediatamente anterior ao período em que se encontra o sujeito nulo/pleno em análise.

- Padrão E: O antecedente está distante ou é um tópico discursivo:

(23) Já fechei o negocio com o **Bentinho**. Peço-te que não te esqueças de trazer a perola quando vieres, sim ? Pago aqui na Europa 50 libras a proporção que [**elle**] precizar , o resto quando eu puder no Rio. (Período 4, Fundo Oswaldo Cruz)

Em todos esses contextos sintáticos, esperamos encontrar uma diminuição nos índices de sujeito nulo ao longo do tempo, que não vai se dar da mesma forma: acreditamos haver uma maior retenção de sujeito nulo nos contextos considerados de "controle", como o Padrão A e nos contextos em que o sujeito está associado a um tópico, como no Padrão E.

Com relação às sentenças coordenadas, decidimos controlá-las, mas apresento os seus resultados separadamente por dois motivos: em geral, as análises quantitativas sobre sujeito nulo não levam em consideração as sentenças coordenadas, pois se trata de contexto de apagamento do sujeito por elipse e esse "sujeito nulo" pode aparecer também em línguas de sujeito não nulo, como o Inglês. Entretanto, decidimos controlar tais contextos para ver se há diminuição nos índices de sujeito nulo também nesses contextos, como resultado da implementação da mudança.

Passemos à apresentação e à discussão dos resultados quantitativos.

3. Resultados e Discussão

A análise de dados está organizada em duas partes: em primeiro lugar, apresento os resultados gerais do percentual de sujeito nulo e pleno ao longo do tempo; em seguida os resultados advindos da rodada de regra variável, com o uso do Programa GoldVarb, com a discussão dos fatores selecionados como relevantes para o sujeito nulo.

Foram analisados 1766 dados de sujeito de terceira pessoa de referência definida, ao longo de sete períodos de tempo. O índice geral de sujeito nulo foi de 54%; o que não nos diz muita coisa se estamos considerando o encaixamento da mudança linguística. Desse modo, faz-se necessária a apresentação dos resultados gerais ao longo do tempo, por data de nascimento do missivista.

No Gráfico 1, apresento a evolução dos índices de sujeito nulo e pleno ao longo do tempo, considerando sentenças matrizes, completivas, adverbiais e relativas: podemos ver que os índices de sujeito diminuem ao longo do tempo, saindo de 70% nas cartas dos missivistas nascidos no primeiro quartel do século XIX e chegando a 38% nas cartas dos missivistas nascidos na segunda metade do século XX. Tais resultados, de certo modo, estão consoantes com os resultados de Duarte (1993; 2019) para os sujeitos de terceira pessoa em peças brasileiras: na amostra de peças, há a

diminuição dos índices de sujeito nulo que sai de 83% nas peças de 1845 e chega a 55% nas peças de 1992.

Nos dados das cartas, chama a atenção a subida entre os períodos 3 e 4, que destoam dos resultados dos períodos adjacentes. No período 4, encontram-se majoritariamente os filhos e filhas do casal da Família Pedreira Ferraz Magalhães, que foram todos enviados a conventos e monastérios para estudar. Assim, acredito que essa subida possa se dar em virtude de uma associação com a norma padrão e ao modelo do PE. Não cabe aqui, entretanto, fazer uma análise mais profunda da sócio-história dos missivistas, em virtude de não ser esse o foco deste artigo. Tal análise ficará para trabalhos futuros, em virtude dos temas de pesquisa sendo desenvolvidos dentro do Projeto de Pesquisa. Passemos à visualização da distribuição percentual do sujeito nulo e pleno ao longo do tempo pela data de nascimento dos missivistas:

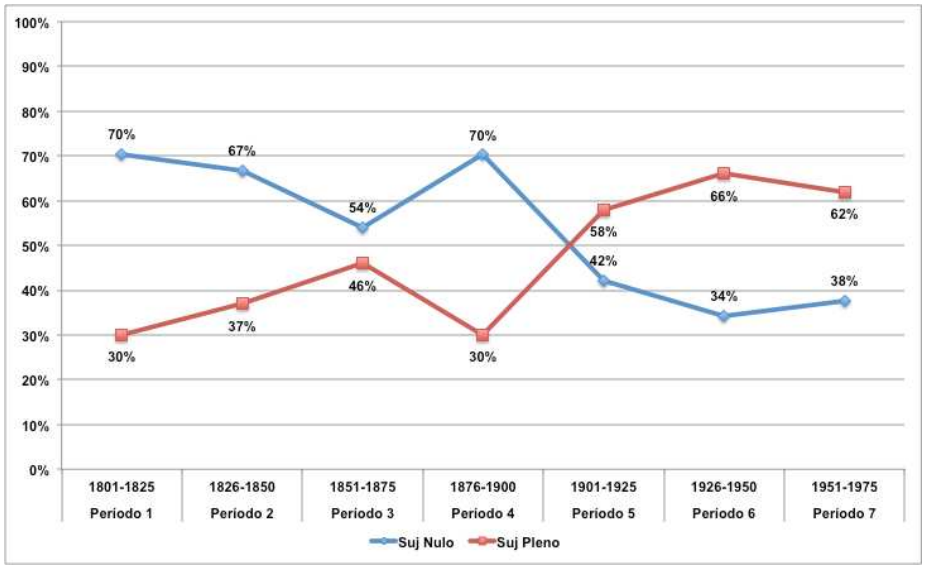


Gráfico 1: Evolução dos sujeitos nulos e plenos ao longo do tempo por data de nascimento do missivista

Os resultados gerais permitem fazer uma análise de regra variável para saber quais contextos favorecem o sujeito nulo nesta amostra. Passemos, portanto, para os resultados da regra variável: realizamos uma rodada geral, com valor de aplicação sujeito nulo, para captar os contextos que favorecem o sujeito nulo. Pode parecer contraditório numa pesquisa descritiva optar por controlar os fatores para sujeito nulo, com base num *corpus* do PB, que apresenta, sabidamente, contextos mais restritos ao sujeito nulo, do que línguas de sujeito nulo canônicas, como o Italiano e o PE. Entretanto, a ideia aqui é ver os contextos que mais favorecem o sujeito nulo para saber se eles estão relacionados, de certo modo, às análises gerativistas sobre o tema, que norteiam os grupos de fatores con-

trolados na análise. Pela quantidade de dados, optei por fazer uma rodada com todos os dados, sem separar por período de tempo. Desse modo, a data de nascimento do missivista é uma das variáveis independentes de análise. Como não fiz rodadas separadas pelo período de tempo, após apresentar a tabela com os índices e pesos relativos dos grupos selecionados.

Os fatores selecionados por ordem de mais relevante para menos relevante foram: período de tempo, animacidade/especificidade do referente, padrão estrutural e tipo de sentença. O *input* geral para o sujeito nulo foi de .541, o que espelha, de certo modo, o índice geral de sujeito nulo na amostra (54%).

O primeiro fator a ser selecionado foi período de tempo, conforme se vê com a Tabela 1 a seguir. Vemos que os pesos relativos acompanham a evolução dos índices percentuais de sujeito nulo, conforme já mostrado no Gráfico 1, e vemos a diminuição ao longo do tempo dos índices de sujeito nulo.

Tempo	Oco.	Total	%	P.R.
Período 1	31	44	70%	0.674
Período 2	46	69	67%	0.602
Período 3	75	139	54%	0.543
Período 4	492	700	70%	0.680
Período 5	111	263	42%	0.380
Período 6	76	222	34%	0.312
Período 7	124	329	38%	0.281

Tabela 1: Sujeito nulo *vs.* período de tempo (valor de aplicação sujeito nulo)

O segundo fator selecionado como favorecedor do sujeito nulo foi o traço de animacidade/especificidade do antecedente, como vemos com a Tabela 2 a seguir: notemos que os referentes com traço [-animado/-específico] favorecem o sujeito nulo, tal como ocorre nas línguas canônicas de sujeito nulo. De fato, na análise comparativa do sujeito nulo em amostras de fala do PE e do PB, Duarte (2019) mostra que o maior índice de sujeito nulo é com antecedente [-animado/-específico] e o menor índice de sujeito nulo é com antecedente [+animado/+específico] (58% *vs.* 23,5%). No PE, o antecedente [-animado/-específico] apresenta índice de 100% de sujeito nulo, ao passo que o antecedente [+animado/+específico] apresenta índice de 55% de sujeito nulo.

Podemos observar que os dados das cartas comprovam a escala de referencialidade de Cyrino, Duarte e Kato (2000), no sentido de apresentarem os índices mais altos de sujeito nulo com referentes [-animados/-específicos] (88%) e os referentes [+animados/+específicos] com índice de 50% de sujeito nulo.

Traço	Oco.	Total	%	P.R.
[-animado / -específico]	21	24	88%	0.927
[-animado / +específico]	170	226	75%	0.808
[+animado / -específico]	36	47	77%	0.725
[+animado/ +específico]	728	1469	50%	0.427

Tabela 2: Influência do traço de animacidade e especificidade para o sujeito nulo

A Tabela 2 traz a distribuição percentual do sujeito nulo seguida do Peso Relativo e podemos ver que o maior índice tanto percentual quanto de P.R. para o sujeito nulo está nos contextos em que o referente é [-animado/-específico]. Associando os resultados da Tabela 2 com o Gráfico 2 e a Tabela 3 a seguir, podemos ver como ocorre a evolução do sujeito nulo ao longo do tempo pelo traço de animacidade / especificidade:

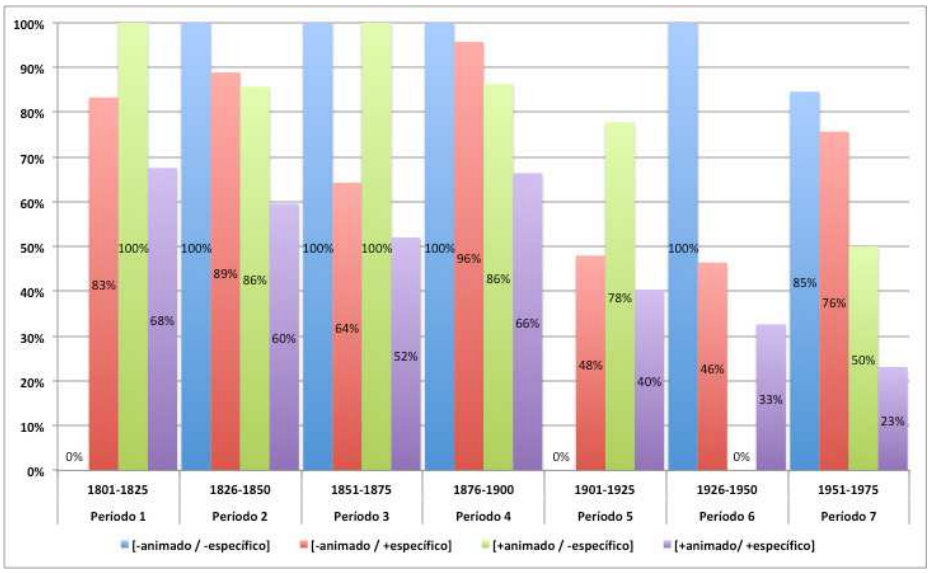


Gráfico 2: Evolução do sujeito nulo por traço de animacidade do referente por data de nascimento do missivista

		Período 1	Período 2	Período 3	Período 4	Período 5	Período 6	Período 7
		1801-1825	1826-1850	1851-1875	1876-1900	1901-1925	1926-1950	1951-1975
[-animado / -específico]	Suj Nulo	0	1	1	7	0	1	11
	Suj Pleno	0	0	0	0	1	0	2
	Total	0	1	1	7	1	1	13
	% Suj Nulo	0%	100%	100%	100%	0%	100%	85%
[-animado / +específico]	Suj Nulo	5	8	9	67	12	13	56
	Suj Pleno	1	1	5	3	13	15	18
	Total	6	9	14	70	25	28	74
	% Suj Nulo	83%	89%	64%	96%	48%	46%	76%
[+animado / -específico]	Suj Nulo	1	6	1	19	7	0	2
	Suj Pleno	0	1	0	3	2	3	2
	Total	1	7	1	22	9	3	4
	% Suj Nulo	100%	86%	100%	86%	78%	0%	50%
[+animado/ +específico]	Suj Nulo	25	31	64	399	92	62	55
	Suj Pleno	12	21	59	202	136	128	183
	Total	37	52	123	601	228	190	238
	% Suj Nulo	68%	60%	52%	66%	40%	33%	23%

Tabela 3: Evolução do sujeito nulo por traço de animacidade do referente e data de nascimento dos missivistas

Se compararmos a evolução do sujeito nulo ao longo do tempo pelo traço de animacidade, veremos que a maior queda nos índices se dá nos referentes com traço [+animado/+específico]: o percentual de sujeito nulo sai de 68% no primeiro período e vai para 33% no último período. Os índices de sujeito nulo com referente [-animado/-específico] mantêm-se altos ao longo do tempo (oscilando entre 85% e 100%), e parecem não apresentar uma curva de mudança, mas uma estabilidade ao longo do tempo.

Chamam a atenção os índices de sujeito nulo nos períodos 5 e 6, [-animado, +específico] (48% e 46%), o que pode indicar uma maior implementação da mudança com traços da gramática brasileira, uma vez que se aproximam mais dos resultados de Duarte (2019) para o PB falado, como podemos exemplificar:

(24) **Este mundo** é tão ingrato para nós dois, **ele** sabe perfeitamente o quanto nos amamos e no entanto finge não perceber. (Período 5, fundo Jayme-Maria)

(25) A tua ausencia e demais voraz, para **o meu pobre coração**, tú somente e mais ninguem é que podes avaliar a dor que **ele** sente, e o anseio que **ele** tem para tornar a sentir as tuas caricias assim que voltares, de ti é que ele espera toda a força e vigor que é para ter vida e poder viver, para amar-te eternamente. (Período 5, Jayme-Maria)

(26) Como veem, **Paris** tem suas compensações e dêste lado, realmente é fantástica a variedade de escolha que **ela** nos oferece. (Período 6, Fundo Frazão Braga)

(27) Enfim êste **diploma** está em casa. Nada de ficar andando, pois tôda a burocracia a ser feita eu já fiz. Repito êle está em casa. Bom não se afobem. Se não acharem darei um jeito, pois tenho meu amigo. (Período 6, Fundo Frazão Braga)

(28) Peça a Adt. para escrever-me de novo. Lamento bastante não ter recebido **sua carta**. De qq maneira agora acredito que [**0**] chegue. (Período 6, Fundo Frazão Braga)

O terceiro fator selecionado foi o padrão estrutural, conforme mostra a Tabela 4 a seguir. Podemos ver que os dois Padrões que mais favorecem o sujeito nulo são o Padrão A e o Padrão B, em que o sujeito nulo ou está numa sentença subordinada e seu referente está na sentença matriz (Padrão A) ou está numa matriz com o seu referente na adverbial anteposta (Padrão B), que são casos em que há c-comando entre o referente e o sujeito:

Padrão	Oco.	Total	%	P.R.
Padrão A	99	120	83%	0.797
Padrão B	33	44	75%	0.730
Padrão C	420	800	53%	0.522
Padrão D	341	666	51%	0.405
Padrão E	62	136	46%	0.457

Tabela 4: Sujeito nulo *vs.* padrão estrutural da sentença (valor de aplicação sujeito nulo)

Cabe aqui mostrar a evolução por tipo de padrão ao longo do tempo, de modo a verificar como se dá diminuição dos índices de sujeito nulo por padrão sentencial ao longo do tempo. O Gráfico 3 a seguir traz os índices de sujeito nulo (versus pleno) por tipo de padrão estrutural e data de nascimento dos missivistas:

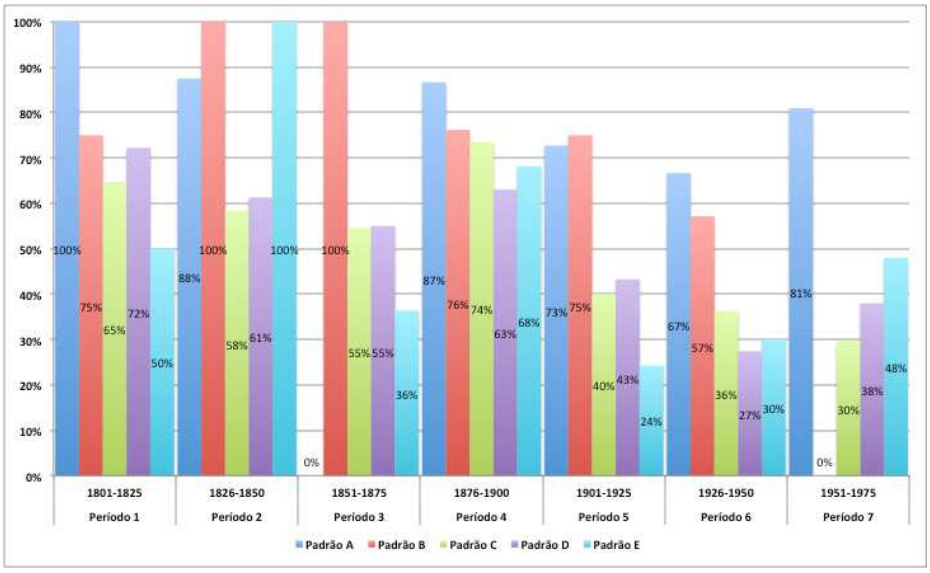


Gráfico 3: Evolução do sujeito nulo por padrão sentencial por data de nascimento do missivista.

Podemos ver que os índices de sujeito nulo diminuem ao longo do tempo, em todos os padrões sentenciais, mas de maneira diferente: por um lado, nos Padrões A e B, em que há um certo "controle" do antecedente sobre o referente do sujeito nulo, os índices são mais altos, variando entre 100%, no primeiro período e 81% no último. Por outro lado, quando o antecedente está numa posição diferente da de sujeito, Padrão D, os índices são mais baixos, saindo de 72% no Período 1 e chegando a 38% no Período 7. Podemos dizer que esses resultados quantitativos espelham as análises gerativistas sobre o licenciamento do sujeito nulo no PB, na medida em que são contextos que a literatura gerativista aponta como os únicos contextos de sujeito nulo – recorde-se que há análises que propõem que o sujeito nulo de uma encaixada é gerado via movimento da encaixada para a matriz (Ferreira, 2000; Rodrigues, 2004; Kato, Martins e Nunes, 2023). Os exemplos a seguir mostram os Padrões A e B, com destaque aos exemplos do Período 5, do casal não-ilustre, que parece ser o que mais se aproxima dos dados de PB da análise de Duarte (2019) e das análises teóricas sobre o sujeito nulo, com os índices mais altos de sujeito nulo nos Padrões A e B:

(29) Já mais agora vendo o nosso filho ocupar uma alta patente como estás ocupando mas **Deus** não quiz, [**0**] achou que lá do céu **ele** está nos vendo nos acompanhando e resando por nós todos não é meu filho (Período 4, Fundo Pedreira Ferraz Magalhães)

(30) Tenho umas novidadesinhas que muito te agradarão, no ultimo dia que subimos juntos, a noite, eu e **minha mãe** estivemos conversando, então ela falou muito a teu respeito, demonstrou estar muito interessada em voce, chegou a chorar, disse que ela e papai foram os culpados por tudo que houve, [**0**] falou tambem que [**0**] tem a certeza que nós não acabamos, e que dentro em breve voce voltará a frequentar a nossa casa. (Jayme)

(31) **Minha mãe** falou também, que não [**0**] se metia em mais nada, porque se fosse o nosso destino nos casavamos mesmo e não adiantava estar se metendo, ela culpa somente a sua irmã, a sua irmã é que foi afronta-la. (Jayme)

(32) **a Hilda** dice que [**0**] manda 10 beijo para você (Maria)

(33) **a Ismenia** quando [**0**] vai no coreio não preciza dizer mais o meu nome a mosa ja sabe (Maria)

(34) Eu a semana passada pasei muito triste e chorando um pouquinho só mais não fiques sangado com migo o que fais isto e eu gostar muito do meu Jayminho **a Ismenia** dice se [**0**] soubece escrever que [**0**] tinha escrito para voce dizendo que eu tinha chorado. (Maria)

Vejamos exemplos do Padrão D, em que o referente está numa posição diferente da de sujeito e não favorece o sujeito nulo:

(35) **O rapaz** não conhece sua futura esposa, são os pais que a escolhem. É todo um ritual; a festa foi realizada "chez lui", casa simples de terra batida e telhado de palha. Todos **o** esperam na rua e qdo êle chega cantam um hino muito bonito. (Período 6, Fundo Frazão Braga)

(36) Muito te agradeço tua **cartinha** de 10 do corrente - No meio de minhas 5 amarguras [**0**] foi um balsamo consoladôr, que por instantes me alegrou e me encheo de satisfação. (Período 4, Fundo Pedreira Ferraz Magalhães)

Cabe mencionar o Padrão E, em que o referente do sujeito é um tópico, em que os índices de sujeito nulo oscilam entre 50% ~100% ~36% ~68% ~24% ~30% ~48% nos sete períodos de tempo analisados, evidenciando uma variação estável neste contexto. Tal resultado pode apontar o que se afirma na literatura sobre o sujeito nulo ser identificado por um tópico ("*topic drop*"), o que aproximaria o PB de línguas orientadas para o tópico.

(37) Vae tambem a sua esposa **a eles** desejo que [**0**] sejam felizes e que se acostumem no clima frio e que nada lhes acuntece de alteração em sua saúde. (Período 4, Fundo Frazão Braga)

(38) **Teu Pai**, Miloca, pelo que eu vejo, não faz idéa o que seja o amôr. Te digo que, parece-me que [**elle**] não quer que tu venhas por isto que hontem quando nós fomos em tua casa elle estava só disendo para Papai: Não achas, seu Cruz, que ellas não podiam estar agora aqui? que cheiro de tinta, isto está impossivel-etc, mas não foi capaz de nos diser que tua Mãi tinha consultado a elle a respeito da volta. (Período 3, Fundo Oswaldo Cruz)

Passemos ao quarto fator selecionado como favorecedor para o sujeito nulo: o tipo de sentença. Podemos ver com a Tabela 5 a seguir que as subordinadas completivas, seguidas das relativas, são os contextos mais favorecedores do sujeito nulo:

Sentença	Oco.	Total	%	P.R.
Completiva	314	481	65%	0.580
Relativa	95	165	58%	0.510
Matriz	470	954	49%	0.470
Adverbial	76	166	46%	0.431

Tabela 5: Sujeito nulo *vs.* tipo de sentença (valor de aplicação sujeito nulo)

Esse resultado, de certo modo, está relacionado ao resultado mostrado para o Padrão sentencial, uma vez que o "controle" ocorre entre o referente na matriz e o sujeito nulo na subordinada. Considerando os índices gerais, sem comparar a evolução ao longo do tempo, vemos que os percentuais são muito próximos entre completivas e relativas e matrizes e adverbiais. Se compararmos a evolução por tipo de sentença por data de nascimento dos missivistas, podemos ver como essa mudança ocorre. Vejamos o Gráfico 4 a seguir, que traz o percentual de sujeito nulo ao longo do tempo por tipo de sentença:

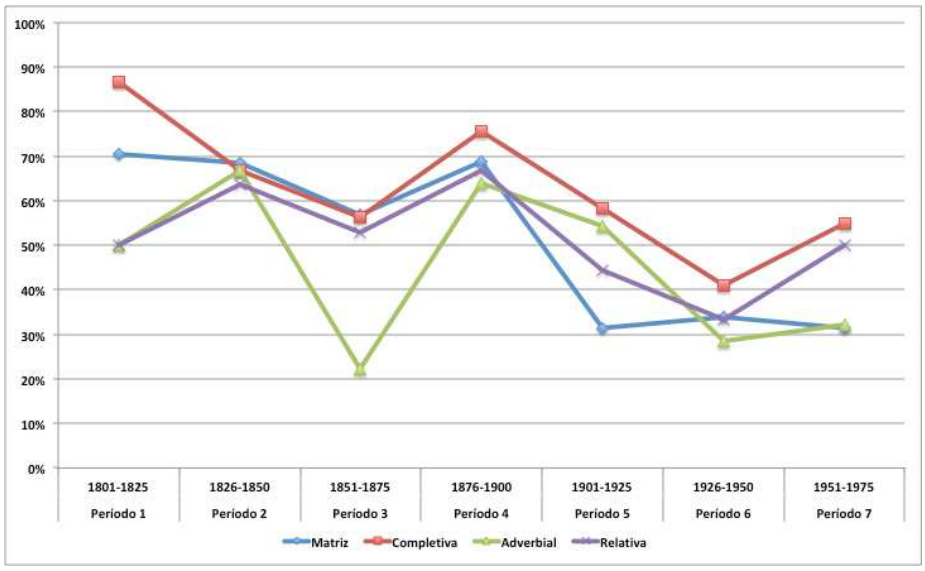


Gráfico 4: evolução do sujeito nulo por tipo de sentença por data de nascimento do missivista

Vemos que os índices de sujeito nulo diminuem ao longo do tempo, entretanto, em alguns contextos os índices são mais altos, como é o caso das completivas, que se mantêm como o contexto favorecedor para o aparecimento do sujeito nulo. A maior queda ao longo do tempo ocorre com as sentenças matrizes, em que o percentual de sujeito nulo sai de em torno de 70% e vai para em torno de 30%.

Cabe trazer aqui o percentual do sujeito nulo nas sentenças coordenadas. Conforme mencionado anteriormente, retirei da análise as sentenças coordenadas, por se tratar de um contexto em que o sujeito nulo pode ser uma elipse e licencia sujeitos nulos em línguas de sujeito não nulo. Entretanto, vale a pena comparar os índices de sujeito nulo nas orações coordenadas ao longo do tempo com as outras sentenças. Comparemos as tabelas 6.1 e 6.2 a seguir: mais uma vez, vemos os índices de sujeito nulo diminuírem ao longo do tempo, mesmo nas coordenadas; entretanto, essa diminuição ocorre de maneira distinta, dependendo do tipo de sentença. Primeiramente, comparando os percentuais de sujeito nulo das coordenadas em relação às outras sentenças, percebemos que os índices de sujeito nulo saem de 98% e chegam a 54-62% nos dois últimos períodos. Com relação às outras sentenças, nenhum dos contextos apresenta índices de sujeito nulo tão altos como esses, incluindo aí as completivas, que são as sentenças com o maior índice de sujeito nulo. A comparação apresentada aqui é importante para mostrar que até no contexto das coordenadas, ocorre a diminuição do percentual de sujeito nulo. Com relação às completivas, o contexto sintático mais favorecedor do sujeito nulo, vemos que os índices são mais altos em comparação com as outras orações, saindo de 87% no Período 1 e chegando a 55% no último período.

Com relação ao período de tempo, vemos que os Períodos 5 e 6 parecem espelhar mais de perto as descrições gerativistas para os contextos de licenciamento de sujeito nulo, uma vez que o contexto mais favorecedor para o sujeito nulo são as completivas, que o sujeito da encaixada é controlado pelo referente na matriz e os índices de sujeito nulo no geral são os mais baixos.

Tipo de Sentença / Período de Tempo		Período 1	Período 2	Período 3	Período 4	Período 5	Período 6	Período 7
		1801-1825	1826-1850	1851-1875	1876-1900	1901-1925	1926-1950	1951-1975
Matriz	Suj Nulo	12	13	42	248	45	44	66
	Suj Pleno	5	6	32	113	98	86	144
	Total	17	19	74	361	143	130	210
	% Suj Nulo	71%	68%	57%	69%	31%	34%	31%
Completiva	Suj Nulo	13	16	22	164	39	16	44
	Suj Pleno	2	8	17	53	28	23	36
	Total	15	24	39	217	67	39	80
	% Suj Nulo	87%	67%	56%	76%	58%	41%	55%
Adverbial	Suj Nulo	4	10	2	22	19	10	9
	Suj Pleno	4	5	7	13	16	25	20
	Total	8	15	9	35	35	35	29
	% Suj Nulo	50%	67%	22%	63%	54%	29%	31%
Relativa	Suj Nulo	2	7	9	58	8	6	5
	Suj Pleno	2	4	8	29	10	12	5
	Total	4	11	17	87	18	18	10
	% Suj Nulo	50%	64%	53%	67%	44%	33%	50%

Tabela 6.1: Evolução do sujeito nulo por tipo de sentença por data de nascimento do missivista

Tipo de Sentença / Período de Tempo		Período 1	Período 2	Período 3	Período 4	Período 5	Período 6	Período 7
		1801-1825	1826-1850	1851-1875	1876-1900	1901-1925	1926-1950	1951-1975
Coordenada	Suj Nulo	40	36	57	223	121	102	158
	Suj Pleno	1	4	12	27	69	86	98
	Total	41	40	69	250	190	188	256
	% Suj Nulo	98%	90%	83%	89%	64%	54%	62%

Tabela 6.2: Evolução do sujeito nulo em sentenças coordenadas por data de nascimento do missivista

Os resultados aqui apresentados revelam, de certo modo, a implementação da remarcação do Parâmetro do Sujeito Nulo na escrita: vemos a diminuição nos índices de sujeito nulo ao longo do tempo: nos Períodos 1, 2 e 3, os índices gerais de sujeito nulo (excluindo-se as coordenadas) se aproximam muito dos índices de sujeito nulo das línguas românicas, mostrados na Tabela 1 acima; a partir do Período 5, temos uma queda nos índices de sujeito nulo mais brusca, que se distancia dos índices de sujeito nulo das línguas românicas e se aproxima do PB falado atual.

Além disso, os contextos sintáticos favorecedores para o sujeito nulo se revelaram como os mesmos contextos já discutidos por Duarte (2019) nas peças brasileiras e nas amostras de fala brasileira e portuguesa: podemos ver praticamente os mesmos contextos de "resistência" do sujeito nulo, a saber:

- o traço de animacidade e especificidade nos contextos de referentes [-animado, -específico] favorece o sujeito nulo, fazendo com que os índices sejam altos nesse contexto, inclusive nos períodos em que o índice de sujeito nulo fica abaixo de 50%;
- os padrões estruturais A e B (em que o sujeito tem sua referência controlada via c-comando) favorecem o sujeito nulo; ao passo que o Padrão D (em que o referente do sujeito está menos acessível) não favorece;
- há uma diminuição geral dos índices de sujeito nulo ao longo do tempo, com a manutenção dos índices altos de sujeito nulo nas completivas e nas relativas;
- por fim, mesmo em contextos de coordenadas, que não é um contexto de sujeito nulo, mas sim de elipse, há a diminuição geral no índice de sujeito nulo.

Se voltarmos aos resultados de Vieira-Pinto (2020), que analisou os sujeitos nulos numa amostra de cartas catarinenses, vemos que os índices gerais de sujeito nulo são mais baixos na amostra de cartas cariocas do que na amostra catarinense. De certo modo, os resultados quantitativos das amostras de cartas, mesmo distantes dos resultados das

amostras de fala, apontam para a observação do encaixamento da mudança na língua escrita.

4. Considerações finais

Este trabalho trouxe uma análise do fenômeno do sujeito nulo em cartas de brasileiros nascidos entre os séculos XIX e XX, com o objetivo principal de observar o encaixamento da mudança na remarcação do PSN na escrita familiar, a fim de responder à pergunta: é possível detectar a mudança na interpretação e licenciamento desse sujeito nulo de terceira pessoa em uma amostra diacrônica de cartas pessoais?

Os resultados a que chegamos permitem indicar que a mudança na interpretação e no licenciamento do sujeito nulo começa a se implementar na amostra diacrônica, na medida em que os contextos mais favorecedores do sujeito nulo são exatamente os contextos que a literatura gerativista aponta como os contextos de sujeito nulo. Ainda não vemos, entretanto, as restrições que as análises gerativistas apontam (por exemplo, o sujeito nulo não é permitido em contextos de ilha). Desse modo, mesmo se tratando de uma amostra de cartas privadas, a hipótese de que os resultados se aproximam da gramática do PB pode ser em parte confirmada, com a diminuição dos índices gerais de sujeito nulo ao longo do tempo e com a sua restrição a contextos sintáticos específicos, em que o sujeito nulo é controlado por um antecedente na sentença matriz.

Os resultados, entretanto, não são conclusivos, tendo em vista que ainda é necessário explicar a subida no índice de sujeito nulo do Período 4: deve-se ao fato de os missivistas estarem em conventos e monastérios e há influência da norma padrão? Se acharmos mais cartas de outros Fundos, o resultado será o mesmo? Todos os missivistas desse período se comportam da mesma maneira? Enfim: precisamos mergulhar em questões mais detalhadas da sócio-história desses missivistas a fim de chegar a uma conclusão.

Além disso, vamos refinar a análise dos contextos sintáticos, a fim de aplicar com mais precisão os testes realizados nos trabalhos teóricos, incorporando as sugestões e referências bibliográficas, principalmente com relação ao controle dos padrões sentencias. Desse modo, é necessário aumentar a amostra, e, assim, aumentar o número de dados por Período, para poder realizar rodadas por período de tempo, a fim de observar quais fatores são mais relevantes em cada período e se a relevância vai mudando ao longo do tempo.

Por fim, poderemos continuar tentando responder à pergunta principal do projeto: quais são as propriedades microparamétricas da mudança na posição e na expressão dos sujeitos pronominais que estão relacionadas entre si e que diferenciam o PB das outras línguas românicas de sujeito nulo na diacronia?

Referências

BAKER, M. C. The acroparameter in a microparametric world. *In:* BIBERAUER, T. *The limits of syntactic variation*. John Benjamins, 2008.

BARBOSA, P.; DUARTE, M. E. L.; KATO, M. A. "Null subjects in European and Brazilian Portuguese". *Journal of Portuguese Linguistics*, vol. 4, n. 2, 2005. Pp 11-52.

CAVALCANTE, S. R. O. O sujeito nulo de referência indeterminada na fala culta carioca. Diadorim (Rio de Janeiro), v. 2, 2007. p. 63-82.

COELHO. I. L.; VIEIRA-PINTO, C. A.; ZIBETTI, E. M. O.; SILVA, G. M. e. Ordem SV, sujeito expresso e objeto nulo: a trajetória da mudança no português de Santa Catarina. Actas do XVIII Congresso Internacional ALFAL – Projetos, Bogotá: Universidade de Bogotá, p. 1-16, 2017.

COELHO, I. L.; CAVALCANTE, S. R. O.; VIEIRA-PINTO, C. A.; MACHADO, A. L.; CRUZ, A. B.; SILVA, G. M. A trajetória da mudança na sintaxe do sujeito e do objeto direto em cartas pessoais catarinenses e cariocas. *In:* Izete Lehmkuhl Coelho; Isabel de Oliveira e Silva Monguilhott; Marco Antonio Rocha Martins; Edair Maria Görski. (Org.). Aspectos sócio-históricos e linguísticos do português escrito em Santa Catarina nos séculos XIX e XX. 1ed.Florianópolis: Editora da UFSC, 2021, v. ? , p. 192-246.

COELHO, I. L. ; VIEIRA-PINTO, C. A. O sujeito nulo em cartas pessoais catarinenses no curso dos séculos XIX e XX (1885-1998). *In:* Juliana Esposito Marins; Mônica Tavares Orsini; Silvia Regina de Oliveira Cavalcante. (Org.). Contribuições à descrição e ao ensino do português brasileiro: da fonética ao discurso, com parada obrigatória na sintaxe. 1ed.São Paulo: Pimenta Cultural, 2021, v. , p. 281-309.

CYRINO, S.; DUARTE, M. E. L. e KATO, M. Visible subjects and invisible clitics in Brazilian Portuguese. In *In:* KATO, M.; NEGRÃO, E. (eds.). *Brazilian Portuguese and the null subject parameter*. Frankfurt: Vervuet-Iberoamericana, 2000. p. 55-73.

DUARTE, M. E. L. "Sujeitos de referência definida e arbitrária: aspectos conservadores e inovadores na escrita padrão". *Revista Linguística – Revista do Programa de Pós-Graduação em Linguística*. UFRJ, vol. 3, n. 1, 2007, pp. 89-115.

DUARTE, M. E. L. "Do pronome nulo ao pronome pleno: a trajetória do sujeito no português brasileiro". *In:* ROBERTS, I.; KATO, M. (org.) *Português brasileiro: uma viagem diacrônica*. Campinas, Ed. UNICAMP, 1993. pp. 107-128. (Reeditado pela Contexto em 2018.)

DUARTE, M. E. L. A evolução na representação do sujeito pronominal em dois tempos. *In:* M. da Conceição de Paiva; M. Eugenia L. Duarte. (Org.). Mudança Lingüística em Tempo Real. Rio de Janeiro: Contra Capa, v. 2003, p. 115-128

DUARTE, M. E. L. *A perda do princípio 'evite pronome' no português brasileiro*. 1995. 140f. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Capinas, 1995.

DUARTE, M. E. L. O sujeito nulo no português brasileiro. *In:* Cyrino, S; Torres-Moraes, M. A. (Org.). Mudança sintática do português brasileiro: perspectiva gerativista. 1a.ed. São Paulo: Contexto, v. 1, 2018. p. 26-71.

DUARTE, M. E. L. O Sujeito nulo referencial no português brasileiro e no português europeu. *In:* GALVES, C.; KATO, M.; ROBERTS, I. *Português brasileiro: uma segunda viagem diacrônica*. Campinas: Editora da Unicamp, 2019, 93-126.

DUARTE, M. E. L. The loss of the avoid pronoun principle in Brazilian Portuguese. *In:* KATO, M.; NEGRÃO, E. (ed.). *Brazilian Portuguese and the null subject parameter*. Frankfurt: Vervuet-Iberoamericana, 2000. p. 17-36.

DUARTE, M. E. L.; MOURAO, G. C.; SANTOS, H. M. Os sujeitos de 3ª. pessoa: revisitando Duarte 1993. *In:* Maria Eugenia L. Duarte. (Org.). O sujeito em peças de teatro (1833-1992): estudos diacrônicos. 1a.ed.São Paulo: parábola Editorial, 2012, v. 1, p. 21-44.

FERREIRA, M. B. *Argumentos nulos em português brasileiro*. 2000. 113f. Dissertação (Mestrado em Lingüística) – Universidade Estadual de Campinas, Campinas, 2000.

FIGUEIREDO SILVA, M. C. *A posição sujeito no português brasileiro*. Campinas: Editora da Unicamp, 1996.

GALVES, C. O enfraquecimento da concordância no Português Brasileiro. *In:* ROBERTS, I.; KATO, M. (org.) *Português brasileiro: uma viagem diacrônica*. Campinas, Ed. UNICAMP, 1993. pp. 31-50. (Reeditado pela Contexto em 2018.)

KROCH, A. Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1989, 1:199-244.

GALVES, C. Revisitando a concordância no Português Brasileiro. *In:* GALVES, C.; KATO, M.; ROBERTS, I. *Português brasileiro: uma segunda viagem diacrônica*. Campinas: Editora da Unicamp, 2019, 127-150.

GALVES, C. A gramática do português brasileiro. *Línguas e Instrumentos Lingüísticos*, n. 1, p. 79-98, 1998.

GALVES, C. *Ensaios sobre as gramáticas do português*. Campinas: Editora da UNICAMP, 2001.

HOLMBERG, A. Is there a little pro? Evidence from finnish. *Linguistic Inquiry*, v. 36, n. 4, 2005. p. 533-564.

HOLMBERG, A. "Null subject parameters". *In:* BIBERAUER, T. *et al.* (ed.) *Parametric Variation: null subjects in minimalist theory*. Cambridge, Cambridge University Press, 2010, pp. 88-124.

KATO, M. The partial *prodrop* nature and the restricted VS order in Brazilian Portuguese. *In:* KATO, M.; NEGRÃO, E. (ed.). *Brazilian Portuguese and the null subject parameter*. Frankfurt: Vervuet-Iberoamericana, 2000, p. 223-258.

KATO, M.; MARTINS, A. M.; NUNES, J. *Português brasileiro e português europeu – Sintaxe Comparada*. 1 ed., São Paulo: Contexto, 2023.

MARINS, J. E. O parâmetro do sujeito nulo: Uma análise contrastiva entre o português e o italiano. Dissertação de Mestrado (PPG-Letras Vernáculas), Faculdade de Letras / UFRJ, 2009.

MARINS, J. E.; SOARES DA SILVA, H.; DUARTE, M. E. L. "Revisiting Duarte (1995): for a gradient analysis of indeterminate subjects in Brazilian Portuguese". *Revista Diadorim*. UFRJ, vol. 19, n. especial, 2017, pp. 140-172.

MODESTO, M. Null subjects without "rich" agreement. *In:* KATO, M. A.; NEGRÃO, E. V. (eds). *Brazilian Portuguese and the null subject parameter*. Frankfurt: Vervuet-Iberoamericana, 2000. p. 147-174.

MOREIRA DA SILVA, S. Études sur la symétrie et l'asymétrie SUJET/OBJET dans le Portugais du Brésil. Tese de Doutorado, Université de Paris VIII, Departament de Linguistique Générale, 1983.

RODRIGUES, C. Brazilian Portuguese and Finnish Referential NullSubjects. *REVISTA DA ABRALIN*, v. 3, 2004. p. 75-11.

SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. *GoldVarb X*: a variable rule application for Macintosh and Windows. University of Toronto: Department of Linguistics, 2005.

SILVA, E. N.; LOPES, C. R. S. O perfil sociolinguístico de um casal não ilustre: uma análise grafemática através da edição de cartas particulares. Confluência (Rio de Janeiro), v. 1, p. 78-104, 2013.

SOARES DA SILVA, H. O Parâmetro do Sujeito Nulo: confronto entre o português e o espanhol. Dissertação de Mestrado (PPG-Letras Vernáculas), Faculdade de Letras / UFRJ, 2006.

TARALLO, F. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. *In:* ROBERTS, I.; KATO, M. A. (org.). *Português brasileiro*: uma viagem diacrônica. Homenagem a Fernando Tarallo. Campinas: Editora da Unicamp, 1993. p. 69-105.

TARALLO, F.; KATO, M. A. Harmonia transsistêmica: variação intra- e inter-linguística. *Preedição*, v. 5, Campinas, p. 315-353, 1989. [Republicado em *Diadorim: Revista de Estudos Linguísticos e Literários*, n. 2, 2007].

TORREGROSSA, J.; ANDREOU, M.; BONGARTZ, C. M. Variation in the use and interpretation of null subjects: A view from Greek and Italian. *Glossa: a journal of general linguistics*, 5(1): 95. 1–28, 2020.

VIEIRA-PINTO, C. A. Trajetória de mudança do sujeito e do objeto direto anafóricos: análise de cartas pessoais brasileiras e portuguesas dos séculos XIX e XX. Universidade Federal de Santa Catarina, Florianópolis. Tese de Doutorado. 2020.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. *Fundamentos empíricos para uma teoria da mudança lingüística*. Tradução de Marcos Bagno. São Paulo: Parábola, 2006 [1968].

http://doi.org/10.5007/1984-8420.2025.e100437

TRAJETÓRIA DE MUDANÇA DO SUJEITO E DO OBJETO DIRETO ANAFÓRICOS EM SANTA CATARINA: ANÁLISE DE CARTAS PESSOAIS DOS SÉCULOS XIX E XX[1]

CHANGE TRAJECTORY OF THE ANAPHORIC SUBJECT AND DIRECT OBJECT IN SANTA CATARINA: ANALYSIS OF PERSONAL LETTERS FROM THE 19TH AND 20TH CENTURIES

Cecília Augusta Vieira Pinto | Lattes | cecilia88augusta@gmail.com
Universidade Federal de Santa Catarina

Resumo: Nesta pesquisa, verificamos a trajetória de mudança da realização do sujeito e da realização do objeto direto anafóricos em cartas pessoais catarinenses escritas ao longo dos séculos XIX e XX, em cinco períodos de tempo, entre 1876 e 2000. Para tanto, baseamo-nos na Sociolinguística Histórica (Conde Silvestre, 2007), fundamentada na Teoria da Variação e Mudança (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972]), levando em conta que analisamos contextos internos e externos à língua que condicionam os usos variados da realização do sujeito e do objeto. Além disso, seguimos alguns postulados da Teoria de Princípios e Parâmetros (Chomsky, 1981, 1994 [1986]; Biberauer *et al.*, 2010; Galves *et al.*, 2012; Biberauer; Roberts, 2017) no levantamento de hipóteses relacionadas aos fatores internos, com base na teoria da Sociolinguística Paramétrica (*Cf.* Tarallo; Kato, 2007 [1989]). Nossa expectativa geral era a de que atestaríamos mudança em tempo real nas cartas catarinenses através da queda do sujeito nulo e do aumento do objeto nulo no curso do tempo. Os resultados a que chegamos mostram um declínio bem sutil do sujeito nulo, mais expressivo na segunda metade do século XX. Sobre o objeto, atestamos mudança em tempo real, considerando o aumento significativo de objeto nulo com o passar dos anos.

Palavras-chave: Sociolinguística Histórica. Teoria da Variação e Mudança. Sociolinguística Paramétrica. Relação entre sujeito e objeto direto no PB. Séculos XIX e XX.

[1] Este artigo parte de um recorte da tese de doutorado da autora.

http://doi.org/10.5007/1984-8420.2025.e100437

Abstract: In this research, we verified the change trajectory in the anaphoric subject and direct object realizations in *Catarinenses* (Brazilian) letters, written in the 19th and 20th centuries in five different periods of time, between 1876 and 2000. To achieve this goal, we take into account the Socio-Historical Linguistics postulates (Conde Silvestre, 2007), grounded in the Theory of Language Change (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972]), since we analyzed the internal and external contexts of the language that constrain the diverse use of the subject and object realizations. In addition, we also took into account some postulates of the Principles-and-Parameter Theory (Chomsky, 1981, 1994 [1986]; Biberauer *et al.*, 2010; Galves *et al.*, 2012; Biberauer; Roberts, 2017) to raise the hypotheses related to the internal factors, thus following the Parametric Sociolinguistics proposed by Tarallo and Kato (2007 [1989]). Our general expectation was that we would be able to show real-time change in *Catarinenses* letters due to null subject use decrease and null object use increase over time. The results showed a subtle decrease in the null subject use, especially since the second half of the 20th century. Regarding the object, we identified real-time change, considering the significant increase in the null object use over the years.

Keywords: Socio-Historical Linguistics. Theory of Language Change. Parametric Sociolinguistics. Relationship between subject and direct object in BP. 19th and 20th centuries.

Introdução

São diversos os trabalhos que descrevem a variação da realização do sujeito anafórico, bem como do objeto direto anafórico no português brasileiro (doravante PB), falado ou escrito (*Cf.* Duarte, 1989, 1993, 1995, 2018; Cyrino, 1997; Costa, 2011; Marafoni, 2004; Barbosa; Duarte; Kato, 2005; Duarte; Mourão; Santos, 2012; Vieira-Pinto, 2015; Marques de Sousa, 2017; entre outros). Em estudos diacrônicos da escrita do PB, os resultados estatísticos revelam, até o século XVIII, uma maior preferência pelo sujeito nulo e um uso bastante significativo do pronome clítico. Ao final do século XX, os percentuais se invertem e as análises sincrônicas apontam para um alto preenchimento do sujeito pronominal (principalmente de primeira e de segunda pessoas), além de uma alta frequência de objeto nulo e a quase extinção do clítico de terceira pessoa na fala.

Tarallo (1983, 1993a, 1993b) e Berlinck (1988, 1989) perceberam que existe uma certa relação com respeito à mudança desses dois fenômenos sintáticos, e isso parece se

refletir na mudança da ordem do sujeito. A partir da simplificação no paradigma flexional verbal do PB, houve um aumento significativo do preenchimento do sujeito pronominal, uma queda do preenchimento do objeto direto anafórico e uma diminuição da ordem VS transitiva. Ou seja, houve um rearranjo do sistema linguístico causado por mudanças que parecem ser encadeadas. O autor defende também que essa relação entre a realização do sujeito e do objeto é assimétrica, no sentido de que se um dos argumentos é nulo, o outro é preenchido. Essa mesma assimetria é verificada em Portugal e, neste ponto, PB e PE se assemelham. Porém, a distinção entre as duas variedades, segundo o autor, está no fato de que houve no PB uma inversão da assimetria entre sujeitos e objetos, e esse percurso de mudança durou pelo menos um século.

Além dos estudos de Tarallo e de Berlinck, temos conhecimento de alguns trabalhos que trouxeram resultados das ocorrências variáveis de sujeito e de objeto em suas amostras (Costa, 2011; Soledade, 2011; Coelho *et al.*, 2017; Othero; Spinelli 2019). Em dados do estágio atual do PB, por exemplo, Othero e Spinelli (2019) analisam entrevistas de Porto Alegre de duas diferentes sincronias (separadas por 20 anos) e verificam tanto ocorrências do sujeito de terceira pessoa quanto usos de objeto direto anafórico. Nas décadas de 1990 e de 2010, os resultados revelam, de um lado, a baixa frequência de sujeito nulo de terceira pessoa (com percentuais em 24% e 21,3%) e, de outro, os altos índices de objeto nulo anafórico (78,1% e 82,1%). Os autores mencionam que os dois fenômenos sintáticos têm andado juntos na história do PB e, por isso, não é de se estranhar que as ocorrências de sujeito e de objeto se encontrem ambas estáveis de uma sincronia para a outra.

Verificando dados do passado, Soledade (2011) também realizou brevemente a comparação entre os dois fenômenos. A autora analisou o objeto direto anafórico (de terceira pessoa) nas mesmas peças teatrais brasileiras em que Duarte (1993) estudou a realização do sujeito das três pessoas do discurso. Os percentuais gerais de sujeito nulo e de objeto nulo dos dois estudos foram apresentados por Soledade e mostram a assimetria citada por Tarallo: enquanto o sujeito nulo tem 80% de ocorrência em 1845, o objeto nulo tem apenas 32%; já no último recorte de tempo, 1992, o sujeito nulo passa a ter percentual em torno dos 25% e o objeto nulo anafórico alcança quase 99%. A partir desses resultados, Soledade nota que a implementação do sujeito pronominal coincide com o aumento do objeto nulo, e isso condiz com a inter-relação existente entre os dois fenômenos citada por Tarallo.

Esses trabalhos trazem alguns indicativos importantes da relação existente entre a trajetória de mudança do sujeito e do objeto que ainda merecem ser explorados. Nesta

pesquisa, pretendemos aprofundar a análise referente à realização do sujeito e do objeto direto anafóricos, descrevendo diacronicamente os dois fenômenos a partir das mesmas amostras e buscando a influência de variáveis linguísticas e extralinguísticas em comum na análise do sujeito nulo e do objeto nulo em cartas catarinenses.

1. Aspectos teórico-metodológicos

Este trabalho se insere no âmbito da Linguística Histórica, que estuda o desenvolvimento da língua com o passar do tempo, focando no modo como ela muda e nas condições associadas a cada mudança (Romaine, 1982; Tarallo, 1990; Labov, 1994; Mattos e Silva, 2004, 2008; Faraco, 2005; Paixão de Sousa, 2006; Conde Silvestre, 2007; entre outros). Seguimos os postulados e a metodologia da Linguística Histórica *stricto sensu* (Mattos e Silva, 2008), mais precisamente da Sociolinguística Histórica (Conde Silvestre, 2007), que leva em conta os pressupostos da Teoria da Variação e Mudança (Weinreich; Labov; Herzog, 2006 [1968]; Labov, 2008 [1972]), cujos principais fundamentos buscam explicações para os processos de variação e mudança ocorridos no passado, levando em conta fatores intra e extralinguísticos.

Assumindo uma teoria linguística formal, seguimos a Teoria Gerativa no que diz respeito à Teoria de Princípios e Parâmetros (*Cf*. Chomsky, 1981, 1994 [1986]; Biberauer *et al*., 2010; Galves *et al*., 2012; Biberauer; Roberts, 2017), tendo, assim, bases para levantar hipóteses internas sobre nossos fenômenos e atestar as restrições que atuam no condicionamento de nossas variantes. Além disso, adotamos neste trabalho o modelo teórico que une a Teoria Gerativa à Teoria da Variação e Mudança, em uma harmonia trans-sistêmica, a Sociolinguística Paramétrica (Tarallo; Kato, 2007 [1989]). A ideia é pensar a variação e a mudança linguística ocorridas em uma mesma língua, no decorrer dos séculos, da mesma forma que a variação entre diferentes línguas, como propõe a Teoria Gerativa.

Nossos dados foram coletados em um conjunto de 17 amostras de cartas catarinenses (pertencentes ao acervo do PHPB-SC)[2], totalizando 491 missivas. Tais cartas foram distribuídas, de acordo com as datas em que foram escritas, em cinco diferentes períodos de tempo dos séculos XIX e XX, contendo cada um deles 25 anos (1876-1900, 1901-1925, 1926-1950, 1951-1975 e 1976-2000). Levantamos todas as ocorrências de sujeito de terceira pessoa, bem como todas as ocorrências de objeto direto anafórico na escrita dos missivistas.

[2] Para a Sociolinguística Histórica, é necessária a investigação das amostras a serem analisadas a fim de buscar a reconstrução do contexto social dos missivistas e o contexto histórico da sociedade em que viviam na época de escrita das cartas. Para conhecer melhor os fatores sociais do conjunto de amostras utilizadas neste trabalho, sugerimos a leitura de Vieira-Pinto (2020), em que buscamos diversas informações que poderiam estar influenciando o uso de nossas variantes.

São duas as variáveis dependentes a serem analisadas neste estudo. A primeira diz respeito à variação do sujeito anafórico, que ocorreu em duas variantes:

Sujeito Nulo

(1) **Maria e Alexandre** vão bem, sómente muito brabos, Ø me deixam quase louca, ainda agora estou escrevendo esta ao som do choro de um dum lado e de outro do outro lado, o que, aliaz, hão de ver pela lettra. (Carta pessoal de Santa Catarina, Amostra Bertaso, Remetente Elza, ano de 1929)

Sujeito pleno

(2) No seu explêndido artigo, você mostra desconhecer o primeiro livro de **Lausimar** "Confidências" poesias, 1942, dedicado a Francisco Barreiros Filho que foi um ídolo pra ela como foi para mim. Ele nunca a recebeu - e ***ela*** tinha uma grande mágoa por não ter recebido nunca palavra dele. (Carta pessoal de Santa Catarina, Amostra Maura de Senna Pereira, ano de 1979)

A segunda variável dependente é a variação do objeto direto anafórico, que também ocorreu em duas variantes:

Objeto nulo

(3) falando nisso vê se me paga **as apostas que ganhei**, podes mandar Ø pelo correio. (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente HB, ano de 1981)

Objeto preenchido (podendo ocorrer nas formas de pronome clítico e de pronome reto)

(4) **pronome clítico:** Tencionava enviar-te **Gazetas de Noticias** por este correio, mas esqueci-***as*** em casa, e por isso só t'as remette rei pelo vapor de 12 do decorrente. (Carta pessoal de Santa Catarina, Amostra Virgílio Várzea, ano de 1890)

(5) **pronome reto:** bem, agora tenho que escrever para **as madames do 342**. Estou bem brava com elas, pois nem se despediram de mim, quando fui procurar ***elas*** domingo a tarde já tinham ido. (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente FL, ano de 1980)

A expectativa geral é a de que haverá diminuição do percentual de sujeito nulo nas cartas à medida que, em um movimento oposto, o objeto nulo for ganhando força, invertendo a assimetria inicial. Sobre o pronome reto na função de acusativo, esperamos que apareça timidamente nas amostras, em sincronias mais recentes (*Cf.* Duarte, 1993; Cyrino, 1997; Costa, 2011; Duarte; Mourão; Santos, 2012; Marques de Sousa, 2017; entre outros).

As variáveis dependentes, realização do sujeito anafórico e realização do objeto direto anafórico, foram controladas a partir de variáveis independentes linguísticas e extralinguísticas, apontadas em estudos anteriores como sendo condicionadoras da variação dos fenômenos investigados. As variáveis intralinguísticas são: "animacidade do referente", "especificidade do referente", "estrutura paralela entre referente e anafórico", "transitividade verbal e estrutura projetada pelo verbo", "construção sintática" e "padrão sentencial". Quanto às variáveis extralinguísticas, controlamos: "relação de intimidade", "temática do trecho da carta", "idade", "notabilidade do missivista" e "missivista".

De modo geral, esperamos apontar a influência das mesmas variáveis linguísticas e extralinguísticas na análise do sujeito nulo e do objeto nulo em um mesmo espaço de tempo.

2. Resultados e discussões

2.1 A realização do sujeito anafórico nas cartas catarinenses

A coleta de dados de sujeito anafórico nas cartas catarinenses resultou em 613 dados[3] em missivas datadas entre 1883 e 1993, dos quais 389 (63,5%) são de sujeito nulo e 224 (36,5%) de sujeito pleno, conforme mostra a Tabela 1:

Tabela 1: Frequência e percentual das variantes do sujeito anafórico nas cartas catarinenses segundo cinco períodos de tempo dos séculos XIX e XX

Sujeito	Período 1 1876-1900	Período 2 1901-1925	Período 3 1926-1950	Período 4 1951-1975	Período 5 1976-2000	Total
	Apl/Total %	Apl/Total %	Apl/Total %	Apl/Total %	Apl/Total %	Apl/Total %
Nulo	82/119 68,9%	34/56 60,7%	88/129 68,2%	36/54 66,7%	149/255 58,4%	389/613 63,5%
Pleno	37/119 31,1%	22/56 39,3%	41/129 31,8%	18/54 33,3%	106/255 41,6%	224/613 36,5%

Fonte: Elaborado pela autora.

[3] Optamos por retirar de nossas rodadas 126 dados de sujeito anafórico de segundas coordenadas, em que o sujeito da primeira oração possui a mesma referência do anafórico. Conforme Duarte (1995, 2018), até mesmo em línguas de sujeito preenchido, o sujeito anafórico de estruturas desse tipo costuma ser nulo.

Nossa hipótese inicial, baseada em Duarte (1993, 2018), era que, nas primeiras três sincronias, o sujeito nulo seria bastante frequente e que seus percentuais cairiam a partir do quarto período, havendo indícios de mudança em tempo real de um sistema majoritariamente de sujeito nulo para um outro, mais variável, de sujeito nulo e pleno. Atestamos parcialmente tal hipótese: somente na última sincronia vemos uma diferença no percentual de sujeito nulo, ainda pequena para atestarmos mudança. A preferência pelo sujeito nulo nas cartas catarinenses ainda é grande, apesar da queda de 10% da primeira para a última sincronia. Os exemplos a seguir ilustram a ocorrência de sujeito nulo e sujeito pleno, respectivamente.

(6) Dia 6 chegou aqui, nossa D. D. Superiora Geral. É boa, o quanto um ente humano póde ser bom. Falei particularmente com **ella**, por ser uma alumna de Curityba. Ø Deu-me "tres bençans", e no dia de meu anniversario deu-me uma imagem. (Carta pessoal de Santa Catarina, Amostra Bertaso, Remetente Elza, ano de 1920)

(7) Logo que recebi sua carta, comuniquei-me com **seu filho Jócio** e solicitei tudo o que êle pudesse conseguir. (Carta pessoal de Santa Catarina, Amostra Maura de Senna Pereira, ano de 1971)

Para a realização das rodadas estatísticas binomiais, decidimos amalgamar os períodos 2 e 3 e os períodos 4 e 5 para, assim, chegarmos a resultados mais expressivos a respeito da trajetória de mudança de nossos fenômenos[4]. Na análise multivariada das 613 ocorrências de sujeito anafórico, tomamos como aplicação da regra a variante sujeito nulo.

Em uma rodada binomial geral[5], observamos que a variável "período de tempo" não foi selecionada pelo programa estatístico como condicionadora do sujeito nulo, não atestando, portanto, mudança em tempo real nas cartas catarinenses, o que vai de encontro ao que havíamos previsto. O programa estatístico Goldvarb X selecionou apenas três variáveis intralinguísticas e uma variável extralinguística como condicionadoras do sujei-

[4] Repensamos a distribuição dos períodos de tempo de nossas amostras, considerando a pouca quantidade de dados coletados nos períodos 2 e 4 e a grande quantidade de variáveis independentes que temos, pois, se mantivéssemos dessa forma, os resultados poderiam não ser tão significativos estatisticamente. Pudemos atestar isso em uma primeira rodada em que, por exemplo, não foi selecionada nenhuma variável como relevante no período 2.

[5] Em Vieira-Pinto (2020), realizamos também rodadas estatísticas separadas a cada período a fim de verificar se as forças internas e externas são as mesmas conforme se avança no tempo.

to nulo, nesta ordem: i) padrão sentencial; ii) transitividade verbal e estrutura projetada pelo verbo; iii) relação de intimidade; e iv) especificidade do referente. Os resultados estão apresentados a seguir, na Tabela 2:

Tabela 2: Frequência, percentual e peso relativo de sujeito nulo nas cartas catarinenses segundo as variáveis "padrão sentencial", "transitividade verbal e estrutura projetada pelo verbo", "relação de intimidade" e "especificidade do referente"

Sujeito	**Todos os períodos juntos**		
	Apl/Total	**%**	**PR**
Padrão sentencial			
Padrão 1	38/41	92,7%	0,87
Padrão 2	184/260	70,8%	0,58
Padrões 3 e 4	166/309	53,7%	0,36
Total	388/610	63,6%	--
Transitividade verbal e estrutura projetada pelo verbo			
Estruturas complexas	44/58	75,9%	0,65
Verbos de 2 complementos	76/107	71%	0,58
Verbos de 1 complemento	163/261	62,5%	0,43
Total	283/426	66,4%	--
Relação de intimidade			
Menos íntimos	39/51	76,5%	0,68
Mais íntimos	341/552	61,8%	0,48
Total	380/603	63%	--
Especificidade do referente			
[-específico]	50/66	75,8%	0,65
[+específico]	339/547	62%	0,48
Total	389/613	63,5%	--
Input 0.674 Log likelihood: -370.275 Significance: 0.021			

Fonte: Elaborado pela autora.

No que diz respeito à variável "padrão sentencial", selecionada em primeiro lugar nesta rodada, nossos resultados corroboram os alcançados por Barbosa, Duarte e Kato (2005) e Duarte, Mourão e Santos (2012). Segundo tais trabalhos, o padrão sentencial 1 – aquele em que o referente se encontra no mesmo período do anafórico e é o sujeito da oração precedente (principal ou subordinada) – e o padrão sentencial 2 – em que o referente é um sujeito ou tópico e se encontra na oração adjacente – condicionam o uso

de sujeito nulo. Em contrapartida, o padrão sentencial 3 – aquele em que o referente é o sujeito de uma oração não adjacente no contexto precedente – e o padrão sentencial 4[6] – em que o referente, posicionado na oração precedente/adjacente, possui função sintática diferente da de sujeito – desfavorecem o uso do apagamento.

Atestando nossa hipótese, o padrão sentencial 1 (exemplo 8 a seguir) teve 0,87 de peso relativo. O padrão 2 (exemplo 9) também foi importante, com 0,58 de PR. Esses dois fatores ficaram bem distantes dos padrões 3 (exemplo 10) e 4 (exemplo 11), que tiveram juntos apenas 0,36 de peso relativo, índice relativamente baixo, conforme havíamos previsto. Exemplos desses quatro padrões estão elencados a seguir:

(8) Pobre Heloisa! esse segundo filho deixou-a em pandarecos. E **ella**, [inint.] e boa, me disse que Ø quer ainda outro filho, um menino. (Carta pessoal de Santa Catarina, Amostra Virgílio Várzea, ano de 1941)

(9) O doutor Francisco Galoti não está atualmente no Rio. Ø Acha-se a serviço em um dos Estados do norte. (Carta pessoal de Santa Catarina, Amostra Diversos para Virgílio Várzea e José Boiteux, Remetente afilhado B, ano de 1932)

(10) Agora voltamos ao trabalho: **Um dos donos do Laboratório** pediu-me para preparar uma aula sobre qualquer assunto que eu preferir, para ministrá-la para alguns funcionários (técnicos), o qual ***ele*** faz questão de assistir. Veja como estou desesperada!! (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente MR, ano de 1984)

(11) Recebemos há dois dias a carta da Elza, aqual muito nos alegrou em saber que vão bem de saude e que Maria já sabe Dizer bobo. Então não sahiu do plano **do papae**, porque ***elle*** disse que Maria ia aprender a diser por primeiro outro nome. (Carta pessoal de Santa Catarina, Amostra Bertaso, Remetente Jayme, ano de 1928)

A variável "transitividade verbal e estrutura projetada pelo verbo" foi selecionada como a segunda mais importante para a variação do sujeito nas amostras, tendo as estruturas complexas, com 0,66 de peso relativo, como as condicionadoras do sujeito nulo (*Cf.* exemplo 12)[7]. As estruturas simples com verbos de dois complementos também foram

[6] Por conta da baixa ocorrência do padrão 3, principalmente quando separamos os dados em três períodos, decidimos amalgamá-lo com o padrão 4 em todas as rodadas estatísticas deste trabalho.

[7] Não tínhamos uma hipótese bem formulada para essa variável, pois a trouxemos dos estudos que trataram da variação do objeto, considerando que os estudos de sujeito de que temos conhecimento não focaram

significativas, alcançando 0,58 de peso relativo (exemplo 13). Já as estruturas simples de um complemento (exemplo 14) são as que menos favorecem o sujeito nulo, com 0,43 de peso relativo.

(12) Estive no médico: meu estômago e fígado estão um pouco atacados. **O médico** aconselhou-me a me moderar nos estudos, pois Ø encontrou-me um pouco esgotado. (Carta pessoal de Santa Catarina, Amostra Arduíno Salami, ano de 1974)

(13) Peço **ao Papae** que quando vier a São Paulo Ø me traga o meu relogio. (Carta pessoal de Santa Catarina, Amostra Bertaso, Remetente Serafim, ano de 1922)

(14) Quanto a Biba, **ela** me procurou no final de janeiro, mas não pudemos nos encontrar. Achei mesmo que ***ela*** tinha uma grande novidade para contar. (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente MR, ano de 1984)

No exemplo (12), a estrutura é complexa porque o verbo é transitivo direto e a estrutura projetada por ele é de objeto direto (nesse caso, [me]) mais o predicativo do objeto [um pouco esgotado]. Em (13), temos uma estrutura simples de dois complementos, em que um deles é o objeto direto [meu relógio] e o outro é o objeto indireto [me]. A estrutura simples de um complemento em (14) ilustra o verbo transitivo direto [tinha] seguido do objeto direto [uma grande novidade].

Quanto à "relação de intimidade", atestamos a nossa hipótese, baseada em Coelho e Nunes de Souza (2014), de que os missivistas menos íntimos de seus destinatários utilizariam mais sujeito nulo do que aqueles que possuíssem uma relação mais próxima. A relação de menos intimidade teve 0,68 de peso relativo, contra 0,48 para a relação mais íntima. O exemplo de sujeito nulo em (15), a seguir, foi retirado da Amostra Arduíno Salami, que tem uma relação de menos proximidade com o Bispo de Lages, a quem presta contas de seus estudos missionários. Já o exemplo de sujeito pleno em (16) foi coletado na Amostra Oscar Rosas, que escreve para Cruz e Sousa, com quem teve uma amizade significativa e duradoura.

apenas em verbos transitivos diretos e bitransitivos, como fazemos aqui. No entanto, em todas as rodadas que fizemos em Vieira-Pinto (2020), em que a variável "transitividade verbal e estrutura projetada pelo verbo" foi selecionada, a tendência foi sempre a mesma (ao contrário do esperado para o objeto): as estruturas complexas condicionam o sujeito nulo e as estruturas simples de um complemento o desfavorecem.

(15) **Meu mano** tem leucemia no sangue, isto é, "câncer no sangue." Já foi desenganado pelos médicos, tanto pode morrer hoje, como amanhã. Ø Já recebeu os Santos Óleos, confessou-se, comungou, assistiu a missa varias vêzes. (Carta pessoal de Santa Catarina, Amostra Arduíno Salami, menos intimidade, ano de 1975)

(16) Dirás ao Varzea que me escreva [...]. Nas coudelorias fluminenses pairam agora duas eguas, pergunta**lhe** se ***elle*** as montou, porque eu ainda não. (Carta pessoal de Santa Catarina, Amostra Oscar Rosas, mais intimidade, ano de 1888)

A variável "especificidade do referente" foi a última a ser selecionada nesta rodada, tendo como condicionador do sujeito nulo o referente com traço [-específico], com 0,65 de peso relativo, atestando nossa hipótese e corroborando estudos realizados anteriormente (Duarte; Mourão; Santos, 2012; Duarte, 2018). O referente com traço [+específico] teve 0,48 de peso relativo, desfavorecendo levemente o uso de sujeito nulo. Os dados a seguir trazem exemplos de referente [-específico] e [+específico], respectivamente.

(17) **Essa gente** é chamada de "farofeiros" porque Ø comem farofa, isto é, comida misturada com farinha, em geral galinha ou carne assada, para melhor se conservar e não derramar o molho. (Carta pessoal de Santa Catarina, Amostra Harry Laus, ano de 1987)

(18) Eu não pensava em fazer mais nada **pela Academia**, depois que ***ela*** me magoou profundamente em 82. Faço uma exceção a teu respeito. (Carta pessoal de Santa Catarina, Amostra Maura de Senna Pereira, ano de 1984)

2.2 A realização do objeto direto anafórico nas cartas catarinenses

A análise do objeto direto anafórico nas cartas catarinenses resultou na coleta de 627 dados, em cartas escritas entre 1880 e 1998, em que grande parte é preenchida pelo pronome clítico (389 ocorrências, o equivalente a 62%). A variante inovadora objeto nulo aconteceu 233 vezes, alcançando 37,2% no geral. Já a outra variante nova, pronome reto na função acusativa, só ocorreu em cinco dados (0,8%). A distribuição dessas variantes nos cinco períodos de tempo está exposta na Tabela 3, a seguir.

Tabela 3: Frequência e percentual das variantes do objeto direto anafórico nas cartas catarinenses segundo cinco períodos de tempo dos séculos XIX e XX

Objeto	Período 1 1876-1900	Período 2 1901-1925	Período 3 1926-1950	Período 4 1951-1975	Período 5 1976-2000	Total
	Apl/Total %	Apl/Total %	Apl/Total %	Apl/Total %	Apl/Total %	Apl/Total %
Clítico	131/178 73,6%	35/56 62,5%	70/100 70%	46/95 48,4%	107/198 54%	389/627 62%
Nulo	47/178 26,4%	20/56 35,7%	30/100 30%	49/95 51,6%	87/198 44%	233/627 37,2%
Pronome reto	Sem dado	1/56 1,8%	Sem dado	Sem dado	4/198 2%	5/627 0,8%

Fonte: Elaborado pela autora.

Nossa expectativa geral era a de que o pronome clítico teria altos índices, principalmente nas três primeiras sincronias aqui analisadas, enquanto veríamos o objeto nulo ganhando força nos dois últimos recortes de tempo. A partir dos percentuais apresentados, nossa hipótese foi atestada, corroborando diversos estudos diacrônicos que analisaram o objeto direto no PB (*Cf.* Cyrino, 1997; Costa, 2011; Soledade, 2011; Marques de Sousa, 2017; entre outros). Seguem exemplos dessas duas variantes:

(19) **A tua Lilly** emigrou, doce pássaro d'amor, para esta tumultuosa cidade. Hoje vou vê-***la*** e à mãe e as flores que elas espalharam pela tua lembrança e pelo teu coração, eu farei com que cheguem ainda vivas e cheirosas junto de ti. (Carta pessoal de Santa Catarina, Amostra Cruz e Sousa, ano de 1889)

(20) Há meses escrevi ao amigo sobre assuntos que em Araranguá haviamos conversado, não tive resposta ainda; mandei lhe **um folhete de meu "Programa de Governo"** tambem não sei se recebeu Ø! (Carta pessoal de Santa Catarina, Amostra Diversos para Virgílio Várzea e José Boiteux, Remetente Alcebiades Seára, ano de 1927)

Sobre o pronome reto na função de acusativo, esperávamos que sua ocorrência fosse tímida nas cartas catarinenses e que aparecesse principalmente nos períodos de tempo mais recentes. Isso realmente aconteceu: foram constatados apenas quatro dados dessa variante nas cartas do período 5, provenientes da amostra Medeiros, de duas missivistas

jovens, pessoas não públicas, que escrevem para sua amiga JT, cujas relações são de mais intimidade:

(21) O **I** almoçou aqui em casa domingo. Parece mentira eu ver ***ele*** sentado ao lado de meu pai conversando. Acho que ele já conquistou a família. (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente FL, ano de 1980)

(22) bem, agora tenho que escrever para as madames do 342. Estou bem brava **com elas**, pois nem se despediram de mim, quando fui procurar ***elas*** domingo a tarde já tinham ido. (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente FL, ano de 1980)

(23) Es estou chateada com o acidente **do ex-namorado da K**, gostaria muito se vocês ajudassem ***ele***. (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente FL, ano de 1980)

(24) Com relação a ida a Lages, já não sei mais, acho que não tenho estrutura para baile no momento, pois es- tou num astral meio ruim e muito abalada com a morte **da Dona Maria**. Parece que não é real, ainda não consegui me convencer. Eu adorava ***ela***, JT! (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente MR, ano de 1984)

O pronome reto em (21) – além de fazer referência a um nome [+animado], [+específico], que está em uma posição não paralela ao anafórico – ocorre em uma estrutura complexa de um verbo sensitivo [ver], que projeta uma oração completiva de particípio [sentado]. Nessa construção, o anafórico [ele] recebe marcação excepcional de caso acusativo. Esse tipo de estrutura complexa, com tais traços semânticos e posição sintática do referente, tende a condicionar o uso de objeto preenchido. Os dados em (22), (23) e (24) já ocorrem em estruturas simples de um complemento, que tenderiam ao uso de objeto nulo; porém, os referentes são [+animado] e [+específico], fortes condicionadores do preenchimento do objeto.

Além dessas quatro ocorrências, fomos surpreendidos com um dado de pronome reto acusativo que aconteceu no início do século XX, mais especificamente em 1902:

(25) O Senhor Jose Ribeiro e Manoel Ribeiro procurarão nos para nos se [inpenhar] com o Senhor para não mandar Sitar ***elles*** que ate o dia 15 de Novembro prochimo que ate esse dia elles lhe remete os Dinheiro de seus impostos de

> capital e assim sendo possível vos pedimos e esperamos ser atendido. (Carta pessoal de Santa Catarina, Amostra Athaÿde, Remetente VA, ano de 1902)

Tal dado advém de um missivista da amostra Athaÿde, cuja idade é por nós desconhecida, que escreve no período 2 para seu chefe sobre o trabalho de contador. As cartas de VA sugerem que sua relação com o chefe é de menos intimidade, o que reforça a nossa surpresa com a presença desse dado inovador. Importante mencionar que esse pronome reto faz referência a um SN [+animado] e [+específico], que é argumento externo, e que o anafórico ocorre em uma estrutura complexa: trata-se de um verbo bitransitivo [sitar] que projeta um objeto direto [elles] e um complemento circunstancial (que nesse caso é nulo). Todos esses fatores juntos influenciam o uso do objeto preenchido, apesar de não ser pela variante canônica pronome clítico.

Seguindo a mesma metodologia nos dois fenômenos estudados nesta pesquisa, passamos a considerar amalgamados os dados dos períodos 2 e 3 e dos períodos 4 e 5 nas rodadas binomiais que apresentamos em seguida. Nas rodadas binomiais do objeto, decidimos retirar os cinco dados de pronome reto acusativo, totalizando 622 dados: 233 objetos nulos (37,5%) e 389 clíticos (62,5%). Em uma primeira rodada, com todos os dados juntos, tínhamos a pretensão de testar se o período de tempo seria selecionado. Nesta ordem, o programa Goldvarb X elegeu como condicionadores do objeto nulo: a "animacidade do referente", o "padrão sentencial", o "período de tempo" e a "especificidade". Os resultados estão apresentados a seguir, na Tabela 4:

Tabela 4: Frequência, percentual e peso relativo de objeto nulo nas cartas catarinenses de todos os períodos juntos segundo as variáveis "animacidade do referente", "padrão sentencial", "período de tempo" e "especificidade do referente"

Objeto	**Todos os períodos**		
	Apl/Total	**%**	**PR**
Animacidade do referente			
[-animado]	212/440	48,2%	0,63
[+animado]	21/182	11,5%	0,21
Total	233/622	37,5%	--
Padrão sentencial			
A – Tópico ou complemento no período adjacente	176/378	46,6%	0,58
B/C – Complemento ou sujeito na oração coordenada	20/58	34,5%	0,48

S – Sujeito no período adjacente	11/46	23,9%	0,45
D/E – Complemento ou sujeito na oração precedente (principal ou subordinada)	16/78	20,5%	0,31
F/G – Complemento ou sujeito na oração não adjacente (com orações intervenientes)	10/58	17,2%	0,28
Total	233/618	37,7%	--
Período de tempo			
Período 1	47/178	26,4%	0,39
Período 2/3	50/155	32,3%	0,47
Período 4/5	136/289	47,1%	0,57
Total	233/622	37,5%	--
Especificidade do referente			
[-específico]	57/105	54,3%	0,60
[+específico]	176/517	34%	0,47
Total	233/622	37,5%	--
Input 0.332 Log likelihood: -345.280 Significance: 0.036			

Fonte: Elaborado pela autora.

Conforme esperávamos, a "animacidade do referente" tem grande importância na variação do objeto direto, seguindo o que foi atestado em diversos estudos sincrônicos e diacrônicos do PB (Duarte, 1989; Cyrino, 1997; Marafoni, 2004; Costa, 2011; -Pinto, 2015; Marques de Sousa, 2017; entre outros): o referente [-animado] condiciona o objeto nulo em 0,63 de peso relativo, contra 0,21 para o referente [+animado]. Os exemplos a seguir representam dados de referente [-animado], que influencia o uso de objeto nulo anafórico, e de referente [+animado], favorecendo o pronome clítico.

(26) Precisaria esse dinheiro, ao menos uma parte, pois deverei dar uma entrada. Pensei o seguinte: se julgarem necessário, me enviem **umas notas promissórias**, que eu assinaria Ø. Depois de Padre pagaria 100,00 por há, até saldar a dívida. (Carta pessoal de Santa Catarina, Amostra Arduíno Salami, ano de 1975)

(27) Não imaginas como te ficarei grato ao saber que te interessaste **pelo marido de minha querida Desdemona**, ajudando-***o*** no que te fór possivel, para que possa sahir-se bem no concurso, e ser nomeado. (Carta pessoal de Santa Catarina, Amostra Araújo Figueredo, ano de 1912)

Para o "padrão sentencial" relacionado ao fenômeno do objeto direto, baseamo-nos nos fatores do estudo de Marques de Sousa (2017). A principal hipótese era a de que o padrão A, em que o referente é um complemento ou um tópico estrutural/discursivo no período adjacente, condicionaria o objeto nulo. Atestamos nossa hipótese, considerando o peso relativo de 0,58 desse padrão, como no exemplo:

(28) Junto envio **uma carta** para o teu irmão – O Sinhô, e supplico lhe remetteres Ø por portador seguro. (Carta pessoal de Santa Catarina, Amostra Tota, ano de 1926)

Atestamos também nossa hipótese de que os padrões D e E (em que o referente está em uma oração precedente no mesmo período composto por subordinação, com função de complemento ou de sujeito) e os padrões F e G (em que o referente complemento ou sujeito está em uma oração não adjacente) seriam desfavorecedores do objeto nulo. Os padrões D e E amalgamados tiveram 0,31 de peso relativo para o objeto nulo, enquanto os padrões F e G tiveram 0,28. Exemplos desses quatro padrões estão elencados a seguir:

(29) Grande parte da edição está imobilizada em Florianópolis e estou tratando de recolhê-**la** para, possivelmente em 1987, reuni-***la*** com o Zenão e o Santo para uma reedição. (Carta pessoal de Santa Catarina, Amostra Harry Laus, ano de 1986)

(30) Vão fazendo companhia a esta **dous sonetos meos**, que não ***os*** quero publicar sem que primeiro tu os aprecie; se é que n'elles há alguma cousa digna disso. (Carta pessoal de Santa Catarina, Amostra Oscar Rosas, ano de 1883)

(31) Faço-lhe este pedido porque entre as autoridades policiaes do districto, há um desaffecto **de Antonio Borges,** capaz de prevalecer-se do cargo para perseguir***l-o*** e vexal-o. (Carta pessoal de Santa Catarina, Amostra Athaÿde, Remetente JL, ano de 1902)

(32) **Ele** colocou que tenho todos os motivos para não aceitá-***lo*** de volta, mas que se eu acreditar nele (mais uma vez!) só poderei ser feliz. (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente MR, ano de 1984)

Em (29), temos ilustrado o padrão D: o referente [la] exerce a função de complemento de [recolher] e está posicionado na sentença precedente que liga a oração subordinada adverbial em que o anafórico está [para reuni-la com o Zenão...]. O padrão E em (30) representa o anafórico [os] posicionado na oração subordinada relativa, que retoma

o sujeito da oração principal [dous sonetos meos]. No exemplo (31), vemos uma oração interveniente entre o referente complemento [de Antonio Borges] e o anafórico [lo], que se encontra em uma oração subordinada adverbial [para perseguil-o]. A interveniência também aparece em (32), porém a função do referente [ele] é de sujeito da oração principal.

O fato de a variável "período de tempo" ter sido selecionada, destacando a última sincronia como condicionadora do objeto nulo em oposição às sincronias mais afastadas, atesta a nossa hipótese de mudança em tempo real relacionada ao objeto nulo nas cartas catarinenses (*Cf*. Labov, 1994).

O período 1 apresentou 0,39 de peso relativo, desfavorecendo o uso de objeto nulo. Em vez dessa variante, os missivistas do século XIX preferem utilizar o pronome clítico. Já o período 4/5 tem 0,57 de peso relativo para o objeto nulo, revelando a evolução dessa variante inovadora no decorrer do tempo e sua alta utilização na escrita dos catarinenses da segunda metade do século XX.

A "especificidade do referente" foi a quarta variável selecionada pelo Goldvarb X, e os resultados atestaram a hipótese de que os referentes [-específico] influenciariam o uso de objeto nulo, com 0,60 de peso relativo (*Cf*. Cyrino, 1997; Marafoni, 2004; Costa, 2011; Marques de Sousa, 2017; entre outros).

(33) Isto escrito, são apenas **palavras**, mas são verdadeiras, sinceras, porque vem de dentro. (<há ↑> se eu pudesse dizer Ø pessoalmente) (Carta pessoal de Santa Catarina, Amostra Medeiros, Remetente HB, ano de 1981)

Quanto aos referentes [+específico], estes não favorecem o objeto nulo, porém não o desfavorecem em grande escala, considerando o peso relativo em 0,47, isto é, o pronome clítico ocorre com mais frequência nesse caso, mas também é possível o apagamento ocorrer com certa frequência, como nos exemplos.

(34) Desejo-te saude e felicidade e bons [inint.] ao Victor e minha querida sobrinha a quem peço dar um beijo por mim nos labios daquella santinha, pedindo que mude o nome de Celestina que tem tenção pôl-a, visto ser um nome muito feio; em uma boneca q'dou-lhe de lembrança designo-**lhe** o nome que devem chama***l-a*** (Basilissa). (Carta pessoal de Santa Catarina, Amostra Brito, Remetente José Maria, ano de 1883)

(35) Sei que nesta altura do tempo, talvez não mais esperasse **minha carta**; porque eu egoisticamente deixei para escrever Ø somente hoje por ser um dia abor-

recido. (Carta pessoal de Santa Catarina, Amostra Vale, Remetente E, ano de 1965)

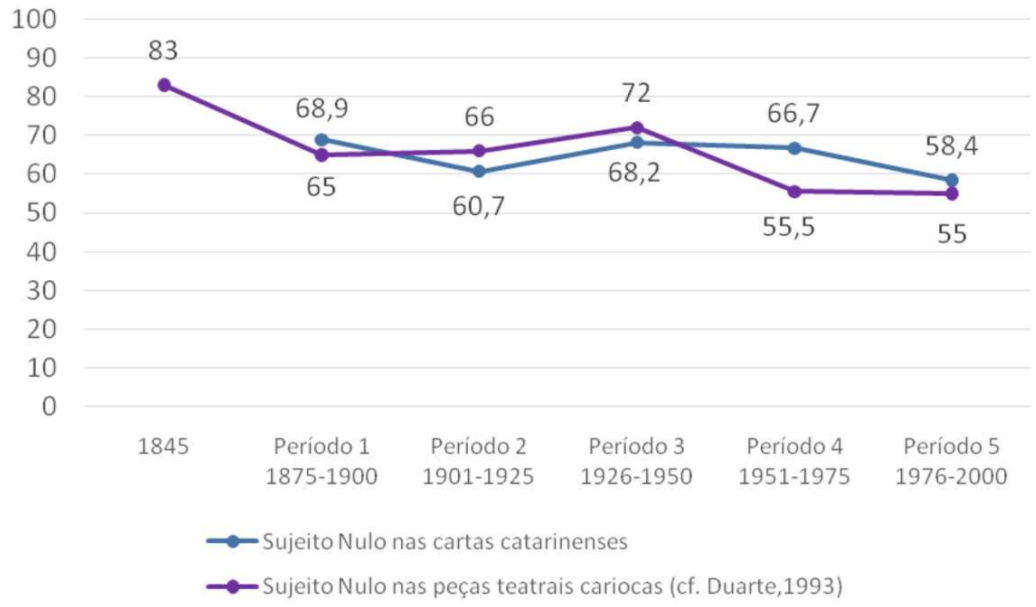
Na próxima seção, comparamos nossos resultados com os de trabalhos que analisaram a realização do sujeito e do objeto direto no decorrer do tempo em peças teatrais do Rio de Janeiro.

2.3 Comparação dos resultados entre as cartas catarinenses e as peças teatrais cariocas

Como vimos até aqui, os resultados das cartas pessoais catarinenses apontam, no geral, para uma estabilidade da mudança do sujeito: somente na última sincronia é que vemos uma diferença no percentual de sujeito nulo, que é 10% mais baixo do que o índice do primeiro recorte de tempo. Sobre o objeto, atestamos mudança em tempo real quando observamos os percentuais de objeto nulo aumentando com o passar dos anos.

Nossos resultados são bem próximos daqueles encontrados por Duarte (1993) nas peças teatrais do Rio de Janeiro no que concerne ao sujeito nulo de terceira pessoa. O Gráfico 1, a seguir, apresenta essa comparação[8].

Gráfico 1: Percentuais de sujeito nulo anafórico nas cartas pessoais catarinenses e nas peças teatrais cariocas (*Cf.* Duarte, 1993) nos séculos XIX e XX



Fonte: Adaptado de Duarte, Mourão e Santos (2012, p. 22).

[8] Considerando que Duarte – e também Marques de Sousa, no próximo gráfico em que comparamos os resultados de objeto – analisa sete diferentes períodos de tempo, voltamos a apresentar nossos percentuais separados em cinco sincronias.

Os percentuais de sujeito nulo de terceira pessoa encontrados pela autora apresentam uma queda acentuada de 1845 (83%) a 1882 (65%), mantendo estabilidade nas sincronias seguintes (66% em 1918 e 72% em 1937). A partir da segunda metade do século XX, os índices decrescem para 55,5% (anos de 1955 e 1975 da autora)[9], permanecendo em 55% em 1992. Quanto às nossas amostras de cartas catarinenses, os resultados seguem praticamente a mesma curva dos percentuais do Rio de Janeiro nos nossos três primeiros recortes, com 68,9%, 60,7% e 68,2% de sujeito nulo, respectivamente. Nas duas últimas sincronias, Santa Catarina é um pouco mais resistente, principalmente no período 4, com 66,7% de apagamento. Já no período 5, o percentual chega bem próximo daquele encontrado para o Rio, em 58,4%. Os resultados dos dois trabalhos representam esse período de transição pelo qual o PB está passando, de um sistema consistente de sujeito nulo para um sistema mais variável, com um aumento lento e gradual de sujeitos plenos (*Cf.* Kato; Duarte, 2014, 2017).

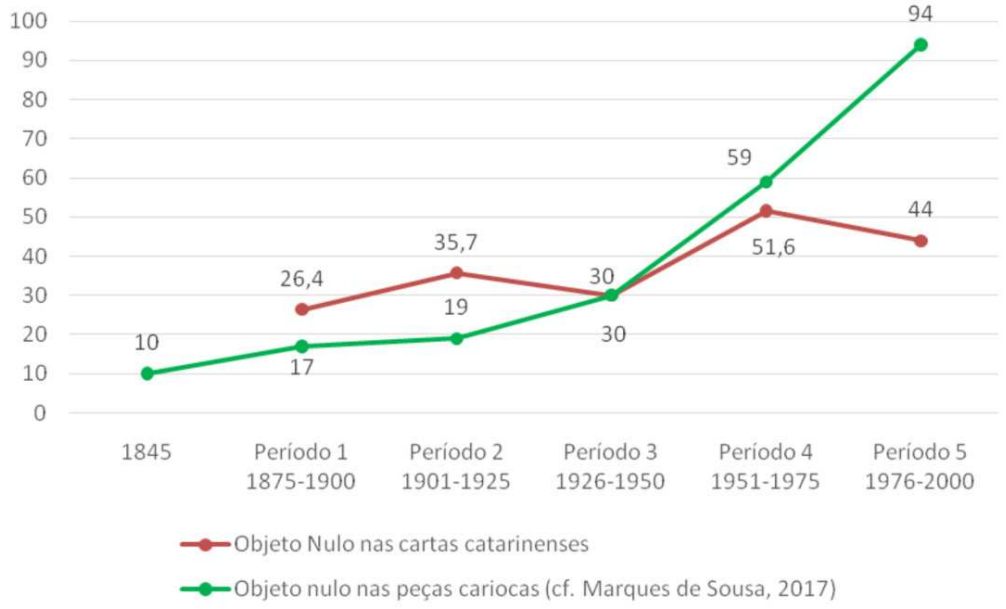
Nessa comparação entre cartas pessoais catarinenses e peças teatrais cariocas, os percentuais de sujeito nulo de terceira pessoa parecem seguir o mesmo movimento de mudança, diferentemente do que acontece com os sujeitos de primeira e de segunda pessoas, que decrescem no Rio de Janeiro (*Cf.* Duarte, 1993) e se mantêm conservadores em Santa Catarina – com índices altos até mesmo nas últimas sincronias analisadas (*Cf.* Coelho; Vieira-Pinto, 2021). O caráter anafórico da terceira pessoa parece ser o elemento que mais aproxima a escrita das duas regiões (atestando o que Duarte costuma chamar de "mudança sintática democrática"). Ademais, as forças linguísticas que atuam na terceira pessoa das cartas catarinenses também são de mesma natureza das que foram atestadas nas peças cariocas, em especial o padrão sentencial e a especificidade.

Da mesma forma, para o fenômeno do objeto direto, encontramos semelhanças com os resultados atestados no trabalho de Marques de Sousa (2017), que também analisa peças de teatro cariocas[10]. Vejamos a comparação entre as cartas pessoais de Santa Catarina e as peças de teatro do Rio de Janeiro apresentada no Gráfico 2.

[9] Na apresentação do Gráfico 1, amalgamamos as sincronias de 1955 e 1975 de Duarte (1993) a fim de comparação com nossos recortes de tempo. Em seu trabalho, os percentuais de sujeito nulo foram de 59% em 1955 e 52% em 1975.

[10] Exceto no período que corresponde ao final do século XIX, Duarte (1993) e Marques de Sousa (2017) analisam o mesmo corpus de peças teatrais (uma peça por período). Apenas na década de 1880 é que Duarte utiliza a peça "Como se fazia um deputado" (1882), de França Junior, e Marques de Sousa coleta dados na peça "As Doutoras" (1889), do mesmo autor.

Gráfico 2: Percentuais de objeto nulo nas cartas pessoais catarinenses e nas peças teatrais cariocas (*Cf.* Marques de Sousa, 2017), nos séculos XIX e XX



Fonte: Adaptado de Marques de Sousa (2017, p. 93).

Os resultados do autor mostram a evolução do objeto nulo, que ganha força a cada sincronia, chegando à quase completação da mudança no final do século XX. Os percentuais de apagamento se mantêm praticamente estáveis nos três primeiros recortes, com 10% em 1845, 17% em 1889 e 19% em 1918; mas, a partir da década de 1930, a mudança passa a ser bastante significativa, com 30% em 1937, 59% em 1955/1975[11] e 94% em 1992. Tais resultados atestam com clareza a mudança de um sistema preferencialmente de pronome clítico – que, nos primeiros períodos analisados, era baseado no português lusitano – para uma gramática legitimamente do PB, em que o objeto nulo é a variante mais utilizada, invertendo o quadro inicial.

As cartas pessoais catarinenses seguem, no geral, a mesma curva de mudança registrada nas peças, com distâncias mais marcadas entre os índices apenas nos períodos 2 e 5. No período 1, apesar de um leve favorecimento do nulo nas cartas, o percentual foi de 26,4%, próximo ao que Marques de Sousa (2017) atestou para o mesmo período. No período 2, no entanto, a diferença encontrada entre as cartas (35,7% de objeto nulo) e as peças (19% apenas) é mais acentuada, o que não seria esperado, já que esse período é

[11] Amalgamamos as sincronias de 1955 e 1975 de Marques de Sousa na apresentação do Gráfico 2 a fim de comparação com nossos recortes de tempo. O objeto nulo teve 66% em 1955 e 52% em 1975.

relacionado ao Modernismo, e a ocorrência da variante inovadora costuma ser frequente (*Cf.* Cyrino; Torres Morais, 2018) – o que deveria ter ocorrido nas peças teatrais. Talvez esse contraste possa ser explicado por conta do enredo da peça estudada por Marques de Sousa, "O simpático Jeremias", de Gastão Tojeiro: o personagem principal é um aprendiz de filósofo, e é característica de sua fala uma linguagem mais erudita, que é destacada em comparação com a fala de outros personagens. Pode ser que, por conta disso, os usos de pronome clítico tenham sido mais frequentes do que o esperado.

Os percentuais voltam a ser semelhantes nos períodos 3 e 4, em que o objeto nulo teve 30% e 51,6% nas cartas catarinenses. Quanto à última sincronia analisada, porém, não podemos afirmar a mesma curva, pois as missivas apresentam 44% de objeto nulo e as peças cariocas, 94%. Marques de Sousa (2017) compara seus resultados de 1992 com outros trabalhos que analisaram a fala do PB e verifica que os percentuais das variantes seguem a mesma tendência. Ao que tudo indica, Miguel Falabella, autor da peça em questão, parece ter tomado a língua falada do PB como base para seu texto escrito, uma excelente característica, que confirma o fato de que as peças retratam a sociedade do momento, inclusive em seus usos linguísticos. De qualquer forma, as cartas catarinenses desse mesmo período, apesar de não revelarem os mesmos percentuais atestados na fala (até mesmo de Santa Catarina), transparecem a implementação do objeto nulo na escrita, conforme nossas expectativas.

Sobre as variáveis selecionadas, os dois trabalhos constataram o período de tempo como importante fator – atestando mudança em tempo real do objeto nulo nas peças cariocas e nas cartas catarinenses –, além de terem sido também destacadas a animacidade e a especificidade como condicionadoras da categoria vazia.

3. Considerações finais

Nesta pesquisa, fizemos uma análise comparativa entre a realização do sujeito e do objeto direto anafóricos, verificando a trajetória desses dois fenômenos sintáticos em cartas pessoais catarinenses escritas ao longo dos séculos XIX e XX, separadas em cinco períodos de tempo. Sobre a análise do sujeito de uma sincronia para a outra, os percentuais são praticamente estáveis, havendo diferença apenas no último período analisado. Mesmo com a queda de 10% do primeiro para o último recorte de tempo, a preferência pelo sujeito nulo na escrita dos catarinenses ainda é grande. No que tange à análise do objeto direto, atestamos o objeto nulo ganhando força em Santa Catarina à medida que o tempo vai passando, indo a favor de nossa expectativa inicial. Sobre as forças que condicionam

as categorias vazias, não registramos as mesmas variáveis externas atuando nas mudanças, mas somente algumas forças internas. As variáveis linguísticas selecionadas em comum, tanto para o sujeito quanto para o objeto, são: o "padrão sentencial" e a "especificidade do referente". Por fim, uma comparação dos resultados das cartas catarinenses com aqueles obtidos em amostras de peças teatrais cariocas (Duarte, 1993; Marques de Sousa, 2017) aponta para o fato de que os percentuais de sujeito de terceira pessoa e de objeto direto anafórico parecem seguir o mesmo movimento de mudança nas duas regiões.

Referências

BARBOSA, P.; DUARTE, M. E. L.; KATO, M. A. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, v. 4, p. 11-52, 2005.

BERLINCK, R. de A. *A ordem V SN no português do Brasil*: sincronia e diacronia. 1988. 265 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1988.

BERLINCK, R. de A. A construção V SN no português do Brasil: uma visão diacrônica do fenômeno da ordem. *In*: TARALLO, F. (org.). *Fotografias sociolingüísticas.* São Paulo: Pontes, 1989. p. 95-112.

BIBERAUER; T.; HOLMBERG, A.; ROBERTS, I.; SHEEHAN, M. *Parametric variation*: null subjects in minimalist theory. Cambridge: Cambridge University Press, 2010.

BIBERAUER; T.; ROBERTS, I. Parameter setting. *In*: LEDGEWAY, A.; ROBERTS, I. (orgs.). *The Cambridge Handbook of Historical Syntax*. Cambridge: Cambridge University Press, 2017. p. 134-162.

CHOMSKY, N. *Lectures on Government and Binding*. Holland: Foris Publications, 1981.

CHOMSKY, N. *O conhecimento da língua* – sua natureza, origem e uso. Tradução Anabela Gonçalves; Ana Teresa Alves. Lisboa: Caminho, 1994 [1986].

COELHO, I. L.; NUNES DE SOUZA, C. M. Uma proposta metodológica para o tratamento da variação estilística em textos escritos. *In*: GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (orgs.). *Variação estilística*: reflexões teórico-metodológicas e propostas de análise. Florianópolis: Insular, 2014. p. 163-199.

COELHO. I. L.; VIEIRA-PINTO, C. A.; ZIBETTI, E. M. O.; SILVA, G. M. e. Ordem SV, sujeito expresso e objeto nulo: a trajetória da mudança no português de Santa Catarina. *Actas do XVIII Congresso Internacional ALFAL* – Projetos, Bogotá: Universidade de Bogotá, p. 1-16, 2017.

COELHO, I. L; VIEIRA-PINTO, C. A. O sujeito nulo em cartas pessoais catarinenses no curso dos séculos XIX e XX (1885-1998). *In:* MARINS, J. E.; ORSINI, M. T.; CAVALCANTE, S. R. de O. (orgs.). *Contribuições à descrição e ao ensino do português brasileiro:* da fonética ao discurso, com parada obrigatória na sintaxe. São Paulo: Pimenta Cultural, 2021. p. 281-309.

CONDE SILVESTRE, J. C. *Sociolingüística Histórica*. Madrid: Editorial Gredos, 2007.

COSTA, S. *O (não) preenchimento do objeto anafórico na língua portuguesa*: análise diacrônica do PB e do PE dos séculos XIX e XX. 2011. 268 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2011.

CYRINO, S. M. L. *O objeto nulo no Português do Brasil*: um estudo sintático-diacrônico. Londrina: Ed. da UEL, 1997.

CYRINO, S. M. L.; TORRES MORAIS, M. A. (org.). *História do português brasileiro* - mudança sintática do português brasileiro: perspectiva gerativista. São Paulo: Contexto, 2018.

DUARTE, M. E. L. Clítico acusativo, pronome lexical e categoria vazia no Português do Brasil. *In*: TARALLO, F. (org.). *Fotografias sociolingüísticas.* São Paulo: Pontes, 1989. p. 19-34.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. *In*: ROBERTS, I.; KATO, M. A. (orgs.). *Português brasileiro*: uma viagem diacrônica. Campinas: Ed. da UNICAMP, 1993. p. 107-128.

DUARTE, M. E. L. *A perda do princípio Evite Pronome no português brasileiro*. 1995. 151 f. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1995.

DUARTE, M. E. L. O sujeito nulo no português brasileiro. *In*: CYRINO, S.; TORRES MORAIS, M. A. (orgs.). *História do português brasileiro* - mudança sintática do português brasileiro: perspectiva gerativista. São Paulo: Contexto, 2018. p. 26-71.

DUARTE, M. E. L.; MOURÃO, G. C.; SANTOS, H. M. Os sujeitos de 3ª pessoa: REvisitando Duarte 1993. *In*: DUARTE, M. E. L. (org.). *O sujeito em peças de teatro (1833-1992)*: estudos diacrônicos. São Paulo: Parábola, 2012. p. 21-44.

FARACO, C. A. *Lingüística Histórica*: uma introdução ao estudo da história das línguas. São Paulo: Parábola Editorial, 2005.

GALVES, C.; CYRINO, S.; LOPES, R.; SANDALO, F.; AVELAR, J. (org.). *Parameter theory and linguistic change*. Oxford: Oxford University Press, 2012.

KATO, M. A.; DUARTE, M. E. L. Restrições na distribuição de sujeitos nulos no português brasileiro. *Veredas: Sintaxe das línguas brasileiras*, v. 18/1, p. 1-22, 2014.

KATO, M. A.; DUARTE, M. E. L. O sujeito no português brasileiro e sua tipologia. *In*: PILATI, E.; SALLES, H. L.; NAVES, R. (orgs.). *Novos olhares para a gramática do português brasileiro*. Campinas: Pontes Editores, 2017. p. 13-42.

LABOV, W. *Padrões sociolingüísticos*. Tradução Marcos Bagno; Maria Marta Pereira Scherre; Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].

LABOV, W. *Principles of linguistic change* – Internal factors. Cambridge: B. Blackwell, 1994.

MARAFONI, R. L. *A realização do objeto direto anafórico*: um estudo em tempo real de curta duração. 2004. 112 f. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.

MARQUES DE SOUSA, A. A. *As realizações do acusativo anafórico no português europeu e brasileiro*: um estudo diacrônico. 2017. 205 f. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017.

MATTOS E SILVA, R. V. *Ensaios para uma sócio-história do português brasileiro*. São Paulo: Parábola Editorial, 2004.

MATTOS E SILVA, R. V. *Caminhos da lingüística histórica*: ouvir o inaudível. São Paulo: Parábola Editorial, 2008.

OTHERO, G. A.; SPINELLI, A. C. Um tratamento unificado da omissão e da expressão de sujeitos e objetos diretos pronominais de 3ª pessoa em português brasileiro. *Cadernos de Estudos Linguísticos*, v. 61, p. 1-30, 2019.

PAIXÃO DE SOUSA, M. C. Linguística histórica. *In*: PFEIFFER, C. C. *Introdução às ciências da linguagem* – linguagem, história e conhecimento. Campinas: Pontes Editores, 2006. p. 11-48.

ROMAINE, S. *Socio-Historical Linguistics* - Its status and methodology. Cambridge: Cambridge University Press, 1982.

SOLEDADE, C. L. V. *A realização do objeto direto anafórico em peças de autores brasileiros dos séculos XIX e XX*: dados empíricos para a observação da mudança no português brasileiro. 2011. 113 f. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

TARALLO, F. *Relativization strategies in brazilian portuguese*. 1983. 273 f. Tese (Doutorado em Linguística) – University of Pennsylvania, Filadélfia,1983.

TARALLO, F. *Tempos lingüísticos*: itinerário histórico da língua portuguesa. São Paulo: Ática, 1990.

TARALLO, F. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. *In*: ROBERTS, I.; KATO, M. A. (orgs.). *Português brasileiro*: uma viagem diacrônica. Campinas: Ed. da UNICAMP, 1993a. p. 35-68.

TARALLO, F. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. *In*: ROBERTS, I.; KATO, M. A. (orgs.). *Português brasileiro*: uma viagem diacrônica. Campinas: Ed. da UNICAMP, 1993b. p. 69-105.

TARALLO, F.; KATO, M. A. Harmonia trans-sistêmica: variação inter e intralinguística. *Diadorim*, v. 2, p. 13-42, 2007 [1989].

VIEIRA-PINTO, C. A. *Variação do objeto anafórico acusativo na fala de Florianópolis.* 2015. 164 f. Dissertação (Mestrado em Linguística) - Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2015.

VIEIRA-PINTO, C. A. *Trajetória de mudança do sujeito e do objeto direto anafóricos:* análise de cartas pessoais brasileiras e portuguesas dos séculos XIX e XX**.** 2020. 391 f. Tese (Doutorado em Linguística) - Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 2020.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança lingüística*. Tradução Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].



http://doi.org/10.5007/1984-8420.2025.e100437

A ORDEM DO VERBO NA HISTÓRIA DO PB: QUESTÕES PARA PERIODIZAÇÃO[1]

VERB ORDER IN THE HISTORY OF BP: ISSUES FOR PERIODISATION

Paulo Ângelo Araújo-Adriano | Lattes | pauloangelo@usp.br
Universidade de São Paulo

Williane Corôa | Lattes | williscoroa@gmail.com
Universidade do Estado da Bahia

Resumo: O objetivo principal deste artigo é apresentar novos elementos para a periodização do português brasileiro (PB). Situamo-nos no campo disciplinar da Sintaxe diacrônica, na perspectiva gerativista, e analisamos a ordem do verbo em textos – anotados morfológica e sintaticamente – que fazem parte do Corpus Tycho Brahe-Brasil, com fins de oferecer novos elementos sobre a periodização do português no Brasil. Escritos por indivíduos nascidos no Brasil, os textos dão conta de um grande período: desde o século 17 até o século 21. Apresentamos evidências de que o português escrito no Brasil apresenta três grandes períodos: o Período I, cuja gramática reflete propriedades do português clássico, considerada "a língua das caravelas", por Galves (2007); Período II, cujos verbos lexicais e auxiliarem perdem movimento, deixando de subir até a periferia esquerda da sentença, mas subindo até uma posição de tempo; e o Período III, em que há adicional perda de movimento em relação aos verbos lexicais, mas não aos auxiliares. Esses resultados parecem validar a hipótese defendida por Ribeiro (1998) e Galves (2007) de que as origens do PB são anteriores ao século XIX e auxiliam a repensar a periodização da língua portuguesa no Brasil.

Palavras-chave: Periodização do português brasileiro. Ordem do verbo. Movimento do verbo.

[1] Este artigo foi apresentado na 30ª Conferência do GELNE (Grupo de Estudos Linguísticos e Literários do Nordeste), realizado em Feira de Santana, na Universidade Estadual de Feira de Santana, em dezembro de 2024. Agradecemos ao público pelas perguntas e sugestões. Além disso, estendemos nossa gratidão aos revisores anônimos por suas contribuições, perguntas e comentários sobre este texto.

Abstract: The main goal of this paper is to present new elements for the periodisation of Brazilian Portuguese (BP). Under the field of Diachronic Syntax, from a generative perspective, we analyse verb order in morphologically annotated texts, which are part of the Tycho Brahe-Brasil Corpus, in order to offer new insights into the periodisation of Portuguese in Brazil. The texts cover a wide period, from the 17th century to the 21st century, and were written by individuals born in Brazil. We present evidence that written Portuguese in Brazil exhibits three major periods: Period I, whose grammar reflects properties of Classical Portuguese, referred to as "the language of the caravels" by Galves (2007); Period II, during which lexical verbs and auxiliaries lose movement, ceasing to move to the left periphery of the sentence but rising to a Tense position; and Period III, in which there is additional loss of movement in relation to lexical verbs but not to auxiliaries. These results appear to validate the hypothesis advocated by Ribeiro (1998) and Galves (2007) that the origins of BP precede the 19th century and help to reconsider the periodisation of the Portuguese language in Brazil.

Key-words: Brazilian Portuguese periodization. Verb order. Verb movement.

Introdução

Quando comparado ao português europeu (PE), o português brasileiro (PB) apresenta diferenças consideráveis que fazem alguns autores tratarem essas línguas como diferentes (Kato, 2012). Mas em que momento da história o PB teria começado a se diferenciar do PE? Se voltarmos nosso olhar sobre a periodização da língua portuguesa no Brasil, várias propostas têm se sucedido.

Um dos primeiros pesquisadores a avaliar a complexidade do caso linguístico brasileiro foi Serafim da Silva Neto (1986[1950]), que traçou, a partir da história social da colônia, a história da língua portuguesa na América. O autor dividiu a história da língua portuguesa no Brasil em três fases: a primeira do início da colonização (1532) até a expulsão dos holandeses (1654); a segunda, inicia-se em 1654 e vai até 1808 com a chegada do Príncipe Regente e da Corte portuguesa ao Brasil; e a terceira fase, a partir de 1808, é marcada pela chegada da Corte portuguesa ao Rio de Janeiro que transformou profundamente a vida na colônia.

Outra proposta de periodização é de Marlos de Barros Pessoa (2003). O grande diferencial dessa proposta em relação à anterior é a tentativa de unir marcos sócio-históricos com estágios de desenvolvimento linguístico. Segundo Pessoa (2003), entre 1534 e 1750

imperam o multilinguismo e a formação de variedades linguísticas rurais. A essa fase, segue-se um estágio de koineização de diferentes variedades ou pré-koineização da língua comum (1750-1808), um período de formação de variedades urbanas (1808-1850) e um subperíodo de estabilização das variedades urbanas e da língua comum (1850-1922), que, englobados, correspondem à segunda fase de gestação do PB. Posteriormente, na terceira fase de desenvolvimento do PB, a partir de 1922, há um período de elaboração da língua literária.

Lobo (2003) também propõe uma periodização para a história linguística no Brasil. Sua proposta fundamenta-se na história demográfico-linguística brasileira, no crescimento populacional associado ao processo de urbanização do país e no processo de escolarização associado ao processo de estandardização linguística. Os critérios escolhidos são, portanto, de natureza sociolinguística. Para Lobo (2003), a periodização para a história linguística no Brasil divide-se em duas grandes fases: a primeira, na qual prevalece o multilinguismo generalizado, cuja característica é a não urbanização, a não escolarização e a não estandardização linguística; e a segunda fase, na qual prevalece o multilinguismo localizado, cuja característica é a escolarização, a urbanização e a estandardização linguística.

Pautando-se em critérios linguísticos, sem, no entanto, separar a evolução linguística e a sócio-história dos falantes, Volker Noll (2008) critica as propostas anteriormente apresentadas, que se limitam quase exclusivamente a fatos históricos, e define sua proposta da seguinte maneira: de 1500 a 1550, fase inicial, em que a língua portuguesa é trasladada para o Brasil; de 1550 a 1700, primeira fase formativa, em que é possível ver a formação das primeiras características da língua portuguesa no Brasil; de 1700 a 1800, fase diferenciadora, na qual a formação do PB se consolida e a diferenciação das variedades europeia e brasileira começa a aparecer; de 1800 a 1950, fase de desenvolvimento da escrita e do ensino assinalada pela introdução da imprensa, pela implantação do ensino público oficial e a criação do ensino superior e pela diferenciação progressiva da norma europeia; e, de 1950 em diante, fase de nivelação, marcada pela evolução dos meios de comunicação, pela introdução da televisão e a urbanização progressiva.

Por fim, Dante Lucchesi (2017) propõe uma periodização para o PB partindo da história sociolinguística do Brasil. Fundamenta-se na história linguística e social do Brasil e considera um conjunto mais amplo de línguas presentes na formação da sociedade brasileira. A grande inovação desta proposta é o recuo temporal para antes da chegada dos portugueses ao Brasil.

Desse modo, para Lucchesi (2017), a primeira fase da história sociolinguística do Brasil, denominada de Tupinização da Costa, inicia-se após o ano mil, com a expansão tupi pelo litoral, e encerra-se no ano de 1532, com o efetivo início da colonização do Brasil pelos portugueses. A segunda fase, o Multilinguismo generalizado, estende-se de 1532 a 1695 e é caracterizada pela baixa densidade demográfica e pelo plurilinguismo. A terceira fase, a Homogeneização Linguística, inicia-se em 1695, com a descoberta do ouro em Minas Gerais, e estende-se até a Revolução de 1930, subdivindo-se em três momentos: (1) de 1695-1808, marcada pelo primeiro surto de urbanização do Brasil com o ciclo do ouro e avanço da pecuária; (2) de 1808-1850, em que se intensifica o processo de urbanização com a chegada da Corte portuguesa ao Brasil e a Independência política; (3) de 1850-1930, cujas marcas são o fim do tráfico negreiro, a imigração maciça de europeus e asiáticos, e o aprofundamento da normatização purista e lusitanizante. A quarta fase, chamada de Nivelamento linguístico, de 1930 aos dias atuais, assinala-se pela crescente industrialização e urbanização do país, pela difusão da norma urbana culta, apagando as marcas do contato na norma popular. Essas diferentes propostas para a periodização do PB estão organizadas abaixo.

Propostas de periodização para o português brasileiro				
Serafim da Silva Neto (1986)				
1ª fase 1532-1654	2ª fase 1654-1808	3ª fase 1808-dias atuais		
Marlos de Barros Pessoa (2003)				
1ª fase 1534-1750	2ª fase 1750-1922	3ª fase 1922-dias atuais		
Tânia Lobo (2003)				
1ª fase 1500-1850	2ª fase 1850-dias atuais			
Volker Noll (2008)				
1ª fase 1500-1550	2ª fase 1550-1700	3ª fase 1700-1800	4ª fase 1800-1950	5a fase 1950-dias atuais
Dante Lucchesi (2017)				
1ª fase 1000-1532	2ª fase 1532-1695	3ª fase 1695-1930	4ª fase 1930-dias atuais	

Apesar de as diversas propostas elencadas considerarem aspectos da história externa e/ou interna da língua e tentarem explicar as diferenças entre o PB e PE, nenhuma considera aspectos sintáticos, marcadamente onde o PB e o PE se diferem mais fortemente. Outro aspecto a se notar é que nenhuma proposta aponta para o momento em que o PB emergiu.

Por outro lado, estudos como os de Tarallo (1993), Ribeiro (1998), Galves (2007) e Corôa (2022), apesar de não proporem uma periodização, buscam responder em que momento o PB emerge. Esses trabalhos examinaram aspectos sintáticos definidores da gramática do PB, como a reorganização do sistema pronominal – que leva à perda dos clíticos e à reorganização do sistema de possessivos –, a mudança da ordem VS em sentenças declarativas e interrogativas, o enfraquecimento do sistema de concordância e a perda do sujeito nulo.

Tarallo (1993) observou três processos de mudança no PB em comparação ao PE: (i) a perda da referência pronominal, (ii) alterações nas estratégias de pronominalização e (iii) a reorganização dos padrões sentenciais básicos. Com relação a (i), no PB, houve perda da referência pronominal, por isso o sistema dos pronomes plenos e dos pronomes clíticos se reorganizou. Como consequência de (i), o PB muda também as estratégias de relativização, com as relativas *piedpiping* – superficialmente idêntica às relativas encontradas na norma padrão –, as relativas com pronome lembrete – cuja posição da lacuna é preenchida por uma forma pronominal correferente – e a relativa cortadora – quando tanto a preposição governante quanto o sintagma relativizado estão ausentes. O terceiro processo de mudança (iii) também está ligado ao primeiro e ao segundo. Segundo Tarallo (1993), o PB sofreu uma grande reversão em suas estratégias de pronominalização, resultando, entre outras coisas, no enrijecimento do padrão canônico de ordem de palavras em direção a SVO. Todas as mudanças sinalizadas estão situadas, segundo Tarallo (1993), no final do século 19.

Já Ribeiro (1998), Galves (2007) e Corôa (2022) propõem uma retroação temporal, considerando aspectos da gramática dos primeiros portugueses que desembarcaram no Brasil. Ribeiro (1998) defende que possivelmente houve uma mudança no PB na virada do século 18 para o 19, mas que muitos dos fatos analisados como resultantes de uma mudança do PB no século 19 se originaram no século 16. Por isso, propõe que o português do Brasil seja constituído de, pelo menos, duas gramáticas diferentes, a dos séculos 16-18 e a dos séculos 19-20. Os fenômenos analisados por Ribeiro (1998) para embasar

sua proposta estão ligados à colocação pronominal. Galves (2007) retoma o questionamento de Ribeiro (1998) e, a partir de um estudo sistemático da colocação de clíticos na história do português, defende que o PB teria evoluído a partir da gramática do português médio – fase gramatical intermediária entre o português arcaico e o português moderno.

Adicionalmente, Corôa (2022) analisou o efeito V2 e suas propriedades correlatas: colocação de clíticos pronominais, sujeito e objeto nulos em dados dos séculos 17 e 18. Os resultados encontrados indicaram uma gramática bastante próxima da gramática do português clássico entre os brasileiros nascidos no século XVII, com mudanças significativas, como a perda do efeito V2, a generalização da próclise e a ocorrência de objetos nulos tal como acontece no PB: com leitura referencial em dados do século 18, o que sugere que a gramática do PB emerge no século 18.

Ainda, Cardoso, Andrade e Carneiro (2023), com base em dados de colocação de clíticos, apresentam a hipótese de que, entre o português clássico e a constituição do português brasileiro moderno em sua vertente prestigiada, houve uma variedade linguística adquirida pelos brancos nascidos no Brasil, chamada pelos autores de português colonial brasileiro que, embora semelhante com o português clássico, possui distanciamentos face a este de modo que não constituem marcas evidentes de uma gramática vernacular brasileira.

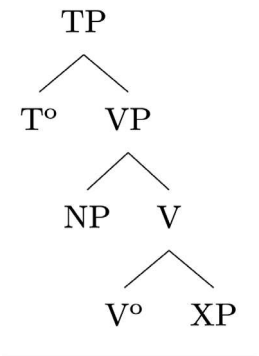
Seguindo esses trabalhos, este artigo busca responder a seguinte questão: em que momento a vertente brasileira do português emergiu? A partir de evidências intralinguísticas em nível sintático, propõe-se uma periodização para o PB. O fenômeno escolhido para análise é a ordem do verbo ao longo da história a partir de um conjunto de dados dos séculos 17 ao 21. Segundo Galves (2018), o estudo sobre a evolução da posição do verbo ao longo do tempo tanto no PB quanto no PE é importante, pois pode nos ajudar a entender a dinâmica da mudança nesses dois sistemas, ao lado da perda do efeito V2. Assim, o ordenamento verbal em relação ao objeto, propriedade de uma gramática V2, também é considerado, apesar de o foco principal do trabalho ter sido observar o posicionamento do verbo em relação a advérbios fixos.

A hipótese defendida alinha-se às hipóteses de Ribeiro (1998), Galves (2007) e Corôa (2022), defendendo a virada do século 17 para 18 como momento em que a gramática brasileira emerge.

1 Ordem de palavras: o verbo

Dentro do quadro da Gramática Gerativa, universalmente, nas línguas naturais, o verbo lexical concatena-se com seus argumentos para formar o sintagma verbal (VP, *Verbal Phrase*), como representado em (1). A despeito dessa configuração universal, a ordem do verbo em relação a certos elementos fixos na sentença é variável. Considerando os exemplos em (2), é possível observar que o verbo nas línguas românicas e germânicas se comportam de maneira diferente. O verbo *sleep* ('dormir') aparece à direita do advérbio *often* ('frequentemente') em inglês; em francês, entretanto, se o verbo segue o advérbio *souvent* ('frequentemente'), a sentença é agramatical, mas quando o precede, a sentença é gramatical.

(1)



(2) a. Inglês

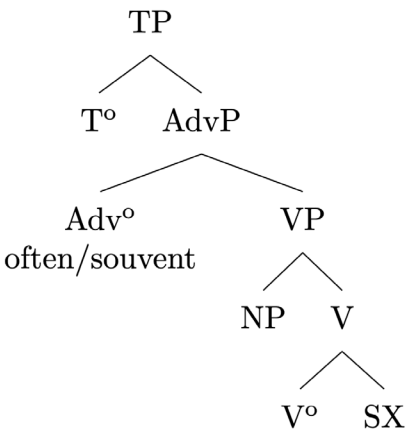
John	(*sleeps)	often	sleeps	early.
J.	*dorme*	*frequentemente*	*dorme*	*cedo*

b. Francês

Jean	dort	souvent	(*dort)	tôt.
J.	*dorme*	*frequentemente*	*dorme*	*cedo*

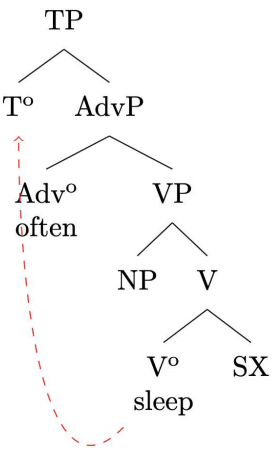
A diferença atestada acima, em que o verbo precede o advérbio em francês, mas o segue em inglês, apesar de o verbo nas as línguas ter a mesma estrutura, foi implementada na literatura gerativa como sendo resultado da presença/ausência de uma regra de movimento verbal. Alguns autores sugeriram que elementos como os advérbios e a negação eram fixos na sentença (cf. Edmonds, 1978; Pollock, 1989; Cinque, 1999; Belletti, 1990): como argumentado por Pollock (1989), advérbios do tipo *often*/*souvent* seriam adjuntos de VP, isto é, concatenar-se-iam universalmente acima de VP, como ilustrado em (3).

(3)

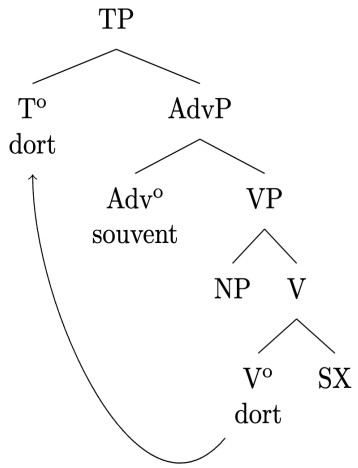


Portanto, o fato de o verbo finito lexical inglês obrigatoriamente seguir o advérbio *often*, diferentemente da contraparte francesa *souvent* em que o verbo o precede, revelaria que nas línguas germânicas o verbo não deixa o VP *via* movimento sintático. Isso seria diferente nas línguas românicas, cujo movimento sintático do verbo seria aplicado. Assim, a ordem ADV_V (*often sleeps*) no inglês seria explicada pela ausência de movimento à flexão (sinalizada pela linha pontilhada em (4a)), enquanto a ordem francesa V_ADV, pela presença de movimento, como ilustrado em (4b), a seguir.

(4) a. Inglês



b. Francês



Adotando essa proposta de que certos elementos na estrutura da sentença são fixos, Cinque (1999) argumenta haver cerca de 40 advérbios fixos e ordenados acima do sintagma verbal que semanticamente lexicalizam o especificador de diferentes projeções

funcionais. Dentro dessa literatura especializada, a Cartografia Sintática, justamente por mapear advérbios na estrutura, a hipótese é a de que o diferente ordenamento do verbo em relação a advérbios entre as línguas seria resultado de uma variação translinguística envolvendo o local para onde o verbo se move.

(5) A hierarquia universal das projeções funcionais de TP – versão do português brasileiro (Adaptado de Cinque, 1999, p. 106; Ledgeway & Lombardi, 2005, p. 81; Cinque, 2006)
ALTOS: [frankly $Mood_{SpeechAct}$ > [luckily $Mood_{Evaluative}$ > [allegedly $Mood_{Evidential}$ > [probably $Mod_{Epistemic}$ > [once T_{Past} > [then T_{Future} > [perhaps $Mood_{Irrealis}$ > [necessarily ModNecessity > [possibly $Mod_{Possibility}$ > [usually $Asp_{Habitual}$ > [finally $Asp_{Delayed}$ > [tendentially $Asp_{Predispositional}$ > [again $Asp_{Repetitive(I)}$ > [often $Asp_{Frequentative(I)}$ > [willingly $Mod_{Volition}$ > [quickly $Asp_{Celerative(I)}$ >
BAIXOS: [already $T_{Anterior}$ > [no longer $Asp_{Terminative}$ > [still $Asp_{Continuative}$ > [always $Asp_{Perfect}$ > [just $Asp_{Retrospective}$ > [soon $Asp_{Proximative}$ > [briefly $Asp_{Durative}$ > [(?) $Asp_{Generic/Progressive}$ > [almost $Asp_{Prospective}$ > [suddenly $Asp_{Inceptive}$ > [obligatorily $Mod_{Obligation}$ > [in vain $Asp_{Frustrative}$ > [(?) $Asp_{Conative}$ > [completely $Asp_{SgCompletive(I)}$ > [everything $Asp_{PlCompletive}$ > [well Voice > [early $Asp_{Celerative(II)}$ > [? $Asp_{Inceptive(II)}$ > [again $Asp_{Repetitive(II)}$ > [often $Asp_{Frequentative(II)}$ > … *v*P

Assim, de modo a precisar o local de pouso do verbo e compará-lo entre as línguas, a hierarquia dos advérbios foi metodologicamente dividida em duas áreas (cf. Ledgeway & Lombardi, 2005, p. 81): advérbios em uma posição mais alta, lexicalizando projeções funcionais entre $Asp_{Celerative(I)}$ e $Mood_{SpeechAct}$, advérbios em uma posição mais baixa, mais próxima do local onde o verbo é concatenado, lexicalizando projeções entre $T_{Anterior}$ e $Asp_{Frequentative(II)}$.

Apesar de o empreendimento da Cartografia ter sugerido uma profusão de projeções funcionais para a arquitetura da gramática, tal empreendimento parece essencial para um melhor diagnóstico sobre movimento do verbo. Como discutido anteriormente, após Pollock (1989) e antes do surgimento do modelo cartográfico, era assumido que o verbo lexical finito no inglês não se movia de dentro de VP, considerando seu ordenamento em

relação a advérbios do tipo *often* (cf. (2)). Após o modelo da Cartografia (Cinque 1999), o padrão inglês foi melhor compreendido. Ao ser adotada uma hierarquia mais articulada, tal como a proposta em (5) e (6), podemos observar que o verbo em inglês, contrariamente ao que fora proposto, se move, mas esse movimento é para uma posição baixa (cf. Tescari Neto, 2012; Roberts, 2017), passando apenas por sobre a projeção funcional de Voice, onde se lexicaliza o advérbio *well* ('bem'). Dessa forma, a hierarquia de Cinque (1999), que adotamos neste trabalho, parece funcionar melhor como um diagnóstico mais preciso do movimento do verbo.

(6) John (*well) sings **well**.
J. bem canta bem

A fim de discutir questões para a periodização do português brasileiro considerando o movimento do verbo na sua diacronia, assumimos a hierarquia na sua versão portuguesa tal como sugerida por Tescari Neto (2019). A partir dela, podemos determinar a posição do verbo na história do PB, dependendo do seu ordenamento em relação aos advérbios fixos e ordenados.

(7) A hierarquia universal das projeções funcionais de TP – versão do português brasileiro (Tescari Neto 2019: 3567)
ALTOS: [francamente $Mood_{SpeechAct}$ > [surpreendentemente $Mood_{Mirative}$ > [felizmente $Mood_{Evaluative}$ > [evidentemente $Mood_{Evidential}$ > [provavelmente $Mod_{Epistemic}$ > [uma vez T_{Past} > [então T_{Future} > [talvez $Mood_{Irrealis}$ > [necessariamente $Mod_{Necessity}$ > [possivelmente $Mod_{Possibility}$ > [normalmente $Asp_{Habitual}$ > [finalmente $Asp_{Delayed}$ > [tendencialmente $Asp_{Predispositional}$ > [novamente $Asp_{Repetitive(I)}$ > [frequentemente $Asp_{Frequentative(I)}$ > [de/com gosto $Mod_{Volition}$ > [rapidamente $Asp_{Celerative(I)}$ >
BAIXOS: [já $T_{Anterior}$ > [não…mais $Asp_{Terminative}$ > [ainda $Asp_{Continuative}$ > [sempre $Asp_{Perfect}$ > [apenas $Asp_{Retrospective}$ > [(dentro) em breve $Asp_{Proximative}$ > [brevemente $Asp_{Durative}$ > [(?) $Asp_{Generic/Progressive}$ > [quase $Asp_{Prospective}$ > [repentinamente $Asp_{Inceptive}$ > [obrigatoriamente $Mod_{Obligation}$ > [à toa $Asp_{Frustrative}$ > [(?) $Asp_{Conative}$ > [completamente $Asp_{SgCompletive(I)}$ > [tudo $Asp_{PlCompletive}$ > [bem Voice > [cedo $Asp_{Celerative(II)}$ > [do nada $Asp_{Inceptive(II)}$ > [de novo $Asp_{Repetitive(II)}$ > [frequentemente $Asp_{Frequentative(II)}$ > …

Por hipótese, a hierarquia acima fornece um diagnóstico preciso sobre a posição para onde o verbo se move. Apesar de estar dividida em advérbios altos e baixos, parece que em PB apenas os advérbios baixos devem ser tomados como bons diagnósticos para subida do verbo. Tescari Neto (2019, p. 3573) nota que advérbios altos, como *provavelmente*, em Spec,$Mod_{Epistemic}$, não podem seguir o verbo, quando estão em posição final (cf. (8a)), apesar de o poderem na presença de um NP objeto, como em (8b). Adicionalmente, assumindo que elipse de VP é licenciada por movimento do verbo (Matos & Cyrino, 2001), Tescari Neto (2019) também mostra que um advérbio alto não é recuperável na elipse de VP, o que aponta para a hipótese de que (8b) não deve ser gerado por movimento do verbo por sobre *provavelmente*, corroborando a afirmação de que "advérbios altos não podem ser utilizados, em PB, como diagnósticos para o movimento do V" (Tescari Neto, 2019, p. 3574).

(8) a. *João mente **provavelmente**.
b. João comia **provavelmente** arroz.

(9) O José comia **provavelmente** arroz e a Maria também comia [-]
a. [-]: ~~comia~~ arroz
b. [-]: ***provavelmente** arroz

Para refinarmos o diagnóstico de movimento do verbo ao longo do tempo, é preciso discutir casos em que um determinado posicionamento não é evidência necessária de ausência/presença de movimento do verbo. Como se vê em (10), em PB, *ainda* pode tanto preceder quanto seguir o verbo finito. A precedência do verbo por *ainda* pode ser explicado por deslocamento à periferia esquerda[2] (como sugerido por Andriani, 2016, p. 240 para o ancora em um dialeto italiano), como se vê em (10a), não como uma ausência de movimento do verbo. Diferentemente, em relação a esse advérbio, devemos tomar como evidência de movimento verbal apenas o ordenamento em que o verbo precede *ainda*, como em (10b).

[2] Esse deslocamento à periferia esquerda é corroborado pela precedência do advérbio a um sujeito quantificador na posição baixa de Foc, como *ninguém*.
(i) No começo, ainda *ninguém* tinha tentado elaborar uma versão falada para essa questão dos novos pronomes (https://www.extraclasse.org.br/geral/2021/10/a-gramatica-da-inclusao/)
(ii) Ainda *ninguém* está passando mal por enquanto, mas fizeram isso na maldade, só pode ser e quero que alguém se pronuncie a respeito (https://www.reclameaqui.com.br/burger-king/que-bebida-e-essa_EpeVUiu6bxaZ4qCR/)

(10) a. O Luciano **ainda** está se recuperando da cirurgia complexa que fez recentemente (cf. também "O Luciano está **ainda** se recuperando da cirurgia complexa que fez recentemente"[3])
b. O relator está **ainda** fazendo novos ajustes no texto, mas promete não mexer em questões polêmicas do projeto, como a venda de bebidas alcoólicas nos estádios[4].

No caso do advérbio *sempre*, há certas complicações porque esse advérbio apresenta diferentes leituras (Tescari Neto, 2013; Araújo-Adriano, 2018), uma assertiva (11a), uma temporal/aspectual (11b-c) e outra de padrão de comportamento (11d). Cada leitura é ativada a depender da projeção funcional em que *sempre* é concatenado: em $Mood_{SpeechAct}$, $Asp_{Perfect}$ ou $Asp_{Frequentative(II)}$, respectivamente. Para o presente trabalho, controlamos o posicionamento do verbo em relação a *sempre* na leitura temporal/aspectual, semanticamente associado a $Asp_{Perfect}$.

(11) a. **Sempre** quero ver se tens coragem para isso! (Lopes, 1998, p. 7)
b. Dane-se, eu vou **sempre** estar lá. Ninguém me deu esse lugar. Eu conquistei.
c. **Sempre** dei o melhor de mim, **sempre** te tratei muito bem.
d. Você vem **sempre** aqui?

O advérbio *bem* só pode aparecer pós-verbalmente, como em (12a). A precedência de *bem* ao verbo dispara uma leitura específica de foco (12b). Nesse sentido, como colocam Tescari Neto e Forero Pataquiva (2020, p. 495, NR 4), a posição pré-verbal de *bem* envolveria, depois do movimento – obrigatório – do verbo por sobre *bem*, movimento desse advérbio para Spec,Foc, e alçamento do sujeito para Spec,Top.

(12) a. A Maria (*bem) fala **bem** inglês.
b. O João **bem** sabe o que deve fazer (não sabe mal!) (Tescari Neto e Forero Pataquiva (2020: 495, NR 4)

[3] https://oglobo.globo.com/cultura/noticia/2022/06/luciano-szafir-e-internado-com-obstrucao-intestinal-apos-retirada-de-bolsa-de-colostomia.ghtml
[4] http://www.espn.com.br/noticia/232257_relator-volta-a-alterar-texto-da-lei-geral-da-copa-para-incluir-idosos-na-cota-social

Para o presente trabalho, observamos o posicionamento do verbo em relação a alguns advérbios baixo, quais sejam: *já*, *ainda*, *sempre*, *quase* e *bem*, que lexicalizam, respectivamente, as projeções de $T_{Anterior}$, $Asp_{Continuative}$, $Asp_{Perfect}$, $Asp_{Prospective}$ e Voice.

Outro refinamento necessário sobre o diagnóstico de movimento do verbo deve levar em consideração evidências como posição do sujeito e fronteamento de objetos. Isso dito, consideramos o fronteamento de objetos, pois o português em fases pretéritas apresentava uma gramática do tipo V2 flexível, como relatado na literatura (Ribeiro, 1995; Torres Morais, 1995; Paixão de Sousa, 2004; Gibrail, 2010; Antonelli, 2011; Galves e Paixão de Sousa, 2017; Galves e Gibrail, 2018; Galves, 2020). Desse modo, apenas o movimento do advérbio não nos ajudaria a caracterizar bem o que acontece nesses sistemas.

As línguas V2 têm duas propriedades principais: movimento do verbo para o núcleo de C e o movimento de apenas um sintagma para a posição de Spec,CP (Haegeman, 1996; Holmberg, 2015). Enquanto Vikner (1995) foi o primeiro a mostrar que, em sistemas V2, o verbo se move para a camada CP, periferia esquerda da sentença, Rizzi (1997) trouxe várias evidências de que o CP contém estruturas internas adicionais, ou seja, o CP inclui uma série de projeções funcionais distintas (13):

(13) ForceP> TopP *> FocP> TopP *> FinP

TopP hospeda tópicos, normalmente informação velha e de alguma forma disponíveis e salientes discursivamente. Essa posição aparece duas vezes, embora possa ocorrer um número indefinido de vezes. FocP abriga focos, expressando informações novas e em algumas línguas restringe-se ao foco contrastivo. FinP é uma projeção que reflete finitude, pelo fato de muitas línguas terem diferenças entre complementadores finitos e não finitos (Rizzi, 1997).

Considerando a hipótese do CP cindido e as hierarquias propostas por Rizzi (1997), alguns pesquisadores (Roberts, 2004; Holmberg, 2015) defendem que o efeito V2 estaria ligado ao traço de [+/-FINITUDE], pois, nessas línguas, FinP hospeda a sonda-φ e o traço de borda responsáveis por V2.

Com base em Benincà (1995), Wolfe (2018, 2019) expande essa proposta. Para ele, todos os sistemas V2 têm um movimento do verbo para Fin, porém, em uma subclasse das línguas V2, a projeção Force tem as mesmas propriedades de Fin. Antes de Wolfe, porém, Pinto (2011) já tinha proposto uma articulação entre Força e Finitude para explicar a variação no efeito V2[5].

[5] A estrutura proposta por Pinto (2011) difere da de Wolfe (2015) com relação à localização do núcleo

Para Wolfe (2018, 2019), as línguas V2 estariam divididas entre aquelas cujo local de pouso do verbo é ForceP e as outras cujo local de pouso é Fin. Assim, teríamos uma tipologia das línguas V2 de acordo com os locais de pouso:

(14) a. V-para-Fin (línguas com V2 menos estrito/mais flexível)
b. V-para-Force (línguas com V2 estrito) (Wolfe, 2018)

A intuição por trás dessa proposta é que "o núcleo com os traços responsáveis pelo fenômeno V2 encontra-se mais baixo na estrutura funcional e, portanto, permite a lexicalização de uma matriz de projeções funcionais estruturalmente mais altas que o verbo movido" (Wolfe, 2016, p. 14). Assim, as ordens V>2 são possíveis se os elementos mais à esquerda forem gerados *in-situ*. Nos sistemas FinV2, como nada *a priori* exclui um conjunto sucessivo dessas projeções que hospedam constituintes simultaneamente, o resultado dessas operações seria um sistema V2 "flexível" cujas ordens V1 e V>2 podem ocorrer além do verbo na segunda posição.

As línguas ForceV2 teriam um traço de borda que exige o movimento do verbo de Fin para Force e o movimento de um XP já fronteado para o Spec-ForceP. Nessas línguas ForceV2, as ordens V>2 são possíveis, porém são menos comuns. Isso se deve à área mais acima de Force estar associada a traços semântico-pragmáticos que codificam a atitude do falante.

Como o movimento do verbo é apenas metade do efeito V2 e outra metade é o fronteamento de um constituinte independentemente de sua categoria ou função sintática para a posição pré-verbal, iremos considerar a especificidade do fronteamento de objetos (Lightfoot, 1995; Holmberg, 2015) que é uma evidência em favor de V2, uma vez que esse tipo de construção apresenta uma configuração estrutural bastante específica OVS, ao mesmo tempo em que funciona como um bom diagnóstico de movimento do verbo.

Na próxima seção, apresentamos a metodologia utilizada para conduzir a presente pesquisa.

2 Metodologia (coleta dos dados, buscas e *corpus*)

Os dados desta pesquisa foram extraídos de 11 textos de teatro e 1 cantiga – anotados morfologicamente – e de 346 Cartas e 1330 Atas da Câmara Municipal de Salvador – anotados morfológica e sintaticamente, oriundos do *Corpus Tycho Brahe-Brasil* (cf. Araújo-Adriano; Corôa, 2022).

Frame.

Quadro 1 – Informações dos textos que compuseram o ***corpus*** de análise

Nascimento	Autor	Código Tycho Brahe	Publicação do texto[6]	Tipo Textual
1602	Rui de Carvalho Pinheiro	va_013	1650-1684	Carta
1630	Domingos Garcia de Aragão	va_014	1653	Carta
1650	Pedro Dias Pereira	va_017	1699-1710	Carta
1670	João de Couros Carneiro Filho	va_016	1699-1710	Carta
1670	Manuel Silveira de Magalhães	va_018	1710-1730	Carta
1672	Manuel Pessoa de Vasconcelos	va_017	1699-1710	Carta
1672	Manuel Pessoa de Vasconcelos	va_018	1710-1730	Carta
1700	João de Couros Carneiro Neto	va_018	1728-1730	Carta
1705	Antônio José da Silva	s_007	1737	Comédia
1710	Jerônimo Sodré Pereira	va_019	1751-1765	Ata
1720	Joaquim Rodrigues da Silveira	va_019	1751-1765	Ata
1720	Manuel José de Azevedo	va_020	1765-1775	Ata
1721	José Álvaro Pereira Sodré	va_020	1765-1775	Ata
1725	João Duarte Silva	va_020	1765-1775	Ata
1740	Domingos Caldas Barbosa	b_011	1798	Cantiga
1815	Martins Pena	p_003	1833	Comédia
1815	Martins Pena	p_004	1845	Comédia
1829	José de Alencar	a_008	1857	Comédia
1838	França Junior	j_001	1883	Comédia
1855	Artur de Azevedo	a_009	1891	Comédia
1880	Gastão Tojeiro	t_001	1920	Comédia
1934	Gianfrancisco Guarnieri	g_011	1957	Comédia
1956	Miguel Falabella	f_003	1990	Comédia
1978	Paulo Gustavo	g_012	2006	Comédia
1966	Paulo Sacaldassy	s_006	2007	Comédia

As atas e cartas da Câmara Municipal de Salvador (308.649 palavras) foram escritas por 18 escrivães, ao longo do século 17 e da primeira metade do século 18. Já as peças de teatro foram publicadas entre o século 18 e o século 21 e a cantiga no século 18 (233.183 palavras).

[6] Apesar de usarmos o título publicação dos textos, no caso das atas e cartas da Câmara Municipal de Salvador, consideramos o período de produção dos documentos.

O total de palavras analisadas, portanto, foi de 541.832. Esses textos estão todos disponíveis e com acesso livre no site do *Corpus* (https://www.tycho.iel.unicamp.br/corpus/catalogo.html). A seguir apresentamos o Quadro 2, com as informações por século, considerando a distribuição dos tipos textuais.

Quadro 2 – Distribuição do ***corpus*** por século e tipo textual

Século	Tipo textual	Quantidade de palavras
17	Cartas	90.676
18	Cartas	108.607
18	Atas	109.366
18	Comédia teatral	27.224
18	Cantiga	18.445
19	Comédia teatral	86.842
20	Comédia teatral	67.597
21	Comédia teatral	41.623

Como adiantado na introdução, o fenômeno escolhido para a presente análise é a ordem do verbo e o fronteamento de objetos entre os séculos 17 e 21. Considerando que o posicionamento do verbo em relação a advérbios é um diagnóstico fidedigno para movimento do verbo (cf. seção 1), foram computadas ocorrências de sentenças em que o verbo lexical e o verbo auxiliar precediam ou seguiam, imediatamente ou não, certos advérbios diagnósticos para o movimento do verbo, mais especificamente os advérbios baixos *já*, *ainda*, *sempre*, *quase* e *bem*. Isso porque expressões adverbiais como *hoje em dia*, *todos os dias*, *nunca*, não fornecem evidências muito claras de onde se posicionam na hierarquia aqui assumida e, consequentemente, não fornecem evidências claras no que diz respeito ao movimento do verbo na história do português.

Assim, baseado na ordem, a precedência de um verbo em relação ao advérbio, por exemplo, *quase*, em Spec-Asp$_{\text{Prospectivo}}$, equivaleria dizer que, da sua posição de base, isto é, dentro de *v*P, o verbo se moveu (pelo menos) para a seguinte projeção funcional (*por exemplo,* Asp$_{\text{Genérico/Progressivo}}$), de modo a gerar a ordem V_*quase*. Contrariamente, se o verbo é encontrado à direita do advérbio, gerando o ordenamento *quase*_V, por hipótese, somos levados a evidenciar uma instância mais baixa do movimento do verbo.

Em relação às buscas no *corpus*, para os casos de precedência do verbo em relação ao advérbio, template V_ADV, buscamos pela etiqueta VB-X, em que VB equivale a verbo lexical e X uma flexão qualquer, como D e P – etiqueta para pretérito perfeito e presente, respectivamente (*e.g.*, *comeu* = VB-D; *come* = VB-P). Para o template em questão, esse

verbo precede – não necessariamente uma precedência imediata, daí a etiqueta Precedes – um ADV, específico – daí a etiqueta iDoms –, por *bem*, por exemplo.

(15) a. Algoritmo de busca fornecido para encontrar V_ADV
((VB-D|VB-I|VB-P|VB-R|VB-RA|VB-SD|VB-SP|VB-SR|VB-F|
Precedes ADV|ADV-*)
AND (ADV|ADV-* iDoms bem|já|ainda|sempre|quase))

b. Algoritmo de busca fornecido para encontrar ADV_V
((ADV|ADV-* Precedes VB-D|VB-I|VB-P|VB-R|VB-RA|VB-SD| VB-SP|VB-SR|VB-F)
AND (ADV|ADV-* iDoms bem|já|ainda|sempre|quase))

Para os verbos auxiliares, buscamos pelas etiquetas ET-X|SR-X|HV-X|TR-X, em que ET equivale ao verbo *estar,* SR, ao verbo *ser*, HV, ao verbo *haver* e TR equivale ao verbo *ter* e X uma flexão qualquer, como P e D (*e.g., está* = ET-P; *tinha* = TR-D). Para o template ADV_AUX, um ADV precede – não necessariamente uma precedência imediata, daí a etiqueta Precedes – um verbo auxiliar. A busca pelas referidas etiquetas retorna quaisquer usos dos verbos *estar*, *ser*, *haver* e *ter*, não necessariamente auxiliares; como em *João tinha muito dinheiro*, em que *tinha é encontrado pela busca TR-D, por exemplo*. Foi necessário, portanto, fazer uma filtragem manual para haver apenas casos desses verbos no seu uso auxiliar, evidenciado quando precedem um verbo não finito (gerúndio, particípio e infinitivo[7]).

(16) a. Algoritmo de busca fornecido para encontrar ADV_V
((ET-D|ET-I|ET-P|ET-R|ET-RA|ET-SD|ET-SP|ET-SR|ET-F|SR-D|SR-I|SR-P|SR-R|SR-RA|SR-SD|SR-SP|SR-SR|SR-F|HV-D|HV-I|HV-P|HV-R|HV-RA|HV-SD|HV-SP|HV-SR|HV-F|TR-D|TR-I|TR-P|TR-R|TR-RA|TR-SD|TR-SP|TR-SR|TR-F Precedes ADV|ADV-*)
AND (ADV|ADV-* iDoms bem|já|ainda|sempre|quase))

b. Algoritmo de busca fornecido para encontrar ADV_V
((ADV|ADV-* Precedes ET-D|ET-I|ET-P|ET-R|ET-RA|ET-SD|ET-SP|ET-SR|ET-F|SR-D|SR-I|SR-P|SR-R|SR-RA|SR-SD|SR-SP|SR-SR|SR-F|HV-D|HV-I|HV-P|HV-R|HV-RA|HV-SD|HV-SP|HV-SR|HV-F|TR-D|TR-I|TR-P|TR-R|TR-RA|TR-SD|TR-SP|TR-SR|TR-F)AND (ADV|ADV-* iDoms bem|já|ainda|sempre|quase))

[7] Para critérios de auxiliaridade, cf. Lunguinho (2011), Araújo-Adriano (2018).

Com os dados de verbos ET/HV/SR em mãos, foi necessário fazer uma filtragem manual dos dados, visto que a busca não retorna apenas dados de estar/haver/ser como auxiliares.

Como ressaltado na seção anterior, se assumirmos que o português apresenta uma gramática V2 flexível, devemos considerar outras evidências para além do posicionamento do verbo. A questão sobre o local de pouso do verbo, ou seja, saber se o verbo se moveu para determinada posição, pode ser medida, entre outros, a partir da posição dos advérbios e do fronteamento de objeto que, consequentemente, levam à inversão do sujeito e ao alçamento do verbo. Neste artigo, iremos considerar também o fronteamento de objeto como evidência de um sistema V2 e, consequentemente, da perda desse sistema. Na próxima seção, trazemos os resultados. Os exemplos reportados na seção de resultados trarão o código do texto na Plataforma *Tycho Brahe*.

3 Resultados

Nesta seção, apresentamos os resultados referentes à posição do verbo em relação àqueles advérbios que podem dar luz sobre o movimento do verbo e o fronteamento de objetos na história do PB. Em geral, encontramos poucos dados em que o advérbio modifica o verbo e também poucos casos de fronteamento de objetos. Alguns foram úteis para diagnosticar movimento do verbo, outros não. Para distinguir verbo lexical de verbo auxiliar, usamos V para aquele e AUX para este. Apresentamos os dados agrupados da seguinte forma: século 17, séculos 18 e 19 e séculos 20 e 21, seguindo a hipótese levantada.

3.1 Século 17

Observando o posicionamento do verbo em relação ao advérbio *já*, pudemos notar que a maior parte dos dados com tal advérbio foi seguindo o item verbal, independente do tipo de verbo, quer lexical, quer auxiliar. No primeiro caso, 82% das ocorrências eram com V_*já* (17a) e apenas 18% com *já*_V (17b). No caso dos verbos auxiliares, a posposição do advérbio foi um posicionamento categórico (17c). Abaixo, trazemos alguns exemplos[8].

(17) Advérbio ***já***

a. e mais quando corre **já** por nossa conta o sal que Vossa Majestade nos largou pelo mesmo contrato, só se nos dá por ele uma limitada porção (va_017).

[8] Por recomendação de um dos pareceristas, reduzimos o número de exemplos.

b. O alívio que temos em ser governados por Vossa Majestade cujas honras grandezas **já** gozamos (va_013).
c. e por isso se tem movido **já** muitas demandas e desavenças pesadas entre si porque uns dizem que por provisão real antiga se não pode fundar novo engenho senão uma légua distante do outro, (va_016)[9].

No que diz respeito ao advérbio *ainda*, também observamos preferência por apresentar uma posição posterior ao verbo. A despeito de a precedência de *ainda*, encontrado nos dados, não poder ser tomada como evidência de ausência de movimento do item verbal, considerando que *ainda* pode se mover para a periferia esquerda da sentença (cf. seção 1), o item verbal precedendo o advérbio é evidência de movimento verbal. É isso que observamos com o verbo lexical, precedendo o advérbio *ainda* em 82% das ocorrências (18a). No caso dos verbos auxiliares, a posposição do advérbio foi um posicionamento encontrado em 60% das ocorrências (18b).

(18) **Advérbio *ainda***
a. e exerce **ainda** com toda a verdade e satisfação (va_016).
b. Esta disposição porém sem nova declaração de Vossa Majestade se foi observando **ainda** ao depois do dito ano de mil setecentos e treze (va_017).

Observando o posicionamento do verbo em relação ao advérbio *sempre*, pudemos constatar que a maior parte dos dados com tal advérbio foi seguindo o item verbal, independente do tipo de verbo, seja lexical, seja auxiliar. No primeiro caso, 68% das ocorrências eram com V_*sempre* (19a) e apenas 32% com *sempre*_V (19b). No caso dos verbos auxiliares, a posposição do advérbio foi um posicionamento também preferido, com 80% dos casos (19c), contra 20% do ordenamento *sempre*_AUX (19d). Abaixo, trazemos alguns exemplos.

(19) **Advérbio *sempre***
a. **Ficaremos** sempre reconhecidos e em tudo obedientes aas reais ordens de Vossa Majestade que Deus nos guarde como desejamos (va_017).

[9] O fato de não só o verbo auxiliar, mas também seu complemento verbal não finito precederem o advérbio *já* é evidência de que esses dois verbos se moviam juntos para uma posição mais alta na sentença. A esse respeito, cf. Araújo-Adriano e Cyrino (2022).

b. E **sempre** se praticou excetuando algumas vezes em que os ditos sindicantes as tomaram pelas acharem por tomar, (va_016).
c. E sendo todas estas tomadas na forma acima foram **sempre** aprovadas pelos sindicantes que as examinaram (va_013).
d. e com este privilégio e Instituição exorbitante **sempre** se há de contender restritamente (va_016).[10]

Em relação a *quase*, observamos que foi categórico o posicionamento do advérbio anteposto ao verbo auxiliar – não encontramos dados com o advérbio e um verbo lexical. A seguir, apresentamos alguns dos poucos exemplos que encontramos.

(20) **Advérbio *quase***
a. E assim chegaram neste ano os contratos a um nunca visto crescimento aplicando-se com tal eficácia a tudo o que toca à defesa e segurança desta praça, que a Fortaleza de Santo Antonio do Carmo que havia principiado seu antecessor está **quase** posta em sua última perfeição, (va_16).

Já em relação ao advérbio *bem*, encontramos verbos lexicais e auxiliares tanto o precedendo quanto o seguindo. Todos os casos de precedência eram casos em que *bem* veiculava uma leitura de foco, não sendo, portanto, evidência para ausência de movimento do verbo (cf. também seção 1). Nesse tipo de sentença, o advérbio veicula uma leitura aspectual. De acordo com Martins (1994), advérbios monossilábicos, quando em posição pré-verbal, perdem sua denotação original, o que chama de "esvaziamento semântico" desses elementos.

Por outro lado, encontramos evidência do verbo lexical precedendo *bem*, o que nos leva a afirmar que havia movimento verbal por sobre esse advérbio. Apesar de não termos encontrado nenhuma ocorrência de *bem* seguindo ou precedendo verbos auxiliares, o fato de o verbo auxiliar preceder advérbios mais altos que *bem*, por exemplo, *já* (17), *ainda* (18), *sempre* (19) e *quase* (20), por transitividade (Cinque, 1999, p. 33), pode-se dizer que o verbo funcional também se movia por sobre *bem*.

[10] Tratamos o verbo *haver* em *haver de* como um auxiliar, conforme demonstra Araújo-Adriano (2019, p. 75-86).

(21) **Advérbio *bem***

a. De que até aqui temos dito se mostra **bem** que uma das grandes obrigações em que a Vossa Majestade estamos é conservar-nos tantos anos o dito Governador (va_017).

3.1.1 Evidências de um sistema V2

Em línguas V2, o fronteamento de objeto é uma opção que altera a posição em que o sujeito ocorre; isso não acontece em línguas que não são V2. Quando o objeto é deslocado para a posição pré-verbal em línguas V2, a ordem gerada é OVS (Antonelli, 2001).

Nos dados do século 17, encontramos os seguintes casos de fronteamento de objetos diretos (em itálico, sinalizamos o sintagma que exerce função de sujeito; em sublinhado, função de objeto):

(22) a. <u>Algumas quantias que achamos</u> **receberam** *os ditos tesoureiros* (va_017).

b. <u>Esta mercê e singular esmola que esperamos</u> nos **conceda** a grandeza, e justiça de Vossa Majestade, (va_015).

c. e <u>a dita quantia</u> **deram** algumas pessoas obrigadas de afabilidade e bom modo de quem nos governa com amor e cristandade, que a não ser isto não permite o tempo cobrar-se coisa alguma, (va_015).

Em nossos dados, quando há na sentença fronteamento de objeto e sujeito expresso, categoricamente, o sujeito ocorre em posição pós-verbal. Por ser uma opção marcada, a ordem OVS é bastante rara. No *corpus* analisado, há apenas 43 ocorrências dessa ordem no século 17, de um total de 1551 casos de sentenças V2, o que soma cerca de 2,7% das sentenças.

3.1.2 Síntese dos dados do século 17

Para concluir a apresentação dos dados do século 17, na Tabela 1, sumarizamos as ocorrências do verbo lexical e auxiliar em relação aos advérbios.

Tabela 1 – Percentual do verbo lexical e auxiliar em relação a advérbios no século 17

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	18%	18%	32%		
V_adv	82%	82%	68%		100%
adv_AUX		40%	20%		
AUX_adv	100%	60%	80%	100%	

Com base nos percentuais apresentados, podemos observar que o verbo lexical se movia por sobre *já*, *ainda*, *sempre* e *bem*. Já os verbos auxiliares se moviam por sobre *já*, *ainda*, *sempre*, *quase* e, por transitividade, *bem*.

3.2 Século 18 e 19

Em relação ao posicionamento de *já* nos séculos 18 e 19, observamos preferência pela sua anteposição ao verbo lexical, ocorrendo em 86% das ocorrências no primeiro século (23b). Nota-se que os 14% de V_*já* (23a) podem ser tratados como resquícios de uma gramática V2, presente ainda no século 18, mas ausente no século 19, considerando 100% (23c) das ocorrências com o verbo lexical seguindo *já*.

(23) **Advérbio *já***

a. Quase dentro da mesma cidade tem Dona Izabel Maria Guedes de Britto uma rocinha que poderá ter em todo o seu circuito quatrocentas braças de terra as quais possuia **já** medidas e demarcadas pelo Desembargador Jozé da Costa Corrêa e cercadas com um valado (va_020).

b. Cala-te Sancho, cala-te, que **já** lá vai: és fiel companheiro! (s_004)

c. Nós **já** sabemos como foi o engano, neste armário … (p_004)

d. Só por eu cumprimentar o moço **já** o estava namorando. (b_011)

e. Não poder deferirmos por estar o negócio pendente como este esteja **já** decidido pela sentença que Vossa Mercê nos mandou (va_019).

f. **Já** está deitando o chapéu. (a_008)

Diferentemente dos verbos lexicais que, no século 18 apresentavam preferência por seguir o advérbio *já*, ainda no século 18 havia certa preferência pelo verbo auxiliar preceder *já*: isso ocorreu no *corpus* em 73% dos casos (23e). É no século 19 que a precedência de *já* a verbos auxiliares também passa a ser categórica, ocorrendo em 100% dos dados no segundo século (23f).

Os dados do século 18 podem ser tomados como evidência de que, nesse período, há um processo de competição de gramáticas, com resquícios de uma gramática clássica que vai entrando em declínio e uma gramática nova que vai surgindo, visto que, no PB, os verbos finitos aparecem consistentemente à direita do advérbio *já* (Modesto, 2001) e uma sentença em que a ordem *já*_V acontece só é possível em PB, se houver uma leitura de foco, diferente do que se verifica no português clássico.

Mais uma evidência de que há resquícios de uma gramática V2 no século 18 é a ocorrência de fronteamento de objetos na ordem O_V_S (24) (em um percentual de 0,19% das sentenças analisadas). No século 19, não encontramos dados de fronteamento de objetos.

(24) a. E o mesmo **disse** Mateus Pereira dos Santos Cavalcante (va_019).
b. E este termo **registara** *o escrivão da nova contribuição voluntária* nos seus livros para por ele haverem os ditos soldados os seus pagamentos (va_020)
c. Amira adornada de graça e lindeza, com que *a natureza* Seus dons **espalhou** (S_007).

Na pesquisa diacrônica, encontramos *ainda* seguindo o verbo lexical tanto no século 18 (25a), em 13% dos casos, quanto no século 19 (25b), em 19% dos dados, o que sugere que o verbo lexical se movia a uma posição acima da projeção funcional lexicalizada por esse advérbio. Apesar de termos encontrado posposição do verbo auxiliar e lexical em relação a *ainda*, não nos mostrando nada em relação ao movimento do verbo, tanto no século 18 (25c) quanto no 19 (25d) temos evidência de verbos auxiliares precedendo o advérbio, o que configura evidência de movimento dos auxiliares por sobre *ainda*.

(25) **Advérbio *ainda***
a. O requerimento sobre se mandar bater moeda provincial se não tem ainda proposto no conselho por razão de se ter ocupado com Rio e Minas, e Índia e não entrar **ainda** com as dependências da Bahia e estar reservado para se tratarem todas depois da saída da nau da Índia. (va_019)
b. Não observei **ainda**! (a_008)
c. O requerimento sobre se mandar bater moeda provincial se não tem **ainda** proposto no conselho por razão de se ter ocupado com Rio e

Minas, e Índia e não entrar ainda com as dependências da Bahia e estar reservado para se tratarem todas depois da saída da nau da Índia. (va_019)
d. Tenho **ainda** que ir arranjar mandado de despejo contra uma viúva, minha inquilina, que há três não me paga o aluguel da casa. (a_009)

Em relação ao advérbio *sempre*, há certa mudança no comportamento dos verbos ao longo dos séculos 18 e 19. Durante o primeiro século, os verbos lexicais não apresentavam tanta tendência de preceder o advérbio, precedendo-o em 35% (26a), sendo esse posicionamento altamente preferível no século 19 (26c), aparecendo em 82% dos dados. Já no que se refere aos verbos auxiliares, no século 18[11], a posposição do advérbio ao auxiliar ocorreu em todas as aparições do advérbio *sempre* (26d), assim como no século 19 (26e).

(26) **Advérbio *sempre***
a. E que aos mesmos criadores se lhe dariam **sempre** que entrassem com os seus gados, para os picarem nos açougues. (va_018)
b. **Sempre** se dará ao Senhor de Engenho, e Lavradores de destes livremente todo o gado, que lhes for necessário para o serviço do dito engenho. (va_019)
c. O senhor nunca nos incomoda, dá-nos **sempre** muito prazer. (j_001)
d. Pois seguro-te, meu bem, que apesar de tudo hei de ser **sempre** firme, constante, e leal. (s_007)
e. Se eu soubesse que havia de ser **sempre** tão feliz, casar-me-ia cinquenta. (p_004)

Não encontramos nos dados ocorrência de *quase* com verbos lexicais, apenas com auxiliares. Seu posicionamento em relação ao advérbio foi categórico: anteposição, tanto no século 18 quanto no 19, como se vê em ((27a) e (27b)), abaixo. Apesar de não termos evidência de *quase* com verbos lexicais, nesse período, o verbo lexical precedia advérbios mais altos que *quase*, como, *sempre* (26), *ainda* (25), o que, por transitividade, sugere que precedia também *quase*.

[11] Tanto no século 18 quanto 19, houve casos em que o advérbio *sempre* precedeu o verbo auxiliar. Foram casos de *sempre* veiculando uma leitura confirmativa, os quais não foram computados (cf. Lopes, 1998; Tescari Neto, 2013; Araújo-Adriano, 2022).

(27) **Advérbio *quase***

a. Foi, que Dom Tibúrcio com a pena de se ver acometido de três mulheres, como vossa mercê sabe, à vista das noivas e do sogro, tomou tal paixão, que lhe deu esta noite uma cólica, e está **quase** indo-se por um fio. (s_007)

b. Os recursos que eu possuía estão **quase** inteiramente esgotados. (a_009)

Quanto ao posicionamento do verbo em relação ao advérbio *bem*, tanto a posposição quanto a anteposição foram atestadas. Assim como no século 17, as anteposições ao verbo eram casos de *bem* na leitura de foco. Por outro lado, encontramos evidência do verbo lexical precedendo *bem*, o que nos leva a afirmar que havia movimento verbal por sobre esse advérbio nos séculos 18 (28a) e 19 (28b).

(28) **Advérbio *bem***

a. Dizem **bem**, que o mundo não é capaz de sustentar aquele globo esférico da formosura; e assim o ar é a pátria daquela estrela de Vênus. (s_004)

b. Vossa Excelência passa **bem**. (j_001)

Apesar de não ter sido atestada nenhuma ocorrência de *bem* seguindo ou precedendo verbos auxiliares, sua precedência em relação a advérbios mais altos que *bem*, como por exemplo, *sempre* (26), *ainda* (25), por transitividade, nos permite afirmar que, por hipótese, o verbo funcional também se movia por sobre *bem* tanto no século 18 quanto no 19.

3.2.1 Síntese dos dados dos séculos 18 e 19

Para concluir a apresentação dos dados dos séculos 18 e 19, na Tabela 2 e 3, sumarizamos as ocorrências do verbo lexical e auxiliar em relação aos advérbios. Podemos observar que o verbo lexical se move por sobre os advérbios *bem*, *sempre*, *ainda* e, por extensão, *quase*, advérbio mais baixo que os dois últimos, mas não se move por sobre *já*. Por outro lado, o verbo auxiliar se move por sobre advérbios mais altos que *bem*, como *quase*, *sempre*, *ainda* e por sobre *já* apenas no século 18, mas não mais no século 19.

Tabela 2 – Percentual verbo lexical e auxiliar em relação a advérbios nos séculos 18

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	86%	87%	65%		
V_adv	14%	13%	35%		100%
adv_AUX	27%	50%			
AUX_adv	73%	50%	100%	100%	

Tabela 3 – Percentual verbo lexical e auxiliar em relação a advérbios nos séculos 19

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	100%	81%	18%		
V_adv		19%	82%		100%
adv_AUX	100%				
AUX_adv		100%	100%	100%	

3.3 Século 20 e 21

No *corpus*, encontramos apenas uma ocorrência de *já* seguindo o verbo lexical, como se vê em (29a). Esse uso, entretanto, não é o prototípico do advérbio em questão. Aqui, *já* parece veicular uma ideia de *logo*, *daqui a pouco*. Em todos os outros usos do advérbio, na sua leitura caraterística de anterioridade ao evento, o verbo lexical e auxiliar precedia-o nos séculos 20 e 21 (29b e 29c).

(29) **Advérbio *já***

a. A Marcelina ainda não foi, mas vai **já**. (t_001)

b. Todo mundo **já** conhece essa história, Selma … (f_003)

c. Você **já** está trocando as pernas. (f_003)

Na pesquisa diacrônica, encontramos *ainda* seguindo, tanto no século 20 quanto no século 21, apenas o verbo auxiliar (30a), o que sugere que o verbo auxiliar se move nesse período a uma posição acima da projeção funcional lexicalizada por *ainda*. Foram também encontradas ocorrências de *ainda* precedendo o verbo lexical e auxiliar, o que não serve de evidência contra a presença de movimento verbal. Entretanto, uma vez que encontramos posposição de *ainda* apenas em relação a auxiliares, é possível afirmar que a ausência de ocorrência de posposição do advérbio em relação a verbos lexicais é evidência de ausência dessa configuração no período analisado.

(30) **Advérbio *ainda***

Eles quase se mataram hoje de manhã, eu vi que ia **ainda** sobrando para mim e pulei fora. (f_003)

Em relação a *quase*, foi atestada posposição apenas em relação a verbos auxiliares (31a), mas não com verbos lexicais, com os quais o advérbio apareceu posposto em todos os casos (31b).

(31) **Advérbio *quase***

a. Seu Otávio está **quase** brigando no botequim! (g_011)

b. Reclamações do prédio todo – a síndica **quase** deu na minha cara. (f_003)

Diferentemente dos séculos 18 e 19, a preferência de o verbo lexical seguir o advérbio *sempre* nos séculos 20 e 21 diminuiu: nesse período, o ordenamento *sempre*_V ocorreu em 80% das ocorrências (32a) e em apenas 20%, com V_*sempre* (32b). Diferentemente do verbo lexical, os verbos auxiliares categoricamente precederam o advérbio *sempre* no *corpus* (32c).

(32) Advérbio ***sempre***

a. Eu não sou mais criança, não, Maria Lúcia, e vou logo avisando que eu não vou admitir que vocês me tratem como uma débil-mental como **sempre** fizeram. (f_003)

b. Você tem **sempre** uma desculpa pronta. (t_001)

c. É por isso que a senhora está **sempre** cansada, vive me prometendo pancada! (g_011)

Nos dados dos séculos 20 e 21, o advérbio *bem* categoricamente seguiu o verbo auxiliar e o verbo lexical, como observado nos exemplos (33a) e (33b), abaixo.

(33) **Advérbio *bem***

a. Está com os pulmão arrebentando mas bebe **bem**. (g_011)

b. E como é que a gente vai enxergar **bem** com aqueles lençóis horríveis que tem que usar quando vira fantasma? (g_012)

3.3.1 Síntese dos dados dos séculos 20 e 21

Para concluir a apresentação dos dados dos séculos 20 e 21, nas Tabelas 4 e 5, sumarizamos as ocorrências do verbo lexical e auxiliar em relação aos advérbios observados até então. Podemos observar que o verbo lexical se move apenas por sobre *bem*, mas não por sobre *quase*, *sempre*, *ainda* tampouco *já*. Diferentemente, os verbos auxiliares, para precederem *bem*, *quase*, *sempre* e *ainda*, precisam se mover mais alto que *ainda*, mas não mais, considerando que não podem preceder *já*.

Tabela 4 – Percentual verbo lexical e auxiliar em relação a advérbios nos séculos 20

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	86%	87%	77%		
V_adv	14%	13%	23%		100%
adv_AUX	27%	47%			
AUX_adv	73%	53%	100%	100%	

Tabela 5 – Percentual verbo lexical e auxiliar em relação a advérbios nos séculos 21

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	100%	100%	86%		
V_adv			14%		100%
adv_AUX	100%	64%			
AUX_adv		36%	100%	100%	

3.3.2 Síntese da seção

No exposto nesta seção 3, apresentamos os resultados referentes ao posicionamento do verbo em relação aos advérbios *bem*, *quase*, *sempre*, *ainda* e *já* nos séculos 17, 18, 19, 20 e 21. Com esses dados diacrônicos, observamos que a ordem do verbo mudou ao longo do tempo. Em um primeiro momento, vimos em 3.1 que tanto o verbo lexical quanto o auxiliar podiam se mover por sobre *já* e todos os outros advérbios mais baixos. Esse ordenamento, como exposto em 3.2, já não era mais preferência para os verbos lexicais a partir do século 18, apesar de ter havido ainda alguns resquícios da gramática do século anterior. Essa gramática parecia resistir nos contextos de verbos auxiliares que ainda no século 18, assim como no século 17, precedem preferencialmente o advérbio *já*. No século 19, por outro lado, não há mais resquícios da gramática do século 17: tanto o verbo lexical quanto o verbo auxiliar obrigatoriamente seguem *já*, mas precedem todos os advérbios mais baixos: *ainda*, *sempre*, *quase* e *bem*.

Nos séculos 20 e 21, notamos, em 3.3, mais uma vez uma mudança no ordenamento do verbo: nesse período, verbos lexicais não mais precedem *ainda*, *quase* e *sempre*, mas continuam se antepondo em relação a *bem*, diferentemente dos verbos auxiliares que continuam a preceder todos esses advérbios, mas não *já*, assim como na gramática dos séculos 18 e 19. No Quadro 2, abaixo, sintetizamos os resultados quanto ao ordenamento do verbo português em relação a advérbios ao longo do tempo.

Quadro 2 - posicionamento do verbo lexical e auxiliar em relação a advérbios na diacronia do PB

Século	**Verbo**	**Posicionamento**	
		Precede	**Não precede**
17	lexical	*já, bem, quase, sempre, ainda*	-
	auxiliar	*já, bem, quase, sempre, ainda*	-
18	lexical	*bem, quase, sempre, ainda*	*já*
	auxiliar	*já, bem, quase, sempre, ainda*	
19	lexical	*bem, quase, sempre, ainda*	*já*
	auxiliar	precede *bem, quase, sempre, ainda*	*já*
20	lexical	*bem*	*já, ainda, quase, sempre*
	auxiliar	*bem, quase, sempre, ainda*	*já*
21	lexical	*bem*	*já, ainda, quase, sempre*
	auxiliar	*bem, quase, sempre, ainda*	*já*

4 Discussão

Pelo ordenamento do verbo em relação a advérbios, mostramos na seção anterior que o português brasileiro passou por três períodos: o **Período 1**, compreendendo o período do século 17, cuja gramática gerava um sistema V2 flexível; o **Período 2**, dos séculos 18 e 19, com o verbo lexical e auxiliar não precedendo *já*; e, por fim, o **Período 3**, compreendendo os séculos 20 e 21, com diferença no ordenamento dos verbos em relação ao tipo de verbo – verbos lexical não precedem *já, ainda, quase* e *sempre*, mas precedem *bem*, diferentemente do verbo auxiliar que precede todos esses advérbios, exceto *já*.

Assumindo a hierarquia universal das projeções funcionais e a proposta de Wolfe (cf. seção 1), vemos que no Período 1 os verbos lexicais e auxiliares se moviam para Fin (periferia esquerda da sentença), posicionamento que gera a precedência a todos os ad-

vérbios mais baixos que *já*. Já no Período 2, esse movimento não é mais atestado, movendo-se, verbos auxiliares e lexicais, para $T_{Anterior}$, lugar onde seguem *já*, mas precedem todos os advérbios mais baixos. Por fim, no Período 3, verbos auxiliares continuam se movendo para $T_{Anterior}$, diferentemente de verbos lexicais, que se movem para alguma posição mais baixa que a projeção onde se localiza *ainda* em $Asp_{Continuative}$, *sempre* em $Asp_{Perfect}$, *quase* em $Asp_{Prospective}$ mas mais alta que *bem* em Voice.

Como vimos, o português escrito por brasileiros perde as propriedades de língua V2 flexível no século 18. Com base em Chomsky (2008), o TP/IP apenas manifesta traços-φ e de Tempo (Tense) se for selecionado por C, pois T herda os traços de Tempo de C, que é um núcleo de fase (Roberts, 2019). Biberauer e Roberts (2017), seguindo Ouali (2008), estipulam a possibilidade de C doar, compartilhar ou manter os traços de Tempo de modo a explicar a variação paramétrica nas línguas. Desse modo, as línguas se diferenciariam se (i) C doa os traços de Tempo para T, se (ii) C compartilha os traços de Tempo com T, ou, se (iii) C retém os traços de Tempo. Retomando a proposta de Wolfe (2018, 2019) de que nas línguas V2 o local de pouso do verbo pode ser Force ou Fin, assumimos que Fin, a projeção funcional mais baixa do núcleo C (cf. seção 2), doa, compartilha ou retém os traços de Tempo.

Nas línguas V2 (exceto no islandês em que os traços de T são compartilhados com C), Fin retém o traço de Tempo, ou seja, T não herda esse traço, o que obriga o verbo a subir até a periferia esquerda e, consequentemente, o traço de borda atrai um XP qualquer para seu especificador. Portanto, herdar ou não o traço de Tempo de Fin tem um papel importante que distingue as línguas românicas das línguas germânicas.

Aliado a isso, lançamos mão da discussão levantada por Schifano (2018), que, a partir da estrutura articulada das projeções funcionais em (7), propõe que a hierarquia de Cinque (1999) pode ser reduzida em três grandes zonas funcionais: "uma zona alta, onde M(ood) é codificado, uma zona sentencial medial, onde T(ense) é codificado, e uma zona baixa, onde A(spect) é codificado" (Schifano, 2018, p. 134-135). Dessa forma, o movimento do verbo é visto como mirando zonas na estrutura sintática, por exemplo, Aspect, em vez de posições específicas, por exemplo, $Asp_{Frequentative}$. Para nossa discussão, em vez de uma estrutura articulada de núcleos funcionais, é suficiente assumir que IP é estruturado em zona Aspectual, Temporal e Modal. Assim, no Período 1, o verbo se movia para Fin, no Período 2, para a zona Temporal e no Período 3, apenas o verbo auxiliar continua se movendo para a zona T, enquanto o verbo lexical reduz seu movimento para a zona Aspect.

Junto de Schifano (2018), adotamos a proposta de que a propriedade que desencadeia movimento para essas zonas é a presença de um traço não interpretável [*u*F] em projeções funcionais como Fin, T, A ou M. Assumindo que *v* é sempre interpretável [*i*X] para Fin, T, Asp e M, portanto [*i*Fin], [*i*T], [*i*Asp] e [*i*M], Fin, T, Asp e M podem portar traços interpretáveis [*i*Fin], [*i*T], [*i*Asp] e [*i*M] ou não interpretáveis [*u*Fin], [*u*T], [*u*Asp] e [*u*M], respectivamente (note que nesse sistema traços são sempre valorados, podendo ser ou não interpretáveis). Pelo fato de que traços não interpretáveis devem ser elididos o mais rápido possível para serem lidos nas interfaces (Chomsky, 1995, p. 219–348), a maneira de deletar [*u*Fin], [*u*T], [*u*Asp] e [*u*M] portados pelas sondas Fin, T, Asp ou M é por uma relação de checagem com o alvo *v* via operação de movimento. Dessa maneira, quando as sondas têm traços interpretáveis correspondentes com os traços de *v*, esse alvo não se move: o movimento só é desencadeado se Fin, T, Asp ou M portarem um traço não interpretável, como apresentado abaixo.

(34) Checagem de traços entre *v* e Fin, T, A e M e consequências sintáticas
[Fin/M/T/$A_{[uX]}$ [… $v_{[iX]}$]]: movimento
[Fin/M/T/$A_{[iX]}$ [… $v_{[iX]}$]]: ausência de movimento

Dentro dessa abordagem, uma proposta explicativa para as diferentes ordens do verbo em relação a advérbios e, consequentemente, para as diferentes posições para onde o verbo se move na diacronia do PB seria da maneira como se segue. No período do século 17 em que tanto verbos auxiliares quanto verbos lexicais precediam *já*, gerando um sistema V2, propomos que Fin portava [*u*Fin], sendo o movimento de V e AUX, ambas as categorias com [*i*Fin], o responsável pelo apagamento desses traços não interpretáveis. No período seguinte, compreendendo os séculos 18 e 19, a perda dos traços não interpretáveis de Fin, perda comum em mudança linguística (cf. Roberts e Roussou, 2003), fez com que seus traços [*i*Fin] deixassem de desencadear movimento do verbo para essa posição alta. Como consequência, o verbo tanto lexical quanto auxiliar perde movimento para Fin, gerando um sistema que não é mais V2. Com essa perda de movimento, ambos os tipos de verbo passaram a seguir *já* mas preceder advérbios mais baixos, como *ainda*, *sempre etc*. Isso equivale a dizer que falantes dos séculos 18 e 19 moviam o verbo lexical e auxiliar para a zona T, que portava traços [*u*T]. Mais uma vez, há perda de traços [*u*T] em T, mas apenas em relação a verbos lexicais: verbos auxiliares continuam se movendo para a zona T, checando os traços [*u*T], precedendo advérbios mais baixos que *já*,

diferentemente de verbos lexicais que perderam movimento para a zona T, se movendo apenas para a zona Asp (cf. Cyrino, 2013, para quem verbos lexicais se movem em português contemporâneo até Aspecto), checando seus traços [*u*Asp] e, consequentemente, seguindo advérbios aspectuais como *sempre, quase* mas precedendo advérbios muito baixos, como *bem*. Assim, vislumbramos ao longo da história do PB três períodos:

(57) a. O sistema gramatical do PB no século 17
$[\text{Fin}_{[\text{uFin}]}[\text{Mood}_{[\text{uM}]}[\text{Tense}_{[\text{uT}]}[\text{Aspect}_{[\text{uAsp}]}[v_{\text{AUX}} - v - \text{VP}]]]]]$

b. O sistema gramatical do PB nos séculos 18 e 19
$[\text{Fin}_{[\text{iFin}]}[\text{Mood}_{[\text{iM}]}[\text{Tense}_{[\text{uT}]}[\text{Aspect}_{[\text{uAsp}]}[v_{\text{AUX}} - v - \text{VP}]]]]]$

c. O sistema gramatical do PB nos séculos 20 e 21
$[\text{Fin}_{[\text{iFin}]}[\text{Mood}_{[\text{iM}]}[\text{Tense}_{[\text{iT}]}[\text{Aspect}_{[\text{uAsp}]}[v - \text{VP}]]]]]$
$[\text{Fin}_{[\text{iFin}]}[\text{Mood}_{[\text{iM}]}[\text{Tense}_{[\text{uT}]}[\text{Aspect}_{[\text{uAsp}]}[v_{\text{AUX}}]]]]]$

A perspectiva que adotamos para mudança linguística é uma perspectiva emergentista, advinda da abordagem de Hierarquia de Parâmetros (Roberts 2012, Biberauer & Roberts 2012, 2015, Biberauer & Walkden 2015, Biberauer 2017, Roberts 2019a). Tal abordagem considera juntamente de Chomsky (2005) que, para se alcançar adequação explicativa, a Gramática Universal (GU) precisa ser a mais pobre possível, não envolvendo um número grande de parâmetros que diferenciariam uma língua da outra, tal como era na abordagem paramétrica tradicional (Chomsky, 1981). Contrariamente, a Faculdade da Linguagem seria formada por três fatores: GU (fator 1), dados linguísticos primários (fator 2) e princípios de eficiência computacional (fator 3).

Dada a pobreza da GU, a premissa básica da abordagem de Hierarquia de Parâmetros é que a variação "emerge de aspectos não subespecificados da GU e é estruturada por propriedades do terceiro-fator altamente emergindo da necessidade de aprendizagem eficiente" (Roberts, 2012, p. 321). Assim, parâmetros não fazem parte da GU, mas são determinados por estratégias de aprendizagem durante a aquisição da linguagem (sobre a relação entre mudança e aquisição da linguagem, cf. Lightfoot, 1979), *i.e.*, uma propriedade emergente da interação dos três fatores. A principal hipótese da abordagem de hierarquia de parâmetros é que a variação paramétrica não está presente em todos os itens lexicais, mas restrita a itens funcionais. Consequentemente, a parametrização emerge quando a criança entretém diferentes traços nos núcleos funcionais no léxico, como formulada pela Conjectura Borer-Chomksy, apresentada abaixo:

(35) Conjectura Borer-Chomsky (Baker, 2008, p. 353)
Todos os parâmetros de variação são atribuíveis às diferenças nos traços de itens particulares (*e.g.*, núcleos funcionais) no léxico.

Sob esse ponto de vista, as diferenças entre as línguas são um corolário de agregados de hierarquias de parâmetros, indo de um traço que é altamente espraiado pelos núcleos funcionais a traços que são altamente restritos a não todos mas a alguns núcleos funcionais. Isso dá origem a uma taxonomia de parâmetros (Biberauer & Roberts 2015, 2012, 2016, Biberauer 2017, Roberts 2019):

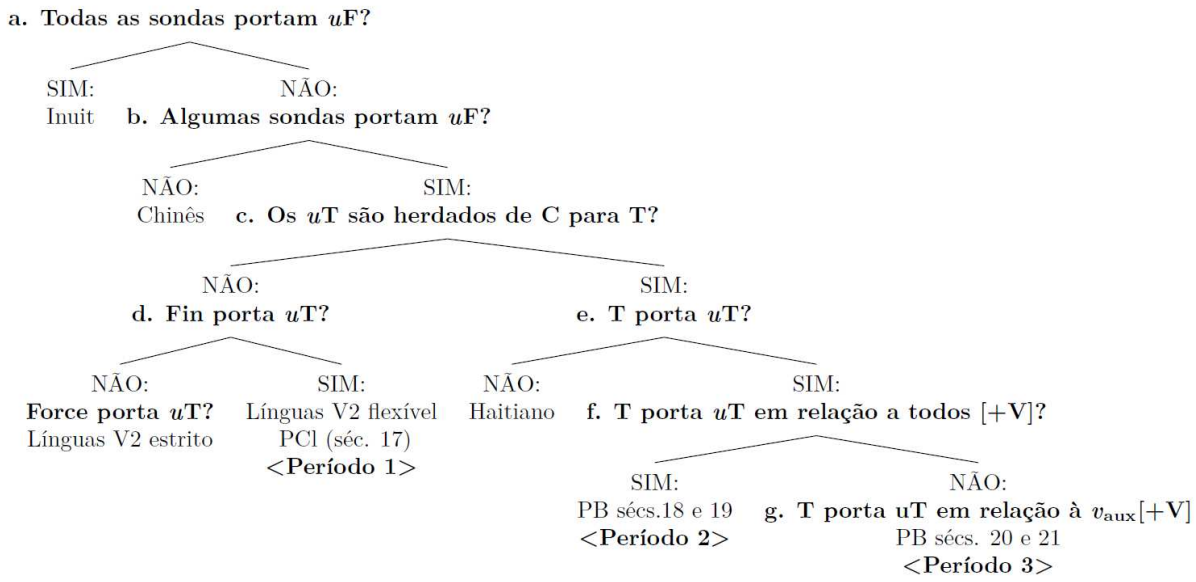
(36) Tipos de parâmetros
Para um dado valor v_i de um traço F parametricamente variante,
a. Macroparâmetros: todos os núcleos de um tipo relevante, e.g., todas as sondas, todos os núcleos fásicos etc compartilham v_i;
b. Mesoparâmetros: todos os núcleos de uma dada classe natural, e.g., [+V] ou uma categoria funcional central compartilham v_i;
c. Microparâmetros: uma subclasse pequena de núcleos funcionais (*e.g.* modais auxiliares, sujeito clíticos), lexicalmente definida compartilha v_i;
d. Nanoparâmetros: um ou mais itens lexicais individuais está/estão especificado/s para v_i.

A taxonomia acima combina núcleos funcionais e traços formais em conjuntos de tamanhos diferentes. Assim, os macroparâmetros e mesoparâmetros são vistos como agregados de tamanhos diferentes de microparâmetros, na medida em que representam diferentes distribuições de traços formais idênticos nos núcleos funcionais. Além disso, essas combinações capturam variações aparentemente discordantes sem proliferar parâmetros ou traços formais e restringem, de modo drástico, o espaço de gramáticas possíveis, criando dependências entre parâmetros (Roberts e Holmberg, 2010).

Como explicitamos anteriormente, a mudança nos três períodos evidenciados pode ser atribuída à presença/ausência de um traço formal que os núcleos funcionais portam. Baseado nos resultados diacrônicos, evidenciamos um período (a) em que Fin porta [*u*Fin] em relação a todos os elementos verbais, outro período em que (b) Fin perde seus traços [*u*Fin], reduzindo o escopo do movimento dos elementos verbais, que

passam a se mover para T, devido a seus traços [*u*T] em relação a v_{aux} e V. Finalmente, evidenciamos um terceiro período (c) em que T limitou seus traços [*u*T] a apenas verbos auxiliares, passando a ser [*i*T] em relação a verbos lexicais. Tendo em vista a abordagem da Hierarquia de Parâmetros, apresentamos a nossa proposta explicativa relativamente à mudança da ordem do verbo no PB.

(37)



No português clássico, as crianças não tinham evidência nos seus dados linguísticos primários que todas as sondas portavam um traço não interpretável (por exemplo, um N não se incorporava a T, como seria o caso se a criança estivesse exposta a uma língua como o inuíte *ti-seuan-mU-ban*/1SG:um-homem-ver-PASSADO 'eu vi o homem', Baker 1996, p. 327). Por isso, é plausível afirmar que as crianças desse período marcavam NÃO para (37a). Sondas que não portam um traço não interpretável é a instaciação de uma língua como o chinês, em que todos os núcleos são "flutuantes", como em *ben shu* CL livro 'livro' (Huang, 2015, p. 8). Uma vez que crianças brasileiras nunca tiveram acesso a esse tipo de *input*, sempre assinalaram SIM para (37b).

A variação entre os diferentes períodos evidenciados do português brasileiro tem origem onde o traço de tempo aparece, isto é, em (37c). Se os traços T ficam retidos em C, Force ou Fin pode portá-los. Se os traços T ficam retidos em Fin, desencadeia o movimento do verbo para Fin, resultando, consequentemente, em um sistema V2 flexível. Se Force, uma projeção mais alta que Fin, porta os traços T, o resultado é um sistema V2 estrito, em que o verbo se move para Force.

O que está em jogo é se os traços T são herdados por T (37c). No português clássico e no português escrito por brasileiros no século 17, os traços de T ficam "retidos" em Fin, evidenciados nos dados linguísticos primários pelo fronteamento do objeto como em *Algumas quantias que achamos* ***receberam*** *os ditos tesoureiros* (va_017), por isso, a criança desse período assinalava NÃO para (37c). Outra evidência de que Fin portava [*u*F] e não Force é a ocorrência de sentenças como *e [destes] [totalmente]* ***ficam*** *isentos os ricos (va_017)*, que apresenta ordem V>3, muito restrita em línguas ForceV2, mas muito comuns em línguas FinV2 (cf. seção 1) .

Já no português escrito por brasileiros no século 18, C doa os traços T para T, evidenciado nos dados linguísticos primários pela perda do fronteamento de objeto como em *Todos* minha paixão **vêm** (b_011), o que leva os escritores nascidos no século 18 a assinalar SIM para (37c), com consequente perda gradativa da propriedade de reter os traços T e passando a doar seus traços para T. Assim, vemos resquícios de um sistema V2 (ainda havendo registros da ordem OVS), principalmente em relação aos verbos auxiliares, como vimos em 3.1, o que pode ser explicado pelo que Kroch (1994; 2001) chama de competição de gramáticas, uma competição entre SIM e NÃO para o parâmetro em (37c).

A competição de gramáticas, também chamada de diglossia sintática, reflete um período de covariação entre opções gramaticalmente incompatíveis na fala e na escrita em uma comunidade de fala. As situações diglóssicas envolveriam o contraste entre a variedade mais conservadora (NÃO para (37c)) e a variedade mais inovadora (SIM para (37c)). Ao ligar o processo de mudança sintática à aquisição (Lightfoot, 1979), assume-se que a mudança sintática se relaciona a uma significativa alteração nos dados linguísticos primários. Essa alteração, embora se fixe de modo abrupto, aparece nos dados via competição de gramáticas, como ocorreu no século 18 (Kroch, 1994; 2001).

Assim, diante dos dados linguísticos primários, tendo marcado SIM para o parâmetro (37c), a criança brasileira então decide se T na sua língua porta ou não traços [*u*T] (37e). Caso fosse exposta a uma língua com morfemas temporais diretamente inseridos em MTA, como no haitiano *Mwen pa kwè pèsonn ap vini* (1SG NEG acreditar ninguém FUT vir 'Eu não acredito que ninguém vá vir'), cuja língua não tem traços que atraiam V (Roberts 2017, p. 333), o valor marcado para (37e) seria NÃO. Na história do português, entretanto, nunca existiram dados linguísticos primários como os do tipo haitiano que levassem o brasileiro oitocentista e novecentista a marcar (37e) diferente de um valor SIM.

Com evidências formadas por dados de precedência em relação a certos advérbios, como *ainda, sempre, quase etc* (*Não dei* ***ainda*** *motivos/Tenho* ***ainda*** *que comprar um par de*

botinas e fazer a barba ou *O senhor nunca nos incomoda, dá-nos* ***sempre*** *muito prazer/Se eu soubesse que havia de ser* ***sempre*** tão feliz, casar-me-ia cinquenta.), os escritores nascidos nos séculos 18 e 19 tinham manifestações suficientes de que T portava traços [*u*T] em relação a todos os elementos [+V], envolvendo tanto o verbo lexical quanto o verbo auxiliar. Como consequência, o valor SIM era marcado para (37f), ou seja, todos os verbos finitos eram sondados por T via [*u*T].

Outra variação entre os diferentes períodos evidenciados do português brasileiro tem origem na reanálise que as crianças dos séculos 18 e 19 fizeram em relação ao traço [*u*T] de T. Até o século 19, a criança brasileira tinha uma gramática estável, com seus dados linguísticos primários apontando em direção à resposta SIM para o parâmetro em (37f), produzindo sentenças com verbos auxiliares, mas também lexicais, precedendo certos advérbios. Por algumas razões externas[12], seja por princípios de economia, seja por contato linguístico, os verbos lexicais no português brasileiro perderam interpretações temporais (cf. Cyrino, 2013; Tescari Neto *et al.*, 2021), o que pode ser a razão pela qual as crianças perderam evidência de [*u*T] em T em relação ao verbo lexical. Isso significa dizer que verbos lexicais, mas não auxiliares, deixaram de estabelecer relação com T, passando a se relacionar apenas projeções mais baixas, como Aspecto (cf. Cyrino, 2013). Portanto, com uma gramática instável, crianças no século 20 e 21 tiveram que realizar seus dados e "mover para baixo" na Hierarquia de Parâmetros para encontrar a melhor solução paramétrica para os dados linguísticos primários.

Consequentemente, os dados com apenas verbos auxiliares, mas não lexicais, precedendo advérbios como *ainda*, *sempre*, *quase* não mais se conformam com a resposta SIM para (37f), fazendo com que as crianças dos séculos 20 e 21 restringissem os elementos sondados por T a apenas verbos auxiliares, estabilizando sua gramática com a resposta NÃO para (37f). Assim, ao longo do tempo, vemos marcações paramétricas distintas, emergindo sistemas linguísticos também distintos:

(38) Português Clássico Colonial (SIM > 38d) >
Português Brasileiro Colonial (SIM > 38f) >
Português Brasileiro moderno (NÃO > 38f).

Nota-se, portanto, que uma abordagem tal como a da Hierarquia de Parâmetros, apesar de ter sido desenvolvida pensando na comparação sincrônica entre línguas (Roberts 2012, Biberauer; Roberts 2012, 2015, Biberauer; Walkden 2015, Biberauer

[12] Questões que, por motivos de escopo, nos limitam a uma discussão mais aprofundada.

2017, Roberts 2019), é também uma abordagem profícua para a diacronia. Assim como cada marcação paramétrica leva a criança a adquirir uma língua-alvo diferente, no âmbito da mudança linguística, cada marcação paramétrica também levaria a criança a uma língua-alvo diferente, apesar de essas línguas-alvo distintas serem estágios da mesma língua em períodos diferentes.

De que modo, então, a análise apresentada contribui para repensarmos as propostas de periodização existentes? Os dados fornecem evidências de que a subida do verbo passou por três períodos de diferenciação: **Período I,** que chamaremos de Português Clássico Colonial, em que a língua transplantada para a colônia apresenta o mesmo padrão do português clássico. Encontramos evidências – subida dos verbos auxiliares e lexicais por sobre o advérbio *já* e ocorrência de ordem OVS – de que os textos escritos por brasileiros ao longo do século 17 apresentam uma gramática V2; **Período II**, que chamaremos de **Português Brasileiro Colonial**[13], em que o verbo lexical perde movimento, deixando de subir por sobre *já*, mas ainda com subida nos verbos auxiliares e resquícios de ordem OVS; **Período III (PB moderno)**, em que a perda do movimento verbal é mais uma evidenciada nos verbos lexicais, mas não nos verbos auxiliares. Como base nisso, propomos a seguinte periodização:

Nova proposta de periodização para o português brasileiro		
Nossa proposta (2025)		
1ª fase Português Clássico Colonial (até 1700)	2ª fase Português Brasileiro Colonial (1700-1899)	3ª fase Português Brasileiro Moderno (1900-atual)

A periodização proposta aqui parte de evidências linguísticas, ou seja, da história interna da língua. Estamos corroborando aqui propostas anteriores de que (i) as raízes do PB emergem no século 18, (ii) há uma fase de transição entre uma gramática clássica, língua das caravelas (Galves, 2007), e a gramática do PB moderno. O *Português Brasileiro Colonial*, fase de transição entre uma gramática clássica e a gramática do PB moderno, não é um reflexo da mudança português clássico > PE, mas sofre influência superficial dessa mudança através do processo de gramatização/normatização que acontece no Brasil, sobretudo, a partir da segunda metade do século 19 (cf. Pagotto, 1998). Defendemos também, com base em Corôa (2022), *Português Brasileiro Colonial* pouco se assemelha a gramática do PE, pois, no século 18, já há indícios de fenômenos idiossincráticos da

[13] Pontuamos, que a denominação adotada, apesar de parecida, diverge do que propõem Cardoso, Andrade e Carneiro (2023), pois concebemos que a gramática do que chamamos de Português Brasileiro Colonial constitui marcas de uma gramática vernacular brasileira, ao contrário do que afirmam aqueles autores.

gramática do PB como restrição da inversão do sujeito aos verbos inacusativos, generalização da próclise nas sentenças matrizes, sujeitos indeterminados e casos de objeto nulo referencial.

Apesar de a nossa proposta tomar por base fatores internos à língua, não deixa de refletir alguns aspectos trazidos nas periodizações de:

(i) Marlos Pessoa de Barros (2003), quando vemos coincidir, em certa medida, a segunda fase da proposta, chamada de etapa da formação do *português comum brasileiro* e o que chamamos de *Português Brasileiro Colonial*;

(ii) Volker Noll (2008), quando coincidem a fase diferenciadora, em que, segundo Noll, há a formação do português brasileiro e diferenciação das variedades europeia e brasileira e o que chamamos de *Português Brasileiro Colonial*.

Por fim, ressaltamos ainda a importância de entrelaçar os dados trazidos nesta investigação com as propostas de periodização elencadas na introdução a fim de melhor compreender os fatores sócio-históricos e linguísticos na diacronia da língua portuguesa no Brasil.

5 Considerações finais

Este artigo discutiu a periodização do português brasileiro, a partir de evidências do diferente ordenamento do verbo em relação a advérbios ao longo da história. Analisamos os resultados referentes ao posicionamento do verbo em relação aos advérbios *bem*, *quase*, *sempre*, *ainda* e *já* pelos séculos 17, 18, 19, 20 e 21. Com base nesses dados diacrônicos, observamos que a ordem do verbo sofreu três mudanças mais significativas ao longo do tempo.

Inicialmente, tanto o verbo lexical quanto o auxiliar podiam preceder *já* e de todos os outros advérbios mais baixos, além de termos evidências de fronteamento de objetos, com o ordenamento OVS, que aponta para uma gramática do tipo V2 (Período 1). Esse ordenamento deixou de ser a preferência para os verbos lexicais a partir do século 18, embora ainda houvesse alguns vestígios da gramática do século anterior. Essa gramática parecia persistir nos contextos com verbos auxiliares, que ainda no século 18, assim como no século 17, preferencialmente precediam o advérbio "*já*". No século 19 (Período 2), entretanto, não há mais vestígios da gramática do século 17: tanto o verbo lexical quanto o auxiliar obrigatoriamente seguem *já*, mas precedem todos os advérbios mais baixos: *ainda*, *sempre*, *quase* e *bem*. Nos séculos 20 e 21 (Período 3), observamos novamente outra mudança na ordenação do verbo. Durante esse período, os verbos lexicais

não mais precedem *ainda*, *quase* e *sempre*, mas continuam se posicionando antes de *bem*. Diferentemente, os verbos auxiliares continuam a preceder todos esses advérbios, exceto *já*, mantendo o comportamento observado na gramática dos séculos 18 e 19.

Formalizamos essa mudança adotando a abordagem da Hierarquia de Parâmetros, propondo que no Período I os traços não interpretáveis em Fin [*u*Fin] desencadeava movimento do verbo para Fin, gerando ordenamento que precedia todos os advérbios analisados e uma gramática V2. Já no Período 2, esses traços são perdidos, com Fin passando a portar traços [*i*Fin]. Nesse período, T portava traços [*u*T] em relação a todos os elementos verbais, desencadeando movimento não só do verbo lexical, mas também dos verbos auxiliares. Por fim, no Período 3, T manteve seus traços em relação a verbos auxiliares, mas restringiu seus traços em relação a verbos auxiliares, passando a portar traços [*i*T], gerando diferenças em relação ao movimento do verbo: apenas verbos auxiliares se movem até a zona T, diferentemente dos verbos lexicais que se movem para uma posição mais baixa que T.

Os dados apontaram a necessidade de revisão das propostas de periodização para a língua portuguesa no Brasil. Por isso, propomos uma periodização que está de acordo com as mudanças significativas reportadas pela ordem do verbo. A periodização proposta divide-se nas seguintes fases: Período I, que chamamos de Português Clássico Colonial; Período II, denominada de Português Brasileiro Colonial; e, a Período III, nomeada de Português Brasileiro Moderno.

Sabemos que são necessárias outras análises, incluindo outros fenômenos sintáticos, que possam confirmar e/ou refutar os dados aqui apresentados. Por ora, deixamos a nossa contribuição sobre a necessidade de voltarmos nosso olhar sobre a sintaxe do período, para que possamos melhor compreender quando a gramática brasileira emergiu.

Referências

ANDRIANI, L. The modal value of ancora/angórə in Barese and northern Apulian varieties. *Cambridge Occasional Papers in Linguistics*, Cambridge, UK, v. 9, p. 235–245, 2016.

ANTONELLI, A. L.. Posição do verbo no português clássico: evidências de um sistema V2. Alfa: Revista de Linguística (São José do Rio Preto), v. 55, n. 2, p. 501–522, jul. 2011.

ANTONELLI, A.L. Sintaxe da posição do verbo e mudança gramatical na história do português europeu. 230f.p. Tese (Doutorado em Linguística)—Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, 2011.

ARAÚJO-ADRIANO, P. Â. The position of the verb with respect to the adverb `sempre' over four centuries: diagnosis for the (loss of high) verb movement in Brazilian Portuguese. *Journal of Historical Syntax*, [*s. l.*], v. 6, n. 3, p. 1–50, 2022.

ARAÚJO-ADRIANO, P. Â.; CORÔA, W. Reconstruindo a história do português do brasil pelo corpus Tycho Brahe Brasil: novos dados, novos olhares. *Revista Linguíftica*, [*s. l.*], v. 17, n. 3, p. 202–227, 2022.

BAKER, M. The macroparameter in a microparametric world. *In*: BIBERAUER, T. (ed.). *The Limits of Syntactic Variation*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2008. p. 351–374.

BELLETTI, A. *Generalized verb movement: aspects of verb syntax*. Torino: Rosenberg & Sellier, 1990.

BIBERAUER, T. & ROBERTS, I. Towards a parameter hierarchy for auxiliaries: diachronic considerations. *Cambridge Occasional Papers in Linguistics* 6(9). 267–294. Place: Cambridge, UK, 2012.

BIBERAUER, T. & ROBERTS, I. Rethinking formal hierarchies: a proposed unification. *Cambridge Occasional Papers in Linguistics* 7(1). 1–31. Place: Cambridge, UK, 2015b.

BIBERAUER, T & WALKDEN, G. (eds.). *Syntax over time: lexical, morphological, and information-structural interactions* (Oxford studies in diachronic and historical linguistics 15). Oxford, United Kingdom: Oxford University Press, 2015.

BIBERAUER, T. Factors 2 and 3: A principled approach. *Cambridge occasional papers in linguistics* 10. 38–65. Place: Cambridge, UK, 2017.

BIBERAUER, T. & ROBERTS, I. Parameter setting. In Adam Ledgeway & Ian G. Roberts (eds.), *The Cambridge Handbook of historical syntax*/Cambridge Handbooks in Language and Linguistics, 234–162. Cambridge: Cambridge University Press, 2017.

CARDOSO, L.; ANDRADE, A. L. D.; CARNEIRO, Z. O português colonial brasileiro: uma nova agenda de pesquisas entre o português clássico e o português brasileiro moderno. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada,* [*s. l.*], v. 39, n. 2, p. 1–27, 2023.

CHOMSKY, N. *Lectures on government and binding: the Pisa Lectures*. Dordrecht, Holland; Cinnaminson, [N.J.]: Foris Publications, 1981. (Studies in generative grammar, v. 9).

CHOMSKY, N. Three Factors in Language Design. *Linguistic Inquiry,* Massachusetts, v. 36, n. 1, p. 1–22, 2005.

CHOMSKY, N. On phases. *In*: FREIDIN, R.; OTERO, C. P.; ZUBIZARRETA, M. L. (ed.). *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. [*S. l.: s. n.*], 2008. (Current studies in linguistics). p. 133–166.

CHOMSKY, N. *The Minimalist Program*. Massachusetts: The MIT Press, 1995.

CINQUE, G. (org.). *Restructuring and Functional Heads. The Cartography of Syntactic Structures*. [*S. l.*]: Oxford University Press, 2006. v. 4

CINQUE, G. *Adverbs and functional heads.* Oxford; New York: Oxford University Press, 1999.

CORÔA, W. *Rastreando as origens do Português Brasileiro: a dinâmica da mudança na escrita de "Homens Bons" na Bahia Colonial.* 2022. 406 f. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, Campinas, SP, 2022.

CYRINO, S. On richness of tense and verb movement in Brazilian Portuguese. *In*: CAMACHO-TABOADA, V. *et al.* (ed.). *Information Structure and Agreement*. Amsterdam: John Benjamins Publishing Company, 2013. p. 297–317.

EMONDS, J. The verbal complex V'-V in French. *Linguistic Inquiry*, Massachusetts, v. 9, n. 2, p. 151–175, 1978.

GALVES, C. A língua das caravelas: periodização do português europeu e origem do português brasileiro. In: CASTILHO, A. d.; MORAES, M. A. T.; LOPES, R.; CYRINO, S.(Ed.). *Descrição, história e aquisição do português brasileiro*. Campinas: Pontes, 2007. p. 513–528.

GALVES, C. Relaxed verb second in classical portuguese. In: WOODS, R.; WOLFE, S. (Ed.). *Rethinking verb second.* Oxford: Oxford University Press, 2020. p. 368–395.

GALVES, C.; PAIXÃO DE SOUSA, M. C. The change in the position of the verb in the history of portuguese: Subject realization, clitic placement, and prosody. *Language, Linguistic Society of America*, v. 93, n. 3, p. e152–e180, 2017. ISSN1535-0665.

GALVES, C.; GIBRAIL, A. Subject inversion intransitive sentences from classical to modern European Portuguese: A corpus-based study. In: MARTINS, A. M.; CARDOSO, A. (Ed.). *Word Order Change*. Oxford: Oxford University Press,2018.

GIBRAIL, A. *Contextos de formação de estruturas tópico foco no Português Clássico*. 2010. 298 f. Tese (Doutorado em Linguística) — Universidade Estadual de Campinas (UNICAMP). Instituto de Estudos da Linguagem, Campinas, 2010.

HAEGEMAN, L. Verb-second, the split cp and null subjects in early Dutch finite clauses. *GenGenP* (Geneva Generative Papers), v. 4, p. 133–175, 1996.

HOLMBERG, A. Verb second. In: KISS, T.; ALEXIADOU, A. (Ed.). *Syntax theory and analysis*: an international handbook of contemporary syntactic research. Berlin:[s.n.], 2015. P. 342–383.

KATO, M. O português são dois… ou três?. *In*: LOBO, T. *et al.* (org.). *Rosae: linguística histórica, história das línguas e outras histórias*. Salvador: EDUFBA., 2012. p. 93–108.

KROCH, A. S. Syntactic Change. *In*: BALTIN, M.; COLLINS, C. (ed.). *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell Publishers, 2001. p. 699–729.

KROCH, Anthony. Morpho-syntactic variation. In: BEALS, Kenneth et al. (Orgs.). *Papers from the 30th Regional Meeting of the Chicago Linguistics Society:* Parasession on Variation e Linguistic Theory, Chicago, v. 2, p. 180-201, 1994.

LEDGEWAY, A.; LOMBARDI, A. Verb Movement, Adverbs and Clitic Positions in Romance. *Probus,* [*s. l.*], v. 17, n. 1, 2005. Disponível em: Acesso em: 20 mar. 2022.

LIGHTFOOT, D. *Principles of diachronic syntax*. Cambridge, UK: Cambridge University Press, 1979.

LUNGUINHO, M. V. da S. Verbos auxiliares e a sintaxe dos domínios não-finitos. 2011. 215 f. Tese – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2011.

LOBO, T. A questão da periodização da história linguística do Brasil. In: CASTRO, I.; DUARTE, I. (Ed.). *Razões e emoção: miscelânea de estudos em homenagem a Maria Helena Mateus.* [S.l.]: Imprensa Nacional-Casa da Moeda, 2003. p. 395–409.

LOPES, A. C. M. Contribuição para o estudo dos valores discursivos de "sempre". *In*: *ACTAS DO XIII ENCONTRO NACIONAL DA APL* 1997. Lisboa: Associação Portuguesa de Linguística, 1998. v. II, p. 3–14.

LUCCHESI, D. A periodização da história sociolinguística do Brasil. *DELTA: Documentação de Estudos em Lingística Teórica e Aplicada*, Fap UNIFESP (SciELO), v. 33, n. 2, p. 347–382, aug 2017.

MARTINS, A. M. Clíticos na história do português. Tese (Doutorado)—Universidade de Lisboa, 1994.

MATOS, G.; CYRINO, S. *Elipse de VP no Português Europeu e no Português Brasileiro*. Universidade Federal do Ceará: [*s. n.*], 2001.

PAIXÃO DE SOUSA, M. C. Língua Barroca: Sintaxe e História do português nos 1600. Tese (Doutorado emLinguística)—Universidade Estadual de Campinas (UNICAMP), Campinas, 2004.

PESSOA, M. d. B. *Formação de uma variedade urbana e semi-oralidade*: o caso do Recife, Brasil. Tubigen: Niemeyer, 2003.

POLLOCK, J.-Y. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry*, [*s. l.*], v. 20, n. 3, p. 365–424, 1989.

RIBEIRO, I. A sintaxe da ordem no português arcaico: o efeito V2. Unicamp. 286f. Tese (Doutorado emLinguística)—Universidade Estadual de Campinas (UNICAMP). Instituto de Estudos da Linguagem, Campinas, 1995.

RIBEIRO, I. A mudança sintática do português brasileiro é mudança em relação a que gramática. In: CASTILHO, A. T. d. (Ed.). *Para História do Português Brasileiro*. São Paulo: Humanitas, 1998. p. 101–119.

RIZZI, L. The fine structure of the left periphery. In: HAEGEMAN, L. (Ed.). *Elements of grammar*. Dordrecht: Springer, 1997. p. 281–337.

ROBERTS, I. Macroparameters and minimalism: A programme for comparative research. *In*: GALVES, C. *et al.* (ed.). *Parameter theory and linguistic change*. Oxford: Oxford University Press, 2012. (Oxford studies in diachronic and historical linguistics, v. s). p. 319–334.

ROBERTS, I. On some languages lacking V-to-I movement. *A Schrift to Fest Kyle Johnson*, Linguistics Open Access Publications, p. 321–335, 2017.

ROBERTS, I. *Parameter hierarchies and Universal Grammar.* Oxford: Oxford University Press, 2019.

ROBERTS, I.; ROUSSOU, A. *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge, UK: Cambridge University Press, 2003. (Cambridge Studies in Linguistics, v. 100).

ROBERTS, I. The C-System in Brythonic Celtic Languages, V2, and the EPP, *In*: Luigi Rizzi (ed.), *The Structure of CP and IP*: The Cartography of Syntactic Structures, Volume 2, New York, NY, 2004; online edn, Oxford Academic, 2023.

ROBERTS, I. *Parameter hierarchies and Universal Grammar*. Oxford: Oxford University Press, 2019.

SAPIR, Edward. *Language*: an introduction to the study of speech San Diego: HJB Books, [1921] 1949.

SILVA NETO, Serafim da. Introdução ao estudo da língua portuguesa no Brasil. 2 ed. Rio de Janeiro: INL, 1986[1951].

TARALLO, F. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX. *In*: ROBERTS, I.; KATO, M. (org.). *Português brasileiro: uma viagem diacrônica*. 3. ed. São Paulo: Editora Contexto, 1993. p. 54–82.

TESCARI NETO, A. On the Movement of Verbal Forms in Romance and English. *Language at the University of Essex (LangUE) Proceedings 2011*, [*s. l.*], p. 54–67, 2012.

TESCARI NETO, A. *On Verb Movement in Brazilian Portuguese: A Cartographic Study*. 2013. 392 f. Tese (Phd em Filosofia) - Università Ca' Foscari di Venezia, Veneza, 2013.

TESCARI NETO, A. Advérbios e o movimento do verbo. *Fórum Linguístico*, [*s. l.*], v. 16, n. 1, p. 3563–3578, 2019.

TESCARI NETO, A.; FORERO PATAQUIVA, F. de P. Do movimento do verbo finito e infinitivo em português brasileiro e espanhol colombiano: microvariação e cartografias. *Cuadernos de la ALFAL*, [*s. l.*], v. 12, n. 2, p. 491–511, 2020.

TESCARI NETO, A.; FORERO PATAQUIVA, F. de P.; WECHSLER, A. L. R. Hipóteses sobre o que desencadearia o movimento do verbo. *In*: III *Encontro de gramática gerativa homenagem a Sonia Cyrino e Maria Eugênia Duarte*, 2021, Online. Anais [...]. Online: [*s. n.*], 2021.

TEYSSIER, P. *História da língua portuguesa*. SãoPaulo: Martins Fontes,1984.

TORRES MORAIS, M. A. *Do português clássico ao português europeu moderno: um estudo diacrônico da cliticização e do movimento do verbo*. 1995. 374 f. Thesis (PhD in Linguistics) - Universidade Estadual de Campinas, Campinas, SP, 1995.

NOLL, V. *O português brasileiro: formação e contrastes*. São Paulo: Globo, 2008.

WOLFE, S. A comparative perspective on the evolution of Romance clausal structure. *Diachronica*, JohnBenjamins, v. 33, n. 4, p. 461–502, 2016.

WOLFE, S. *Verb second in medieval Romance*. Oxford: Oxford Studies in Diachronic, 2018.

WOLFE, S. Redefining the typology of v2 languages: *The view from medieval Romance and beyond*. Linguistic Variation, JohnBenjamins, Amsterdam; NewYork, v. 19, n. 1, p. 16–46, 2019.



http://doi.org/10.5007/1984-8420.2025.e100437

VERB ORDER IN THE HISTORY OF BP: ISSUES FOR PERIODISATION[1]

A ORDEM DO VERBO NA HISTÓRIA DO PB: QUESTÕES PARA PERIODIZAÇÃO

Paulo Ângelo Araújo-Adriano | Lattes | pauloangelo@usp.br
Universidade de São Paulo – USP

Williane Corôa | Lattes | williscoroa@gmail.com
Universidade do Estado da Bahia – UNEB

Abstract: The main goal of this paper is to present new elements for the periodisation of Brazilian Portuguese (BP). Under the field of diachronic syntax, from a generative perspective, we analyse verb order in morphologically annotated texts, which are part of the *Tycho Brahe-Brasil Corpus*, in order to offer new insights into the periodisation of Portuguese in Brazil. The texts cover a wide period, from the 17th century to the 21st century, and were written by individuals born in Brazil. We present evidence that written Portuguese in Brazil exhibits three major periods: Period I, whose grammar reflects properties of Classical Portuguese, referred to as "the language of the caravels" by Galves (2007); Period II, during which lexical verbs and auxiliaries lose movement, ceasing to move to the left periphery of the sentence but rising to a Tense position; and Period III, in which there is additional loss of movement in relation to lexical verbs but not to auxiliaries. These results appear to validate the hypothesis advocated by Ribeiro (1998) and Galves (2007) that the origins of BP precede the 19th century and help to reconsider the periodisation of the Portuguese language in Brazil.

Key-words: Brazilian Portuguese periodisation. Verb order. Verb movement.

Resumo: O objetivo principal deste artigo é apresentar novos elementos para a periodização do português brasileiro (PB). Situamo-nos no campo disciplinar da Sintaxe diacrônica, na perspectiva gerativista, e analisamos a ordem do verbo em textos – anotados

[1] This paper was presented at the 30th GELNE Conference (Grupo de Estudos Linguísticos e Literários do Nordeste), held in Feira de Santana, at the Universidade Estadual de Feira de Santana, in December 2024. We thank the audience for their questions and suggestions. Additionally, we extend our gratitude to the anonymous reviewers for their contributions, questions, and comments on this text.

morfológica e sintaticamente – que fazem parte do *Corpus Tycho Brahe-Brasil*, com fins de oferecer novos elementos sobre a periodização do português no Brasil. Escritos por indivíduos nascidos no Brasil, os textos dão conta de um grande período: desde o século 17 até o século 21. Apresentamos evidências de que o português escrito no Brasil apresenta três grandes períodos: o Período I, cuja gramática reflete propriedades do português clássico, considerada "a língua das caravelas", por Galves (2007); Período II, cujos verbos lexicais e auxiliarem perdem movimento, deixando de subir até a periferia esquerda da sentença, mas subindo até uma posição de tempo; e o Período III, em que há adicional perda de movimento em relação aos verbos lexicais, mas não aos auxiliares. Esses resultados parecem validar a hipótese defendida por Ribeiro (1998) e Galves (2007) de que as origens do PB são anteriores ao século XIX e auxiliam a repensar a periodização da língua portuguesa no Brasil.

Palavras-chave: Periodização do português brasileiro. Ordem do verbo. Movimento do verbo.

Introduction

When compared to European Portuguese (EP), Brazilian Portuguese (BP) exhibits significant differences, leading some scholars to consider them distinct languages (Kato, 2012). However, at what point in history did BP begin to diverge from EP? If we examine the periodisation of the Portuguese language in Brazil, various proposals have emerged over time.

One of the first researchers to address the complexity of the Brazilian linguistic issue was Serafim da Silva Neto (1986 [1950]), who outlined the history of the Portuguese language in the Americas based on the colony's social history. He divided the history of Portuguese in Brazil into three phases: the first, from the beginning of colonisation (1532) to the expulsion of the Dutch (1654); the second, from 1654 to 1808, marked by the arrival of the Prince Regent and the Portuguese Court in Brazil; and the third phase, beginning in 1808, which saw the arrival of the Portuguese Court in Rio de Janeiro, profoundly transforming life in the colony.

Another periodisation proposal is that of Marlos de Barros Pessoa (2003). The primary distinction of this approach compared to the previous one is its attempt to integrate socio-historical milestones with stages of linguistic development. According to Pessoa (2003), between 1534 and 1750, multilingualism and the formation of rural linguistic varieties prevailed. This phase was followed by a stage of koineisation or pre-koineisation

of a common language (1750–1808), a period of urban variety formation (1808–1850), and a sub-period of stabilisation of urban varieties and the common language (1850–1922), which together correspond to the second phase of BP's gestation. Subsequently, in the third phase of BP's development, from 1922 onwards, there was a period of literary language elaboration.

Lobo (2003) also proposed a periodisation for the linguistic history of Brazil. Her approach is based on Brazilian demographic-linguistic history, population growth associated with urbanisation, and the process of schooling linked to linguistic standardisation. The selected criteria are thus sociolinguistic in nature. For Lobo (2003), the periodisation of Brazil's linguistic history is divided into two broad phases: the first, characterised by widespread multilingualism, marked by the absence of urbanisation, schooling, and linguistic standardisation; and the second, in which localised multilingualism prevails, characterised by schooling, urbanisation, and linguistic standardisation.

Grounded in linguistic criteria, whilst still considering the socio-historical evolution of speakers, Volker Noll (2008) critiques the previously presented proposals, which he argues rely almost exclusively on historical events. He defines his periodisation as follows: from 1500 to 1550, an initial phase in which Portuguese is transferred to Brazil; from 1550 to 1700, a formative phase in which the first characteristics of Portuguese in Brazil emerge; from 1700 to 1800, a differentiating phase, during which BP takes shape and the divergence between European and Brazilian varieties begins to appear; from 1800 to 1950, a phase of written language and education development, marked by the introduction of the printing press, the establishment of public education, and the creation of higher education institutions, furthering the differentiation from the European norm; and, from 1950 onwards, a levelling phase, influenced by the rise of mass media, the introduction of television, and increasing urbanisation.

Finally, Dante Lucchesi (2017) proposes a periodisation of BP based on Brazil's sociolinguistic history. His approach is rooted in both linguistic and social history, incorporating a broader range of languages that contributed to the formation of Brazilian society. The major innovation of this proposal is its temporal extension to a period prior to the arrival of the Portuguese in Brazil.

Thus, for Lucchesi (2017), the first phase of Brazil's sociolinguistic history, termed Tupinisation of the Coast, begins after the year 1000, with the expansion of Tupi along the coastline, and ends in 1532 with the effective beginning of Portuguese colonisation in Brazil. The second phase, Generalised Multilingualism, spans from 1532 to 1695

and is characterised by low population density and linguistic plurality. The third phase, Linguistic Homogenisation, begins in 1695 with the discovery of gold in Minas Gerais and extends till the Revolution of 1930, and is subdivided into three stages: (1) from 1695 to 1808, marked by Brazil's first urbanisation surge during the gold cycle and the expansion of cattle farming; (2) from 1808 to 1850, when urbanisation intensified with the arrival of the Portuguese Court and Brazil's political independence; (3) from 1850 to 1930, characterised by the abolition of the transatlantic slave trade, massive European and Asian immigration, and the reinforcement of purist, Lusitanised linguistic norms. The fourth phase, Linguistic Levelling, from 1930 to the present, is marked by increasing industrialisation and urbanisation, as well as the dissemination of the educated urban norm, erasing contact-induced linguistic features in popular varieties. The different proposals for BP's periodisation are summarised below.

Periodisation Proposals for Brazilian Portuguese				
Serafim da Silva Neto (1986)				
1st phase 1532-1654	2nd phase 1654-1808	3rd phase 1808-present		
Marlos de Barros Pessoa (2003)				
1st phase 1534-1750	2nd phase 1750-1922	3rd phase 1922-present		
Tânia Lobo (2003)				
1st phase 1500-1850	2ª phase 1850-present			
Volker Noll (2008)				
1st phase 1500-1550	2nd phase 1550-1700	3rd phase 1700-1800	4th phase 1800-1950	5th phase 1950-present
Dante Lucchesi (2017)				
1st phase 1000-1532	2nd phase 1532-1695	3rd phase 1695-1930	4th phase 1930-present	

Despite the various periodisation proposals considering aspects of the external and/or internal history of the language which attempt to explain the differences between BP and EP, none explicitly address syntactic aspects – precisely where BP and EP diverge most significantly. Another noteworthy aspect is that none of these proposals pinpoint the moment when BP emerged.

Conversely, studies such as those by Tarallo (1993), Ribeiro (1998), Galves (2007), and Corôa (2022), whilst not proposing a periodisation, seek to determine the point at which BP emerged. These studies examine syntactic features that define BP grammar, such as the reorganisation of the pronominal system – leading to the loss of clitics and the restructuring of possessive pronouns –, the change in VS order in declarative and interrogative sentences, the weakening of agreement systems, and the loss of null subjects.

Tarallo (1993) identified three major processes of change in BP compared to EP: (i) the loss of pronominal reference, (ii) modifications in pronominalisation strategies, and (iii) the reorganisation of basic sentence patterns. Regarding (i), BP lost pronominal reference, leading to a reorganisation of both full and clitic pronouns. As a consequence of (i), BP also changed its relativisation strategies, with pied-piping relatives – superficially similar to those in the standard norm – resumptive pronoun relatives – where the gap is filled by a coreferential pronoun – and the so called chopping relatives – where both the governing preposition and the relativised constituent are absent. The third change process (iii) is also linked to the first two. According to Tarallo (1993), BP underwent a significant shift in its pronominalisation strategies, resulting, among other things, in a rigidification of the canonical word order pattern towards SVO. These changes, according to Tarallo (1993), occurred at the end of the 19th century.

Ribeiro (1998), Galves (2007), and Corôa (2022), however, propose a temporal shift further back, considering aspects of the grammar of the first Portuguese settlers in Brazil. Analysing clitic placement, Ribeiro (1998) argues that BP may have undergone a shift at the turn of the 18th to the 19th century, but many of the phenomena analysed as resulting from 19th-century changes in BP actually originated in the 16th century. Therefore, Ribeiro (1998) suggests that Brazilian Portuguese consists of at least two distinct grammars: one from the 16th to the 18th centuries and another from the 19th to the 20th centuries. Galves (2007) revisits Ribeiro's (1998) question and, based on a systematic study of clitic placement in the history of Portuguese, argues that BP evolved

from the grammar of Middle Portuguese – an intermediate grammatical stage between Old Portuguese and Modern Portuguese.

Additionally, Corôa (2022) examined the V2 effect and its correlated properties – clitic placement, null subjects, and null objects – in data from the 17th and 18th centuries. The findings indicate that the grammar of Brazilians born in the 17th century was still closely aligned with Classical Portuguese, with significant changes such as the loss of the V2 effect, the generalisation of proclisis, and the occurrence of referential null objects in 18th-century data, suggesting that BP grammar emerged in the 18th century.

Furthermore, Cardoso, Andrade, and Carneiro (2023), based on clitic placement data, raise the hypothesis that, between Classical Portuguese and the establishment of Modern Brazilian Portuguese in its prestigious variant, there existed a linguistic variety acquired by white Brazilians, which the authors term *Colonial Brazilian Portuguese*. Although this variety remained similar to Classical Portuguese, it exhibited enough divergences to suggest the early development of a distinct vernacular grammar in Brazil.

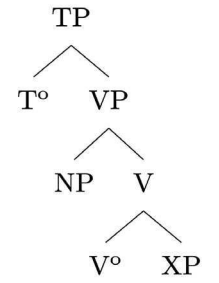
Building upon these studies, this paper seeks to address the following question: When did the Brazilian variety of Portuguese emerge? Based on intralinguistic syntactic evidence, a periodisation of BP is proposed. The phenomenon selected for analysis is verb placement throughout history, based on a dataset spanning from the 17th to the 21st centuries. According to Galves (2018), studying the evolution of verb position over time alongside the loss of the V2 effect in both BP and EP is crucial, as it can help elucidate the dynamics of change in these two systems. Therefore, verb order concerning the object – a feature of V2 grammar – is also considered, although the primary focus remains the positioning of the verb in relation to fixed adverbs. The hypothesis put forward aligns with those of Ribeiro (1998), Galves (2007), and Corôa (2022), advocating that the transition from the 17th to the 18th century marks the emergence of Brazilian Portuguese grammar.

To answer the question addressed in the present study, this paper is structured as follows. In Section 1, we discuss verb placement in relation to certain fixed elements within the sentence. Section 2 outlines the methodology employed to obtain the results. Section 3 presents the research findings, which serve as the basis for Section 4, where we propose a periodisation of Brazilian Portuguese based on verb placement patterns. Finally, Section 5 offers some concluding remarks.

1 Word order: the verb

Within the framework of Generative Grammar, in natural languages, a lexical verb universally concatenates with its arguments to form a Verbal Phrase (VP), as represented in (1). Despite this universal configuration, the position of the verb in relation to certain fixed elements within the sentence varies. Considering the examples in (2), it is possible to observe that verbs in Romance and Germanic languages behave differently. In English, the verb *sleep* appears to the right of the adverb *often*. In French, however, if the verb follows the adverb *souvent* ('frequently'), the sentence is ungrammatical, whereas when it precedes the adverb, the sentence is grammatical.

(1)



(2) a. English
John (*sleeps) often sleeps early.

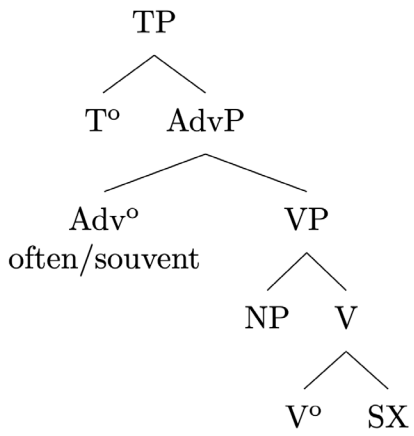
b. French

Jean	dort	souvent	*dort	tôt.
J.	sleep.PRS.3SG	often	sleepPRS.3SG	early

'J. often sleeps early.'

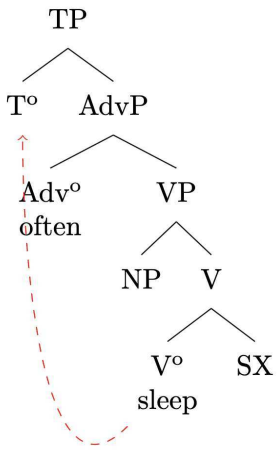
The contrast observed above, where the verb precedes the adverb in French but follows it in English – despite the verb having the same underlying structure in both languages – has been accounted for in Generative literature as a consequence of the presence or absence of a verb movement rule. Some scholars have proposed that elements such as adverbs and negation are structurally fixed within the sentence (cf. Edmonds, 1978; Pollock, 1989; Belletti, 1990; Cinque, 1999). As argued by Pollock (1989), adverbs such as *often* and *souvent* function as VP-adjuncts, meaning that they universally merge above VP, as illustrated in (3).

(3)

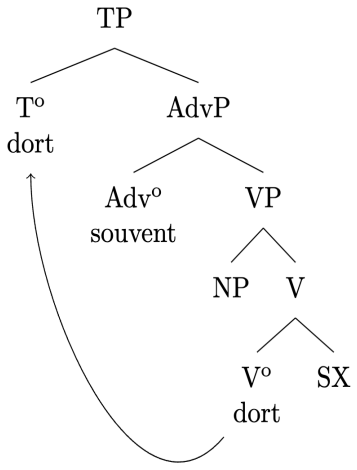


Consequently, the fact that the finite lexical verb in English obligatorily follows the adverb *often*, whilst in French the counterpart *souvent* precedes the verb, suggests that, in Germanic languages, the verb does not leave the VP through syntactic movement. Contrarily, in Romance languages, verb movement applies, yielding different word orders. Thus, the ADV_V order ($often_{ADV}$ $sleeps_{V}$) in English can be explained by the absence of movement to the inflectional domain (marked by the dotted line in (4a)), whereas the French V_ADV order ($dort_{V}$ $souvent_{ADV}$) results from the presence of verb movement, as illustrated in (4b).

(4) a. English



b. French



Building upon the proposal that certain elements in sentence structure are fixed, Cinque (1999) argues that there are approximately 40 adverbs that are rigidly ordered above VP, each semantically lexicalising the specifier of distinct functional projections.

Within this specialised framework – Syntactic Cartography, which systematically maps adverbs in syntactic structure – the hypothesis is that cross-linguistic variation in verb placement relative to adverbs results from differences in the landing site of verb movement.

(5) The universal hierarchy of functional projections of TP (Adaptaded from Cinque, 1999, p. 106; Ledgeway & Lombardi, 2005, p. 81; Cinque, 2006)
ALTOS: [frankly $Mood_{SpeechAct}$ > [luckily $Mood_{Evaluative}$ > [allegedly $Mood_{Evidential}$ > [probably $Mod_{Epistemic}$ > [once T_{Past} > [then T_{Future} > [perhaps $Mood_{Irrealis}$ > [necessarily ModNecessity > [possibly $Mod_{Possibility}$ > [usually $Asp_{Habitual}$ > [finally $Asp_{Delayed}$ > [tendentially $Asp_{Predispositional}$ > [again $Asp_{Repetitive(I)}$ > [often $Asp_{Frequentative(I)}$ > [willingly $Mod_{Volition}$ > [quickly $Asp_{Celerative(I)}$ >
BAIXOS: [already $T_{Anterior}$ > [no longer $Asp_{Terminative}$ > [still $Asp_{Continuative}$ > [always $Asp_{Perfect}$ > [just $Asp_{Retrospective}$ > [soon $Asp_{Proximative}$ > [briefly $Asp_{Durative}$ > [(?) $Asp_{Generic/Progressive}$ > [almost $Asp_{Prospective}$ > [suddenly $Asp_{Inceptive}$ > [obligatorily $Mod_{Obligation}$ > [in vain $Asp_{Frustrative}$ > [(?) $Asp_{Conative}$ > [completely $Asp_{SgCompletive(I)}$ > [everything $Asp_{PlCompletive}$ > [well Voice > [early $Asp_{Celerative(II)}$ > [? $Asp_{Inceptive(II)}$ > [again $Asp_{Repetitive(II)}$ > [often $Asp_{Frequentative(II)}$ > … *v*P

To pinpoint and compare the precise landing site of verbs across languages, the adverbial hierarchy has been methodologically divided into two domains (cf. Ledgeway & Lombardi, 2005, p. 81): adverbs in a higher position, lexicalising functional projections between $Asp_{Celerative(I)}$ and $Mood_{SpeechAct}$, and adverbs in a lower position, closer to the site where the verb is merged, lexicalising projections between $T_{Anterior}$ and $Asp_{Frequentative(II)}$.

Although the Cartographic approach has suggested an extensive array of functional projections within the architecture of grammar, this endeavour appears crucial for a more refined diagnosis of verb movement. As previously discussed, following Pollock (1989) and prior to the emergence of the cartographic model, it was generally assumed that the finite lexical verb in English remained within the VP, based on its placement relative to adverbs like *often* (cf. (2)). However, with the advent of the Cartographic model (Cinque, 1999), the English pattern has been more accurately understood. By adopting a more articulated hierarchy, such as the one proposed in (5), it becomes evident that the English verb, contrary to earlier assumptions, does undergo movement – albeit to a low syntactic

position (cf. Tescari Neto, 2012; Roberts, 2017), landing just above the functional projection of Voice, where the adverb *well* is lexicalised, as seen in (6). Thus, the hierarchy proposed by Cinque (1999), which we adopt in this study, appears to offer a more precise diagnostic for verb movement.

(6) John (*well) sings **well**.

In order to discuss periodisation issues concerning BP through the lens of verb movement in its diachrony, we adopt the hierarchy in its Portuguese version as proposed by Tescari Neto (2019). This framework enables us to determine the position of the verb in BP history, based on its ordering in relation to fixed and hierarchically structured adverbs.

(7) The universal hierarchy of functional projections of TP – Brazilian Portuguese version (Tescari Neto 2019: 3567)
ALTOS: [francamente $Mood_{SpeechAct}$ > [surpreendentemente $Mood_{Mirative}$ > [felizmente $Mood_{Evaluative}$ > [evidentemente $Mood_{Evidential}$ > [provavelmente $Mod_{Epistemic}$ > [uma vez T_{Past} > [então T_{Future} > [talvez $Mood_{Irrealis}$ > [necessariamente $Mod_{Necessity}$ > [possivelmente $Mod_{Possibility}$ > [normalmente $Asp_{Habitual}$ > [finalmente $Asp_{Delayed}$ > [tendencialmente $Asp_{Predispositional}$ > [novamente $Asp_{Repetitive(I)}$ > [frequentemente $Asp_{Frequentative(I)}$ > [de/com gosto $Mod_{Volition}$ > [rapidamente $Asp_{Celerative(I)}$ >
BAIXOS: [já $T_{Anterior}$ > [não…mais $Asp_{Terminative}$ > [ainda $Asp_{Continuative}$ > [sempre $Asp_{Perfect}$ > [apenas $Asp_{Retrospective}$ > [(dentro) em breve $Asp_{Proximative}$ > [brevemente $Asp_{Durative}$ > [(?) $Asp_{Generic/Progressive}$ > [quase $Asp_{Prospective}$ > [repentinamente $Asp_{Inceptive}$ > [obrigatoriamente $Mod_{Obligation}$ > [à toa $Asp_{Frustrative}$ > [(?) $Asp_{Conative}$ > [completamente $Asp_{SgCompletive(I)}$ > [tudo $Asp_{PlCompletive}$ > [bem Voice > [cedo $Asp_{Celerative(II)}$ > [do nada $Asp_{Inceptive(II)}$ > [de novo $Asp_{Repetitive(II)}$ > [frequentemente $Asp_{Frequentative(II)}$ > …

As a working hypothesis, the aforementioned hierarchy provides an accurate diagnostic for identifying the landing site of verb movement. Despite being divided into high and low adverbs, it appears that, in BP, only low adverbs should be considered reliable indicators of verb raising. Tescari Neto (2019, p. 3573) notes that high adverbs, such as *provavelmente* ('probably'), in Spec,$Mod_{Epistemic}$, cannot follow the verb when they occur

in a sentence-final position (cf. (8a)), albeit they may do so in the presence of an object Noun Phrase (NP), as in (8b). Furthermore, assuming that VP ellipsis is licensed by verb movement (Matos & Cyrino, 2001), Tescari Neto (2019) also demonstrates that a high adverb is not recoverable through VP ellipsis, supporting the hypothesis that (8b) should not be derived *via* verb movement over *provavelmente*. This corroborates the claim that "high adverbs cannot be used as diagnostics for verb movement in BP" (Tescari Neto, 2019, p. 3574).

(8) a. *João mente **provavelmente**.
J. lie.PRS.3SG probably
'J. probably lies.'
b. João comia **provavelmente** arroz.
J. eat.PRS.3SG probably rice
'J. probably ate rice.'

(9) João comia **provavelmente** arroz e a Maria também comia [-].
J. eat.PST.3SG probably rice and the M. too eat.PST.3SG
'J. probably are rice and M. also did [-].'
a. [-]: ~~comia~~ arroz
eat.PST.3SG rice
b. [-]: ***provavelmente** arroz
probably rice

To sharpen up the diagnostic of verb movement over time, it is necessary to discuss cases where a given word order does not necessarily provide conclusive evidence for the presence or absence of verb movement. As seen in (10), in BP, *ainda* ('still') may either precede or follow the finite verb. The placement of *ainda* before the verb can be explained by left-peripheral movement[2] (as suggested by Andriani, 2016, p. 240, for *ancora* in an Italian dialect), as illustrated in (10a), rather than by the absence of verb movement. Thus, regarding this adverb, solely the word order where the verb precedes *ainda*, as in (10b), should be taken as evidence of verb movement.

[2] This left-periphery displacement is supported by the precedence of the adverb over a quantifier subject in the low Foc position, such as *ninguém* 'nobody'.
(i) No começo, ainda *ninguém* tinha tentado elaborar uma versão falada para essa questão dos novos pronomes (https://www.extraclasse.org.br/geral/2021/10/a-gramatica-da-inclusao/)
(ii) Ainda *ninguém* está passando mal por enquanto, mas fizeram isso na maldade, só pode ser e quero que alguém se pronuncie a respeito (https://www.reclameaqui.com.br/burger-king/que-bebida-e-essa_EpeVUiu6bxaZ4qCR/)

(10) a. O Luciano **ainda** está se recuperando da cirurgia complexa...
theL. still be.PRS.3SGCLrecover.PROG of.the surgerycomplex
'L. is still recovering from the complex surgery...' (cf. also "O Luciano está **ainda** se recuperando…"[3])
b. O relator está **ainda** fazendo novos ajustes no texto…[4]
The rapporteur be.PRS.3SG still make.PROG new adjustments to.the text
'The rapporteur is still making new adjustments to the text.'

The adverb *sempre* ('always') presents additional complications, as it has different interpretations (Tescari Neto, 2013; Araújo-Adriano, 2018): an assertive reading (11a), a temporal/aspectual reading (11b-c), and a behavioural pattern reading (11d). Each interpretation is activated depending on the functional projection where *sempre* is merged: $Mood_{SpeechAct}$, $Asp_{Perfect}$, or $Asp_{Frequentative(II)}$, respectively. In this study, we control verb placement in relation to *sempre* in its temporal/aspectual reading, which is semantically associated with $Asp_{Perfect}$.

(11) a. Lopes (1998, p. 7)
Sempre quero ver se tens coragem para isso!
really want.PRS.1SG see.INF if have.PRS.2SG courage to this
'I do want to see if you are bold enough to do that'
b. Dane-se, eu vou **sempre** estar lá. Ninguém me deu esse lugar. Eu
Screw-CL, I will always be.INF there. No one CL give.PST.3SG this place I conquistei.
earn.PST.1SG
'Screw it, I will always be there. No one gave me this place. I earned it.'
c. **Sempre** dei o melhor de mim, **sempre** te tratei muito bem.
always give.PST.1SG the best of me always CL treat.PST.1SG very well
'I always gave my best, always treated you very well.'
d. Você vem **sempre** aqui?
you come.PRS.3SG often here?
'Do you come here often?'

[3] https://oglobo.globo.com/cultura/noticia/2022/06/luciano-szafir-e-internado-com-obstrucao-intestinal-apos-retirada-de-bolsa-de-colostomia.ghtml
[4] http://www.espn.com.br/noticia/232257_relator-volta-a-alterar-texto-da-lei-geral-da-copa-para-incluir-idosos-na-cota-social

The adverb *bem* ('well') can only appear post-verbally, as in (12a). Its pre-verbal placement triggers a specific focus-related interpretation (12b). As noted by Tescari Neto and Forero Pataquiva (2020, p. 495, footnote 4), the pre-verbal position of *bem* involves, after the obligatory verb movement over *bem*, additional movement of the adverb to Spec,Foc and the raising of the subject to Spec,Top.

(12) A Maria (*bem) fala **bem** inglês.
the M. well speak.PRS.3SG well English
'M. speaks English well.'

b. Tescari Neto e Forero Pataquiva (2020: 495, NR 4)
O João **bem** sabe o que deve fazer (não sabe mal!)
the J. well know.PRS.3SG the what must do.INF not know.PRS.3SG bad
'João sure knows what he must do (he doesn't know it poorly!).'

For the present study, we analyse verb placement in relation to a selection of low adverbs: *já* ('already'), *ainda* ('still'), *sempre* ('always'), *quase* ('almost'), and *bem* ('well'), which lexicalise, respectively, the projections of $T_{Anterior}$, $Asp_{Continuative}$, $Asp_{Perfect}$, $Asp_{Prospective}$ and Voice.

Another necessary enhancement in diagnosing verb movement must take into account evidence such as subject position and object fronting. With this in mind, we consider object fronting, given that Portuguese, in its earlier stages, exhibited a flexible V2 grammar, as reported in the literature (Ribeiro, 1995; Torres Morais, 1995; Paixão de Sousa, 2004; Gibrail, 2010; Antonelli, 2011; Galves & Paixão de Sousa, 2017; Galves & Gibrail, 2018; Galves, 2020). Thus, adverb movement alone would not sufficiently characterise the syntactic processes occurring in these systems.

V2 languages exhibit two main properties: verb movement to the C head and the movement of only one constituent to Spec,CP (Haegeman, 1996; Holmberg, 2015). Whilst Vikner (1995) was the first to show that in V2 systems the verb moves to the CP layer, the left periphery of the sentence, Rizzi (1997) provided extensive evidence that CP contains additional internal structures, splitting CP into a series of distinct functional projections (13):

(13) ForceP> TopP *> FocP> TopP *> FinP

TopP hosts topics, typically encoding given information that is in some way salient and available in the discourse. This position appears twice in the hierarchy, albeit it may theoretically occur an indefinite number of times. FocP accommodates focus, marking new information, and, in some languages, it is restricted to contrastive focus. FinP is a projection reflecting finiteness, given that many languages exhibit distinctions between finite and non-finite complementisers (Rizzi, 1997).

Considering the split-CP hypothesis and the hierarchies proposed by Rizzi (1997), some scholars (Roberts, 2004; Holmberg, 2015) argue that the V2 effect is linked to the feature of [+/- FINITE], as in these languages, FinP hosts the φ-probe and the edge feature responsible for V2.

Building on Benincà (1995), Wolfe (2018, 2019) expands this proposal, according to which all V2 systems involve verb movement to Fin; however, in a subclass of V2 languages, the Force projection exhibits the same properties as Fin. Prior to Wolfe, Pinto (2011) had already proposed an articulation between Force and Finiteness to account for variation of the V2 effect[5].

For Wolfe (2018, 2019), V2 languages can be categorised according to whether the landing site of the verb is ForceP or FinP. This results in the following typology of V2 languages based on landing sites:

(14) Typology of V2 languages (Wolfe, 2018)
a. V-to-Fin (languages with a less strict/more flexible V2 system)
b. V-to-Force (languages with a strict V2 system)

The intuition behind this proposal is that "the head bearing the features responsible for the V2 property is low in the functional structure, and thus permits lexicalisation of an array of functional projections structurally higher than the moved verb" (Wolfe, 2016, p. 296). Thus, V>2 orders are possible if the leftmost elements are generated *in situ*. In FinV2 systems, since nothing *a priori* excludes a successive set of projections hosting constituents simultaneously, the result of these operations would be a "flexible" V2 system in which both V1 and V>2 orders can occur alongside verb-second placement.

ForceV2 languages would exhibit an edge feature requiring verb movement from Fin to Force and the movement of an already fronted XP to Spec,ForceP. In these ForceV2 languages, V>2 orders are possible, though less common. This is so because the field above

[5] The structure proposed by Pinto (2011) differs from that of Wolfe (2015) regarding the positioning of the Frame head.

Force is associated with semantic-pragmatic features encoding the speaker's attitude.

Since verb movement constitutes only half of the V2 effect, whilst the other half involves the fronting of a constituent – regardless of its category or syntactic function – to the pre-verbal position, we consider the specificity of object fronting (Lightfoot, 1995; Holmberg, 2015). This serves as evidence for V2, as such constructions exhibit a highly specific OVS structural configuration whilst simultaneously functioning as a reliable diagnostic for verb movement. In the following section, we present the methodology employed to conduct the present study.

2 Methodology (data collection, searches and corpus)

The data for this research were extracted from 11 theater texts and 1 ballad – morphologically annotated – and from 346 Letters and 1330 Minutes of the Salvador City Council – morphologically and syntactically annotated, originating from the *Tycho Brahe-Brazil Corpus* (cf. Araújo-Adriano; Corôa, 2022).

Table 1 – Information on the texts of the analysed *corpus*

Birth	Author	Tycho Brahe Code	Publication of the texts[6]	Genre
1602	Rui de Carvalho Pinheiro	va_013	1650-1684	Letter
1630	Domingos Garcia de Aragão	va_014	1653	Letter
1650	Pedro Dias Pereira	va_017	1699-1710	Letter
1670	João de Couros Carneiro Filho	va_016	1699-1710	Letter
1670	Manuel Silveira de Magalhães	va_018	1710-1730	Letter
1672	Manuel Pessoa de Vasconcelos	va_017	1699-1710	Letter
1672	Manuel Pessoa de Vasconcelos	va_018	1710-1730	Letter
1700	João de Couros Carneiro Neto	va_018	1728-1730	Letter
1705	Antônio José da Silva	s_007	1737	Comedy
1710	Jerônimo Sodré Pereira	va_019	1751-1765	Minutes
1720	Joaquim Rodrigues da Silveira	va_019	1751-1765	Minutes
1720	Manuel José de Azevedo	va_020	1765-1775	Minutes
1721	José Álvaro Pereira Sodré	va_020	1765-1775	Minutes
1725	João Duarte Silva	va_020	1765-1775	Minutes
1740	Domingos Caldas Barbosa	b_011	1798	Ballad
1815	Martins Pena	p_003	1833	Comedy
1815	Martins Pena	p_004	1845	Comedy

[6] Although we use the title "publication of the texts", in the case of the minutes and letters of the Salvador City Council, we consider the period of production of the documents.

1829	José de Alencar	a_008	1857	Comedy
1838	França Junior	j_001	1883	Comedy
1855	Artur de Azevedo	a_009	1891	Comedy
1880	Gastão Tojeiro	t_001	1920	Comedy
1934	Gianfrancisco Guarnieri	g_011	1957	Comedy
1956	Miguel Falabella	f_003	1990	Comedy
1978	Paulo Gustavo	g_012	2006	Comedy
1966	Paulo Sacaldassy	s_006	2007	Comedy

The minutes and letters of the Salvador City Council (308,649 words) were written by 18 clerks throughout the 17th century and the first half of the 18th century. The plays were published between the 18th and 21st centuries, and the ballad in the 18th century (233,183 words). The total number of words analysed was 541,832. These texts are all available and freely accessible on the *Corpus website* (https://www.tycho.iel.unicamp.br/corpus/catalogo.html). Below, we present Table 2, with information by century, considering the distribution of the genres.

Table 2 – Distribution of the *corpus* by century and genre

Century	Genre	Number of words
17	Letters	90,676
18	Letters	108,607
18	Minutes	109,366
18	Theatrical comedy	27,224
18	Song	18,445
19	Theatrical comedy	86,842
20	Theatrical comedy	67,597
21	Theatrical comedy	41,623

As mentioned in the introduction, the phenomenon chosen for our analysis is verb order and object fronting between the 17th and 21st centuries. Considering that verb positioning in relation to adverbs is a reliable diagnostic for verb movement (cf. section 1), occurrences of sentences were computed in which the lexical verb and the auxiliary verb preceded or followed, immediately or not, certain adverbs diagnostic for verb movement, more specifically the low adverbs *já* ('already'), *ainda* ('still'), *sempre* ('always'), *quase* ('almost'), and *bem* ('well'). This is because adverbial expressions such as *hoje em dia* ('nowadays'), *todos os dias* ('every day'), and *nunca* ('never') do not provide very clear evidence

of where they are positioned in the hierarchy we assume here and, consequently, do not provide clear evidence with regard to verb movement in the history of Portuguese.

Thus, based on the order, the precedence of a verb in relation to the adverb – for example, *quase*, in Spec,$Asp_{Prospective}$ – would be equivalent to saying that, from its base position, that is, within *v*P, the verb moved (at least) to the following functional projection (for example, $Asp_{Generic/Progressive}$), so as to generate the order V_ *quase* . Conversely, if the verb is found to the right of the adverb, generating the order *quase* _V, by hypothesis, we are led to evidence a lower instance of the verb's movement.

Regarding the searches in the *corpus*, for the cases of adverb verb precedence, template V_ADV, we searched for the tag VB-X, where VB is equivalent to a lexical verb and X is any inflection, such as D and P – tags for past perfect and present, respectively (e.g., *comeu*/*comia* 'ate' = VB-D; *come* 'eats' = VB-P). For the template in question, this verb precedes – not necessarily an immediate precedence, hence the tag Precedes – a specific ADV – hence the tag iDoms –, for *bem* ('well'), for example.

(15) a. Search algorithm provided to locate V_ADV
((VB-D|VB-I|VB-P|VB-R|VB-RA|VB-SD|VB-SP|VB-SR|VB-F| Precedes ADV|ADV-*)
AND (ADV|ADV-* iDoms bem|já|ainda|sempre|quase))
b. Search algorithm provided to locate ADV_V
((ADV|ADV-* Precedes VB-D|VB-I|VB-P|VB-R|VB-RA|VB-SD| VB-SP|VB-SR|VB-F)
AND (ADV|ADV-* iDoms bem|já|ainda|sempre|quase))

For auxiliary verbs, we searched for the tags ET-X|SR-X|HV-X|TR-X, where ET is equivalent to the verb *estar* ('to be'), SR is equivalent to the verb *ser* ('to be'), HV to *haver* ('to have'), and TR to *ter* ('to have'), with X representing any inflection, such as P and D (e.g., *está* 'is' = ET-P; *tinha* 'had' = TR-D). In the ADV_AUX template, an ADV precedes – though not necessarily immediately, hence the *Precedes* tag – an auxiliary verb. The search for these tags retrieves any occurrences of *estar*, *ser*, *haver*, and *ter*, not necessarily in their auxiliary usage; for instance, in *João tinha muito dinheiro* ('João had a lot of money'), *tinha* is identified by the search as TR-D. Therefore, a manual filtering process was required to isolate only instances in which these verbs function as auxiliaries, which is confirmed when they precede a non-finite verb (gerund, participle, or infinitive[7]).

[7] For auxiliary criteria, cf. Lunguinho (2011), Araújo-Adriano (2018).

(16) a. Search algorithm provided to locate AUX_ADV
((ET-D|ET-I|ET-P|ET-R|ET-RA|ET-SD|ET-SP|ET-SR|ET-F|SR-D|SR-I|SR-P|SR-R|SR-RA|SR-SD|SR-SP|SR-SR|SR-F|HV-D|HV-I|HV-P|HV-R|HV-RA|HV-SD|HV-SP|HV-SR|HV-F|TR-D|TR-I|TR-P|TR-R|TR-RA|TR-SD|TR-SP|TR-SR|TR-F Precedes ADV|ADV-*) AND (ADV|ADV-* iDoms bem|já|ainda|sempre|quase))
b. Search algorithm provided to locate ADV_AUX
((ADV|ADV-* Precedes ET-D|ET-I|ET-P|ET-R|ET-RA|ET-SD|ET-SP|ET-SR|ET-F|SR-D|SR-I|SR-P|SR-R|SR-RA|SR-SD|SR-SP|SR-SR|SR-F|HV-D|HV-I|HV-P|HV-R|HV-RA|HV-SD|HV-SP|HV-SR|HV-F|TR-D|TR-I|TR-P|TR-R|TR-RA|TR-SD|TR-SP|TR-SR|TR-F)AND (ADV|ADV-* iDoms bem|já|ainda|sempre|quase))

As highlighted in the previous section, if we assume that Portuguese has a flexible V2 grammar, we must consider other evidence besides the positioning of the verb. The question of where the verb landed, that is, knowing whether the verb moved to a certain position, can be measured, among other things, from the position of the adverbs and the object fronting that, consequently, lead to the inversion of the subject and the raising of the verb. In this article, we will also consider object fronting as evidence of a V2 system and, consequently, of the loss of this system. In the next section, we will present these results. The examples reported in the results section will include the text code in the *Tycho Brahe Platform.*

3 Results

In this section, we present the results regarding the position of the verb in relation to those adverbs and object fronting that can shed light on verb movement in the history of Brazilian Portuguese. In general, we found few instances in which the adverb modifies the verb, as well as few cases of object fronting. Some of these were useful for diagnosing verb movement, while others were not. To distinguish lexical verbs from auxiliary verbs, we use *V* for the former and *AUX* for the latter. The data are grouped as follows: the 17th century; the 18th and 19th centuries; and the 20th and 21st centuries, following the hypothesis proposed.

3.1 The 17th century

Examining the position of the verb in relation to the adverb *já* ('already'), we observed that, in most cases, this adverb followed the verbal element, regardless of whether it was a lexical or an auxiliary verb. In the case of lexical verbs, 82% of occurrences followed the V_*já* pattern (17a), whereas only 18% followed the *já*_V pattern (17b). For auxiliary verbs, postposing the adverb was a categorical position (17c). Below, we present some examples[8].

[8] Following a reviewer's recommendation, we have reduced the number of examples.

(17) Adverb *já*

a. e mais quando corre **já** por nossa conta o sal (va_017).
and more when run.PRS.3SG already for our account the salt
'and more when the salt is already running on us.'

b. O alívio que temos em ser governados por VM
the relief who have.PRS.1PL in be.INF govern.PTC by VM
cujas honras grandezas **já** gozamos (va_013).
whose honours great already enjoy.PRS.1PL
'The relief we have in being governed by VM whose great honours we already enjoy.'

c. e por isso se tem movido **já** muitas
and therefore CL have.PRS.3SG move.PTC already many
demandas (va_016)[9].
lawsuits
'and therefore many lawsuits have already been.'

Regarding the adverb *ainda* ('still'), we also observed a preference for a position after the verb. Despite the fact that the precedence of *ainda*, found in the data, cannot be taken as evidence of the absence of movement of the verbal item, considering that *ainda* can move to the left periphery of the sentence (cf. section 1), the verbal item preceding the adverb is evidence of verbal movement. This is what we observed with the lexical verbs preceding the adverb *ainda* in 82% of the occurrences (18a). In the case of auxiliary verbs, the postposition of the adverb was found in 60% of the occurrences (18b).

(18) Adverb *ainda*

a. e exerce **ainda** com toda a verdade (va_016).
and exercise.PRS.3SG still with all the truth
'he still exercises with all truth'

b. Esta disposição se foi observando **ainda** ao depois (va_017).
this provision CL be.PST.3SG observe.PROG still to the after
'This provision was still observed after'

Observing the position of the verb in relation to the adverb *sempre* ('always'), we were able to see that most of the data with that adverb followed the verbal item, regardless of the type of verb, whether lexical or auxiliary. In the first case, 68% of the occurrences were with V_ *sempre* (19a) and only 32% with *sempre* _V (19b). In the case of auxiliary

[9] The fact that not only the auxiliary verb but also its non-finite verbal complement precedes the adverb is *already* evidence that these two verbs moved together to a higher position in the sentence. In this regard, cf. Araújo-Adriano and Cyrino (2022, forthcoming).

verbs, the postposition of the adverb was also a preferred position, with 80% of the cases (19c), against 20% of the ordering *sempre* _AUX (19d). Below, we provide some examples.

(19) Adverb *sempre*

a. Ficaremos **sempre** reconhecidos (va_017).
stay.PRS.1PL always grateful
'We will always be grateful'

b. E **sempre** se praticou (va_016).
and always CL practice.PST.3SG
'and it has always been practiced'

c. E foram **sempre** aprovadas pelos sindicantes (va_013).
and be.PST.3PL always aproved for union members
'they were always approved by the union members'

d. e com este privilégio **sempre** se há de
and with this privilege always CL have.PRS.3SG of
contender (va_016)[10]
contend.INF
'and with this privilege one will always contend'

Regarding *quase* ('almost'), we observed that the positioning of the adverb before the auxiliary verb was categorical – we did not find data with this adverb and lexical verbs. Below, we present some of the few examples we found.

(20) Adverb *quase*

[...] a Fortaleza de Santo Antonio do Carmo que havia
[...]the Fortress of S. A. of.the C. who have.PST.3SG
principiado seu antecessor está **quase** posta em sua
start.PST.3SG your predecessor be.PRS.3SG almost placed in your
última perfeição, (va_16).
final perfection
'the Fortress of Santo Antonio do Carmo, which had been started by its predecessor, is almost placed in its final perfection'

[10] We treat the verb *haver* in *haver de* as an auxiliary, as demonstrated by Araújo-Adriano (2019, p. 75-86).

Regarding the adverb *bem* ('well'), we found lexical and auxiliary verbs both preceding and following it. All the cases of precedence were cases in which *bem* conveyed a focus reading and were therefore not evidence for the absence of verb movement (cf. also section 1). In this type of sentence, the adverb conveys an aspectual reading. According to Martins (1994), monosyllabic adverbs, when in a pre-verbal position, lose their original denotation, which he calls the "semantic emptying" of these elements.

On the other hand, we found evidence of lexical verbs preceding *bem,* which leads us to affirm that there was verbal movement over this adverb. Although we did not find any occurrence of *bem* following or preceding auxiliary verbs, the fact that the auxiliary verb precedes higher adverbs than *bem*, for example, *já* (17), *ainda* (18), *sempre* (19) and *quase* (20), by transitivity (Cinque, 1999, p. 33), it can be said that the functional verb also moved over *bem*.

(21) Adverb *bem*

De	que	até	aqui	temos	dito	se	mostra	**bem**	que
From	that	to	here	have.PRS.1PL	say.PTC	CL	show.PRS.3SG	well	that
uma das		grandes	obrigações	é		conservar-nos	tantos	anos	
one of the		great	obligations	be.PRS.3SG		keep.INF=CL	many	years	
o	dito	Governador (va_017).							
the	said	Governor							

'From what we have said so far it is clear that one of the great obligations is to keep the said Governor for so many years.'

3.1.1 Evidence of a V2 System

In V2 languages, object fronting is an option that changes the position in which the subject occurs; this does not happen in non-V2 languages. When the object is moved to the pre-verbal position in V2 languages, the order generated is OVS (Antonelli, 2001).

In the 17th century data, we found the following cases of direct object fronting (in italics, we indicate the phrase that performs the function of subject; and we underlined the object function):

(22) a. <u>Algumas quantias</u> <u>que</u> <u>achamos</u> **receberam** *os* *ditos*
Some amounts that think.PST.1PL receive.PST.3PL the so-called
tesoureiros (va_017).
treasurers
'The so-called treasurers received some amounts we think'

b. Esta mercê e singular esmola que esperamos
this mercy and singular alms that hope.PRS.1PL
nos **conceda** *a grandeza, e justiça de VM* (va_015).
CL grant.PRS.SBJ.3SG the greatness, and justice of VM
'This mercy and singular alms that we hope will be granted to us by the greatness and justice of Your Majesty'

In our data, in sentences with object fronting and an expressed subject, categorically, the subject occurs in the post-verbal position.

As it is a marked option, the OVS order is quite rare. In the *corpus* analysed, there are only 43 occurrences of this order in the 17th century, out of a total of 1551 cases of V2 sentences, which adds up to about 2.7% of the sentences.

3.1.2 Summary of 17th century data

To conclude the presentation of the 17th century data, in Table 3, we summarize the occurrences of lexical and auxiliary verbs in relation to the adverbs.

Table 3 – Percentage of lexical and auxiliary verbs in relation to adverbs in the 17th century

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	18%	18%	32%		
V_adv	82%	82%	68%		100%
adv_AUX		40%	20%		
AUX_adv	100%	60%	80%	100%	

Based on the percentages presented, we can observe that the lexical verbs moved through *já*, *ainda*, *sempre* and *bem*. Still, the auxiliary verbs moved through *já*, *ainda*, *sempre*, *quase* and, by transitivity, *bem*.

3.2 18th and 19th centuries

Regarding the positioning of *já* in the 18th and 19th centuries, we observe a preference for its precedence over lexical verbs, occurring in 86% of the occurrences in the 18th century (23b). It is noteworthy that the 14% of V_*já* (23a) can be treated as remnants of a V2 grammar, still present in the 18th century, but absent in the 19th century, considering the 100% (23c) of occurrences of lexical verbs following *já*.

(23) Adverb *já*

a. Quase dentro da mesma cidade tem Dona
Almost within by same city have.PRS.3SG D.
Izabel Maria Guedes de Britto uma rocinha
I. M. G. of B. a farm
que poderá ter em todo o seu circuito
that can.FUT.3SG have.INF in all the this circuit
quatrocentas braças de terra as quais possuia **já**
four hundred fathoms of land the which have.PST.3SG already
medidas e demarcadas (va_020).
measure.PTC and demarc.PTC
'Almost within the same city, Dona Izabel Maria Guedes de Britto has a small farm that may have four hundred fathoms of land in its entire circuit, which had already been measured and demarcated'

b. Cala-te Sancho, cala-te, que **já** lá vai (s_004)
Shut.IMP=CL S., shut.IMP=CL that already there go.PRS.3SG
'Shut up Sancho, shut up, it's already gone'

c. Nós **já** sabemos como foi o engano (p_004)
We already know.PRS.1PL how be.PST.3SG the mistake
'We already know how the mistake was made.'

d. Só por eu cumprimentar o moço **já** o estava
just by I great.INF the boy already CL be.PST.3SG
namorando. (b_011)
date.PROG
'Just by greeting the guy I was already dating him.'

e. Não poder deferirmos por estar o negócio
Not can.INF defer.INF.1PL by be.INF the matter
pendente como este esteja **já** decidido (va_019).
pending as this be.PRS.SBJ.3SG already decide.PTC
'We cannot defer because the matter is pending as it has **already been** decided'

f. **Já** está deitando o chapéu. (a_008).
Already be.PRS.3SG put-down.PROG the hat
'He is already putting his hat down'

Unlike lexical verbs, which in the 18th century showed a preference for following the adverb *já,* there was still a certain preference in the 18th century for auxiliary verbs to precede *já.* This occurred in the *corpus* in 73% of cases (23e). It was in the 19th century that the precedence of *já* over auxiliary verbs also became categorical, occurring in 100% of the data in the second century (23f).

The data from the 18th century can be taken as evidence that there was a process of competition between grammars in that period, with remnants of a classical grammar that was declining and a new grammar that was emerging, since, in BP, finite verbs consistently appear to the right of the adverb *já* (Modesto, 2001) and a sentence in which the order *já* _V occurs is only possible in BP with a focus reading, unlike what can be seen in classical Portuguese.

Further evidence that there are remnants of a V2 grammar in the 18th century is the occurrence of object fronting in the order O_V_S (24) (in 0.19% of the sentences analysed). In the 19th century, we did not find data on object fronting.

(24) a. E o mesmo **disse** *Mateus Pereira dos Santos Cavalcante* (va_019).
and the same say.PST.3SG M. P. of-the S. C.
'And the same was said by Mateus Pereira dos Santos Cavalcante.'

b. E este termo **registara** *o escrivão da nova contribuição*
and this term register.PTP.3SG the clerk of-the newcontribution
voluntária nos seus livros (va_020)
voluntary in-theyour books
'And this term was recorded by the clerk of the new voluntary contribution in his books'

In the diachronic research, we found *ainda* following the lexical verb both in the 18th century (25a), in 13% of the cases, and in the 19th century (25b), in 19% of the data, which suggests that the lexical verb moved to a position above the functional projection lexicalized by this adverb. Although we found postposition of auxiliary and lexical verbs in relation to *ainda,* showing us nothing in relation to the verb movement, both in the 18th century (25a) and in the 19th (25c) we have evidence of auxiliary verbs preceding this adverb, which constitutes evidence of movement of the auxiliaries over *ainda*.

(25) **Adverb *ainda***

a. O requerimento sobre se mandar bater moeda provincial se não
the request about if send.INF mint.INF currency provincial if not
tem **ainda** proposto no conselho por razão de se
have.PRS.3SG still propose.PTC in-the council by reason of CL
ter ocupado com Rio e Minas. (va_019)
have.INF busy with R. and M.
'The request regarding the minting of provincial currency has not yet been submitted to the council due to its engagement with Rio and Minas.'

b. Não observei **ainda**! (a_008)
not observe.PST.3SG yet
'I haven't observed it yet.'

c. Tenho **ainda** que ir arranjar mandado de despejo contra
have.PRS.1SG still that go.INF obtain.INF order of eviction against
uma viúva, minha inquilina, que há três não me paga o
a widow my tenant that have.PST.3SG three not CL pay.PRS.3SG the
aluguel da casa. (a_009)
rent of-the house
'I still have to obtain an eviction order against a widow, my tenant, who has not paid the house rent for three months.'

Regarding the adverb *sempre,* there was a certain change in the behaviour of verbs throughout the 18th and 19th centuries. During the 18th century, lexical verbs did not show a high tendency to precede the adverb, preceding it in 35% (26a), with this position being highly preferable in the 19th century (26c), appearing in 82% of the data. Regarding auxiliary verbs, in the 18th century[11], the postposition of the adverb to the auxiliary occurred in all appearances of the adverb *sempre* (26d), as well as in the 19th century (26e).

(26) **Adverb *sempre***

a. E que aos mesmos criadores se lhe dariam **sempre** que entrassem
and that to-the same creators CL CL give.PST.3PL always that enter.PST.SBJ.3PL
com os seus gados. (va_018)
with the their cattle
'And that the same breeders would be given [them] whenever they brought in their cattle.'

[11] In both the 18th and 19th centuries, there were cases in which the adverb *always* preceded the auxiliary verb. These were cases of *always* conveying a confirmatory reading, which were not computed (cf. Lopes, 1998; Tescari Neto, 2013; Araújo-Adriano, 2022).

b. **Sempre** se dará ao Senhor de Engenho, e Lavradores destes
always CL give.FUT.3SG to-the sir of mill and farmer of-this
livremente todo o gado (va_019)
freely whole the cattle
'All the cattle will always be freely given to the Lord of the Sugar Mill and the Farmers.'

c. O senhor nunca nos incomoda, dá-nos **sempre** muito prazer. (j_001)
the sir never CL bother.PRS.3SG give.PRS.3SG=CL always great pleasure
'The gentleman never bothers us; he always brings us great pleasure.'

d. Pois seguro-te, meu bem, que apesar de tudo hei
because assure.PRS.1SG=CL my dear that despite of everything have.PRS.1SG
de ser **sempre** firme, constante, e leal. (s_007)
of be.INF always firm constant and loyal
'Well, I assure you, my dear, that despite everything, I shall always remain firm, constant, and loyal.'

e. Se eu soubesse que havia de ser **sempre** tão feliz… (p_004)
If I know.PST.SBJ.1SG that have.PST.3SG of be.INF always so happy
'If I knew I would always be this happy, I would marry fifty times.'

We did not find occurrences of *quase with lexical verbs* in the data, only with auxiliaries. Its positioning in relation to the adverb was categorical: anteposition, both in the 18th and 19th centuries, as seen in ((27a) and (27b)), below. Although we have no evidence of *quase* with lexical verbs, in that period, the lexical verb preceded adverbs higher than *quase,* such as *sempre* (26), *ainda* (25), which, by transitivity, suggests that it also preceded *quase.*

(27) Adverb *quase*

a. … e está **quase** indo-se por um fio. (s_007)
and be.PRS.3SG almost go.PROG=CL by one thread
'and is almost hanging by a thread.'

b. Os recursos que eu possuía estão **quase** inteiramente
the resources that I possess.PST.3SG be.PRS.3PL almost entirely
esgotados. (a_009)
exhausted
'The resources I possessed are almost entirely exhausted.'

As for the positioning of the verb in relation to the adverb *bem,* both postposition and anteposition were attested. Just as in the 17th century, the antepositions to the verb were cases of *bem* with a focus reading. On the other hand, we found evidence of lexical verbs preceding *bem,* which leads us to affirm that there was verbal movement over this adverb in the 18th (28a) and 19th (28b) centuries.

(28) Adverb *bem*

a. Dizem **bem**, que o mundo não é capaz de… (s_004)
say.PRS.3PL well that the world not be.PRS.3SG capable of
'They say rightly that the world is not capable of'

b. VE passa **bem**. (j_001)
H.E. spend.PRS.3SG well
'Your Excellency is well.'

Although no occurrence of *bem* following or preceding auxiliary verbs has been attested, its precedence in relation to adverbs higher than *bem* , such as *sempre* (26), *ainda* (25), due to transitivity, allows us to affirm that, hypothetically, the functional verb also moved over *bem* in both the 18th and 19th centuries.

3.2.1 Summary of 18th and 19th centuries data

To conclude the presentation of the data from the 18th and 19th centuries, in Tables 4 and 5, we summarize the occurrences of the lexical and auxiliary verbs in relation to the adverbs. We can observe that lexical verbs move over the adverbs *bem, sempre, ainda* and, by extension, *quase,* an adverb lower than the last two, but does not move over *já.* On the other hand, auxiliary verbs only move over adverbs higher than *bem,* such as *quase, sempre, ainda* and *já*, in the 18th century, but no longer in the 19th century.

Table 4 – Percentage of lexical and auxiliary verbs in relation to adverbs in the 18th century

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	86%	87%	65%		
V_adv	14%	13%	35%		100%
adv_AUX	27%	50%			
AUX_adv	73%	50%	100%	100%	

Table 5 – Percentage of lexical and auxiliary verbs in relation to adverbs in the 19th century

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	100%	81%	18%		
V_adv		19%	82%		100%
adv_AUX	100%				
AUX_adv		100%	100%	100%	

3.3 20th and 21st Centuries

In the *corpus,* we found only one occurrence of *já* following a lexical verb, as seen in (29a). This use, however, is not prototypical of the adverb in question. Here, *já* seems to convey an idea of *soon, in a while.* In all other uses of the adverb, in its characteristic reading of preceding the event, the lexical and auxiliary verb preceded it in the 20th and 21st centuries (29b and 29c).

(29) **Adverb *já***

a. A Marcelina ainda não foi, mas vai **já**. (t_001)
the M. yet not go.PST.3SG but goPRS.3SG right away
'Marcelina hasn't gone yet, but she's going right away.'

b. Todo mundo **já** conhece essa história, Selma … (f_003)
everybody already know.PRS.3SG this story Selma
'Everybody already know this story, Selma.'

c. Você **já** está trocando as pernas. (f_003)
you already be.PRS.3SG change.PROG the legs
'You're already staggering.'

In the diachronic research, we only found *ainda* following auxiliary verbs (30a), both in the 20th and 21st centuries, which suggests that in this period the auxiliary verb moves to a position above the functional projection lexicalized by *ainda.* Occurrences of *ainda* preceding lexical and auxiliary verbs were also found, although this is no as evidence against the presence of verbal movement. However, since we found postposition of *ainda* only in relation to auxiliaries, it is possible to affirm that the absence of occurrence of postposition of the adverb in relation to lexical verbs is evidence of the absence of this configuration in the period analysed.

(30) Adverb *ainda*

Eles quase se mataram hoje de manhã, eu vi que ia
They almost CL kill.PST.3PL today of morning I see.PST.1SG that go.PST.3SG

ainda sobrando para mim e pulei fora (f_003).
still remain.PROG to CL and jump.PST.1SG out

'They almost killed each other this morning. I saw that I might get caught up in it too, so I got out.'

In relation to *quase,* postposition was only attested in relation to auxiliary verbs (31a), but not with lexical verbs, with which the adverb appeared postposed in all cases (31b).

(31) **Adverb *quase***

a. Seu Otávio está **quase** brigando no botequim! (g_011)
S. O. be.PRS.3SG almost fight.PROG in.the bar

'Mr. Otávio is almost getting into a fight at the bar.'

b. Reclamações do prédio todo – a síndica **quase**
Complaints from.the building whole the manager almost
deu na minha cara. (f_003)
hit.PST.3SG in.the my face

'Complaints from the whole building—the manager almost hit me in the face.'

Unlike the 18th and 19th centuries, the preference for the lexical verb to follow the adverb *sempre* decreased in the 20th and 21st centuries: in this period, the order *sempre* _V occurred in 80% of the occurrences (32a) and in only 20% with V_ *sempre* (32b). Unlike lexical verbs, auxiliary verbs categorically preceded the adverb *sempre* in the *corpus* (32c).

(32) Adverb *sempre*

a. Eu não sou mais criança, não, Maria Lúcia, e vou
I not be.PST.1SG more child not M. L. and go.PRS.1SG
logo avisando que eu não vou admitir que vocês me
soon notice.PROG that I not go.PRS.1SG admit.INF that you CL
tratem como uma débil-mental como **sempre** fizeram. (f_003)
treat.PRS.SBJ.3PL as a fool as always have.PST.3PL

'I'm not a child anymore, Maria Lúcia, and I'm telling you right now that I won't allow you all to treat me like a fool as you always have.'

b. Você tem **sempre** uma desculpa pronta. (t_001)
you have.PRS.3SG always an excuse ready
'You always have an excuse ready.'

c. É por isso que a senhora está **sempre** cansada, vive
be.PRS.3SG by this that the madame be.PRS.3SG always tired keep.PRS.3SG
me prometendo pancada! (g_011)
CL promise.PROG blow
'That's why you're always tired—you keep promising to hit me!'

In the 20th and 21st century data, the adverb *bem* categorically followed auxiliary and lexical verbs, as seen in examples (33a) and (33b) below.

(33) Adverb *bem*

a. Está com os pulmão arrebentando mas bebe **bem**. (g_011)
be.PRS.3SG with the lung wreck.PROG but drinkPRS.3SG well
'He's got his lungs wrecked, but he drinks just fine.'

b. E como é que a gente vai enxergar **bem** com aqueles lençóis
and how be.PRS.3SG that the people go.PRS.3SG see.INF well with those sheets
horríveis que tem que usar quando vira fantasma? (g_012)
horribles that have.PRS.3SG that use.PRS.3SG when turn.PRS.3SG ghost
'And how are we supposed to see well with those horrible sheets we have to wear when we turn into ghosts?'

3.3.1 Summary of 20th and 21st centuries data

To conclude the presentation of the data from the 20th and 21st centuries, in Tables 6 and 7, we summarize the occurrences of lexical and auxiliary verbs in relation to the adverbs observed so far. We can observe that lexical verbs move only over *bem,* but not over *quase, sempre, ainda* or *já.* Differently, auxiliary verbs, to precede *bem, quase, sempre* and *ainda,* need to move one position higher than *ainda,* but not more, considering that they cannot precede *já.*

Table 6 – Percentage of lexical and auxiliary verbs in relation to adverbs in the 20th century

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	86%	87%	77%		
V_adv	14%	13%	23%		100%
adv_AUX	27%	47%			
AUX_adv	73%	53%	100%	100%	

Table 7 – Percentage of lexical and auxiliary verbs in relation to adverbs in the 21st century

	já	*ainda*	*sempre*	*quase*	*bem*
adv_V	100%	100%	86%		
V_adv			14%		100%
adv_AUX	100%	64%			
AUX_adv		36%	100%	100%	

3.3.2 Section summary

In this section 3, we have shown the results regarding the positioning of verbs in relation to the adverbs *bem, quase, sempre, ainda* and *já* in the 17th, 18th, 19th, 20th and 21st centuries. With this diachronic data, we observed that the order of the verb has changed over time. At first, we saw in 3.1 that both lexical verbs and auxiliary verbs could move over *já* and all other lower adverbs. This ordering, as explained in 3.2, was no longer preferred for lexical verbs from the 18th century onwards, although there were still some remnants of the grammar from the previous century. This grammar seemed to resist in the contexts of auxiliary verbs that still preferentially precede the adverb *já* in the 18th century, as in the 17th century. On the other hand, there are no more remnants of a 17th-century grammar in the 19th century: both lexical verbs and auxiliary verbs necessarily follow *já,* but precede all lower adverbs: *ainda, sempre, quase* and *bem.*

In the 20th and 21st centuries, we observed, in 3.3, once again a change in the ordering of the verb: in this period, lexical verbs no longer precede *ainda, quase* and *sempre,* but continue to precede *bem,* unlike auxiliary verbs that continue to precede all these adverbs, except *já*, just as in the grammar of the 18th and 19th centuries. In Table 8, below, we summarize the results regarding the ordering of the Portuguese verb in relation to adverbs over time.

Table 8 - positioning of lexical and auxiliary verbs in relation to adverbs in BP diachrony

Century	Verb	Position	
		Precedes	**Does not precede**
17	lexical	*já, bem, quase, sempre, ainda*	-
	auxiliary	*já, bem, quase, sempre, ainda*	-
18	lexical	*bem, quase, sempre, ainda*	*já*
	auxiliary	*já, bem, quase, sempre, ainda*	
19	lexical	*bem, quase, sempre, ainda*	*já*
	auxiliary	precede *bem, quase, sempre, ainda*	*já*
20	lexical	*bem*	*já, ainda, quase, sempre*
	auxiliary	*bem, quase, sempre, ainda*	*já*
21	lexical	*bem*	*já, ainda, quase, sempre*
	auxiliary	*bem, quase, sempre, ainda*	*já*

4 Discussion

Based on the ordering of the verb in relation to fixed adverbs, we demonstrated in the previous section that Brazilian Portuguese has undergone three periods: **Period 1**, covering the 17th century, in which the grammar generated a flexible V2 system; **Period 2**, spanning the 18th and 19th centuries, in which lexical verbs and auxiliaries no longer preceded *já*; and finally, **Period 3**, covering the 20th and 21st centuries, where verb placement differs according to verb type – lexical verbs do not precede *já, ainda, quase,* and *sempre* but do precede *bem*, unlike auxiliary verbs, which precede all these adverbs except *já*.

Assuming the universal hierarchy of functional projections and following Wolfe's proposal (see Section 1), we observe that in Period 1, both lexical and auxiliary verbs moved to Fin (the left periphery of the sentence), a position that results in their precedence over all lower adverbs including *já*. In Period 2, this movement is no longer attested, with both auxiliary and lexical verbs moving instead to $T_{Anterior}$, a position where they follow *já* but precede all lower adverbs. Finally, in Period 3, auxiliary verbs continue to move to $T_{Anterior}$, whereas lexical verbs move to a position lower than the projection where *ainda* is located in $Asp_{Continuative}$, *sempre* in $Asp_{Perfect}$, *quase* in $Asp_{Prospective}$, but higher than *bem* in Voice.

As we have seen, written Portuguese produced by Brazilian authors lost the properties of a flexible V2 language in the 18th century. Based on Chomsky (2008), TP/IP manifests φ-features and Tense features only if selected by C, since T inherits its Tense features from C, which is a phase head (Roberts, 2019). Following Ouali (2008), Biberauer and Roberts (2017) propose that C may either transfer, share, or retain Tense features to account for parametric variation across languages. Accordingly, languages differ depending on whether (i) C transfers Tense features to T, (ii) C shares Tense features with T, or (iii) C retains Tense features. Building on Wolfe's (2018, 2019) proposal that in V2 languages the verb's landing site can be either Force or Fin, we assume that Fin, the lowest functional projection within the C domain (see Section 2), may transfer, share, or retain Tense features.

In V2 languages (except Icelandic, where Tense features are shared with C), Fin retains the Tense feature, meaning that T does not inherit it. Consequently, the verb is forced to move to the left periphery, and as a result, the edge feature attracts an XP to its specifier. Therefore, whether or not Tense features of Fin are inherited plays a crucial role in distinguishing Romance from Germanic languages.

In addition to this, we draw on the discussion raised by Schifano (2018), who, based on the articulated structure of functional projections in (7), proposes that Cinque's (1999) hierarchy can be reduced to three major functional fields: "a high one, where M(ood) is encoded, a clause-medial one, where T(ense) is encoded, and a low one, where A(spect) is encoded" (Schifano, 2018, pp. 134–135). Thus, verb movement is understood as targeting fields within the syntactic structure – e.g., Aspect – rather than specific positions, such as $Asp_{Frequentative}$. For our discussion, instead of an articulated structure of functional heads, it is sufficient to assume that IP/TP is structured into Aspectual, Temporal, and Mood fields. In our proposal, in Period 1, the verb moved to Fin; in Period 2, to the Temporal field; and in Period 3, only auxiliary verbs continue moving to the T field, whilst lexical verbs reduce their movement to the Aspect field.

Following Schifano (2018), we adopt the proposal that the property triggering movement to these fields is the presence of an uninterpretable feature [*u*F] in functional projections such as Fin, T, A, or M. Assuming that *v* is always interpretable [iX] for Fin, T, Asp, and M – hence, [*i*Fin], [*i*T], [*i*Asp], and [*i*M] – Fin, T, Asp, and M may bear either interpretable [*i*Fin], [*i*T], [*i*Asp], and [*i*M] or uninterpretable [*u*Fin], [*u*T], [*u*Asp], and [*u*M]features, respectively (note that in this system, features are always valued but may or may not be interpretable). Since uninterpretable features must be deleted as soon as

possible to be legible at the interfaces (Chomsky, 1995, pp. 219–348), the deletion of [*u*Fin], [*u*T], [*u*Asp], and [*u*M] carried by Fin, T, Asp, or M takes place *via* a checking relation with the goal *v* through movement. Consequently, when probes bear interpretable features matching those of *v*, this goal does not move; movement is only triggered when Fin, T, Asp, or M bear an uninterpretable feature, as outlined below:

(34) Features checking amongst *v* and Fin, T, A, and M: syntactic corollary
[Fin/M/T/A$_{[u\mathrm{X}]}$ [… $v_{[i\mathrm{X}]}$]]: movement
[Fin/M/T/A$_{[i\mathrm{X}]}$ [… $v_{[i\mathrm{X}]}$]]: absence of movement

Within this framework, a possible explanatory proposal for the different verb orders in relation to adverbs and, consequently, for the various positions to which the verb moves in the diachrony of Brazilian Portuguese (BP) can be outlined as follows.

During the 17th century, when both auxiliary and lexical verbs preceded *já*, thus generating a V2 system, we propose that Fin carried [*u*Fin], with the movement of V and AUX – both categories bearing [*i*Fin] – being responsible for the deletion of these uninterpretable features. In the subsequent period, covering the 18th and 19th centuries, the loss of uninterpretable features in Fin, a common occurrence in linguistic change (cf. Roberts & Roussou, 2003), meant that its [*i*Fin] features ceased to trigger verb movement to this high position. As a consequence, both lexical and auxiliary verbs lost movement to Fin, resulting in a system that was no longer V2. With this loss of movement, both types of verbs came to follow *já* but precede lower adverbs such as *ainda*, *sempre*, etc. This suggests that speakers in the 18th and 19th centuries moved both lexical and auxiliary verbs to the T field, which carried [*u*T] features. Once again, in the 20th and 21st centuries, there was a loss of [*u*T] features in T, but only concerning lexical verbs: auxiliary verbs continued to move to the T field, checking [*u*T] features and preceding lower adverbs than *já*, whereas lexical verbs lost movement to the T field, moving only as far as the Asp field (cf. Cyrino, 2013; Araújo-Adriano, 2023, who argue that lexical verbs in contemporary Portuguese move up to Aspect), where they checked their [*u*Asp] features and, consequently, followed aspectual adverbs such as *sempre*, *quase* while preceding very low adverbs such as *bem*. Thus, we observe three distinct periods throughout the history of BP:

(57) a. The grammatical system of BP in the 17th century
[Fin$_{[\mathrm{uFin}]}$[Mood$_{[\mathrm{uM}]}$[Tense$_{[\mathrm{uT}]}$[Aspect$_{[\mathrm{uAsp}]}$[v_{AUX} - *v* - VP]]]]]
b. The grammatical system of BP in the 18th e 19th centuries
[Fin$_{[\mathrm{iFin}]}$[Mood$_{[\mathrm{iM}]}$[Tense$_{[\mathrm{uT}]}$[Aspect$_{[\mathrm{uAsp}]}$[v_{AUX} - *v* - VP]]]]]

c. The grammatical system of BP in the 20th e 21st centuries
[$\text{Fin}_{[iFin]}$[$\text{Mood}_{[iM]}$[$\text{Tense}_{[iT]}$[$\text{Aspect}_{[uAsp]}$[v - VP]]]]]
[$\text{Fin}_{[iFin]}$[$\text{Mood}_{[iM]}$[$\text{Tense}_{[uT]}$[$\text{Aspect}_{[uAsp]}$[v_{AUX}]]]]]

The perspective we adopt on linguistic change is an emergentist one, deriving from the Parameter Hierarchy framework (Roberts, 2012; Biberauer & Roberts, 2012, 2015; Biberauer & Walkden, 2015; Biberauer, 2017; Roberts, 2019a). This approach follows Chomsky (2005) in arguing that, in order to achieve explanatory adequacy, Universal Grammar (UG) must be as minimal as possible, avoiding a large number of parameters that would differentiate one language from another, as was the case in the traditional parametric approach (Chomsky, 1981). Instead, the Faculty of Language is assumed to be shaped by three factors: UG (Factor 1), primary linguistic data (Factor 2), and principles of computational efficiency (Factor 3).

Given the minimal nature of UG, the core premise of the Parameter Hierarchy approach is that variation "arises from under-specified aspects of UG, and is structured by third-factor properties arising largely from the need for efficient learning" (Roberts, 2012, p. 321). Thus, parameters are not part of UG but are determined by learning strategies during language acquisition (on the relationship between change and language acquisition, cf. Lightfoot, 1979), i.e., an emergent property arising from the interaction of the three factors. The central hypothesis of the Parameter Hierarchy approach is that parametric variation does not affect all lexical items but is restricted to functional items. Consequently, parameterisation emerges when the child entertains different features in functional heads within the lexicon, as formulated in the Borer-Chomsky Conjecture, presented below:

(35) The Borer-Chomsky Conjecture (Baker, 2008, p. 353)
All parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon.

From this perspective, cross-linguistic differences are a corollary of aggregates of parameter hierarchies, ranging from a feature that is highly spread across functional heads to features that are highly restricted – not to all but to certain functional heads. This results in a taxonomy of parameters (Biberauer & Roberts, 2015, 2012, 2016; Biberauer, 2017; Roberts, 2019):

(36) Types of parameters

For a given value v_i of a parametrically variant feature F:

a) Macroparameters: all heads of the relevant type share v_i;

b) Mesoparameters: all functional heads of a given category (*e.g.* all verbal heads, all nominal heads, all φ-bearing heads or all finite Cs) share v_i;

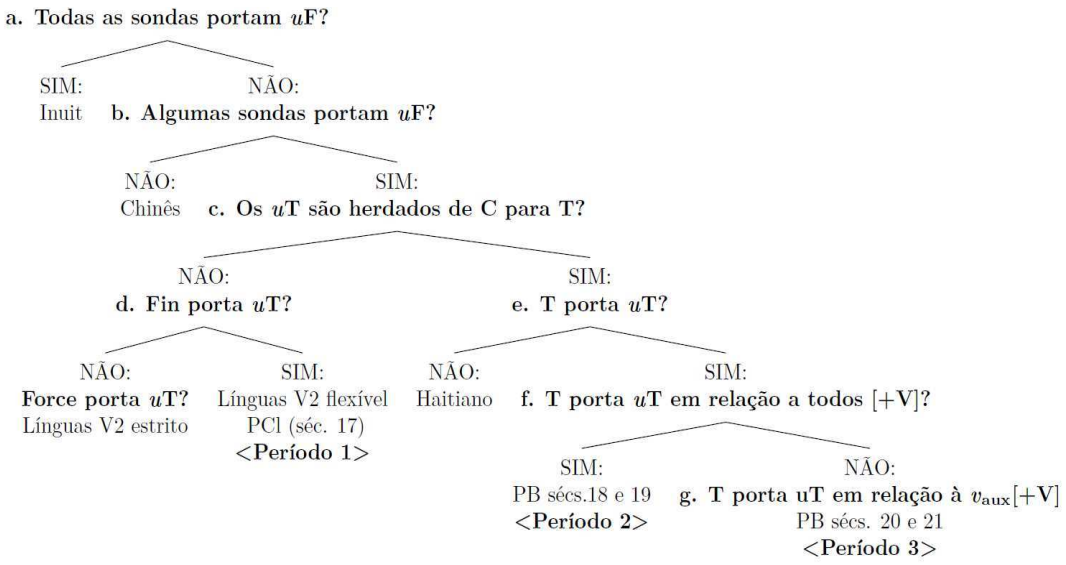
c) Microparameters: a small subclass of functional heads (*e.g.* auxiliaries, pronouns) share v_i;

d) Nanoparameters: one or more idiosyncratic lexical items are specified for v_i.

The taxonomy outlined above combines functional heads and formal features into sets of varying sizes. Thus, macroparameters and mesoparameters are viewed as aggregates of different sizes of microparameters, insofar as they represent distinct distributions of identical formal features across functional heads. Furthermore, these combinations capture seemingly discordant variations without proliferating parameters or formal features and drastically constrain the space of possible grammars by creating dependencies between parameters (Roberts & Holmberg, 2010).

As previously stated, the change observed across the three identified periods can be attributed to the presence or absence of a formal feature born by functional heads. Based on the diachronic results, we identify a period (a) in which Fin bears [*u*Fin] in relation to all verbal elements, followed by a period (b) in which Fin loses its [*u*Fin] features, thereby reducing the scope of verb movement, with verbs subsequently moving to T due to their [*u*T] features in relation to v_{aux} and V. Finally, we observe a third period (c) in which T has restricted its [*u*T] features to auxiliary verbs only, becoming [*i*T] in relation to lexical verbs. Considering the Parameter Hierarchy approach, we present our explanatory proposal regarding the change in verb order in Brazilian Portuguese.

(37)



In Classical Portuguese, children lacked evidence in their primary linguistic data that all probes carried an uninterpretable feature (for instance, an N did not incorporate into T, as would be the case if the child were exposed to a language such as Inuktitut: *ti-seuan-mU-ban/1*SG:*see-man*-PAST 'I saw the man', Baker, 1996, p. 327). Therefore, it is plausible to assert that children of this period marked NO for (37a). Probes that do not carry an uninterpretable feature instantiate a language such as Chinese, where all heads are "floating", as in *ben shu*/CL book 'book' (Huang, 2015, p. 8). Since Brazilian children have never been exposed to this type of input, they have consistently marked YES for (37b).

The variation across the different periods identified in Brazilian Portuguese originates from where the Tense feature appears, as in (37c). If T features remain in C, then Force or Fin may bear them. If T features remain in Fin, this triggers verb movement to Fin, consequently resulting in a flexible V2 system. If Force, a higher projection than Fin, bears the T features, the outcome is a strict V2 system, in which the verb moves to Force.

What is at stake is whether the T features are inherited by T (37c). In Classical Portuguese and the Portuguese written by Brazilians in the 17th century, T features remain "retained" in Fin, as evidenced in primary linguistic data by object fronting, as in (22a). Thus, children of this period marked NO for (37c). Further evidence that Fin, rather than Force, born [*u*F] lies in the occurrence of sentences such as *e [destes] [totalmente]* ***ficam*** *isentos os ricos* (va_017) 'and [from.these] [totally] be.PRS.3PL exempt the rich', which exhibits a V>3 order—highly restricted in ForceV2 languages but common in FinV2 languages (cf. section 1).

By contrast, in the Portuguese written by Brazilians in the 18th century, C donates T features to T, as evidenced in primary linguistic data by the loss of object fronting, as in *Todos minha paixão* ***vêm*** (b_011) 'All my passion come.PRS.3PL'. This leads writers born in the 18th century to mark YES for (37c), with a gradual loss of the property of retaining T features and a shift towards donating them to T. Thus, we observe remnants of a V2 system (with OVS order still being attested), particularly concerning auxiliary verbs, as discussed in 3.1. This can be explained by what Kroch (1994, 2001) refers to as grammar competition – a competition between YES and NO for the parameter in (37c).

Also referred to as syntactic diglossia, the competition of grammars reflects a period of covariation between grammatically incompatible options in both speech and writing within a speech community. Diglossic situations would involve the contrast between the more conservative variety (NO for (37c)) and the more innovative variety (YES for (37c)). By linking the process of syntactic change to acquisition (Lightfoot, 1979), it is

assumed that syntactic change relates to a significant alteration in primary linguistic data. This alteration, although ultimately stabilising in an abrupt manner, appears in the data through the competition of grammars, as observed in the 18th century (Kroch, 1994; 2001).

Thus, when confronted with primary linguistic data and having marked YES for parameter (37c), the Brazilian child then determines whether T in their language bears [*u*T] features or not (37e). If exposed to a language in which temporal morphemes are directly inserted in TMA, as in Haitian Creole *Mwen pa kwè pèsonn ap vini* (1SG NEG believe nobody FUT come 'I do not believe that anyone will come'), a language lacking features that attract V (Roberts, 2017, p. 333), the marked value for (37e) would be NO. In the history of Portuguese, however, primary linguistic data of the Haitian type never existed, which would have led 19th-century Brazilian speakers to mark (37e) with a value other than YES.

Evidence of verbs preceding certain adverbs, such as *ainda* ('still'), *sempre* ('always'), *quase* ('almost') etc. (*e.g.*, *Não dei* ***ainda*** *motivos* 'no give.PST.1SG still reasons'/ *Tenho* ***ainda*** *que comprar um par de botinas e fazer a barba* 'have.PRS.1SG still that buy.INF a pair of boots and do.INF the beard' or *O senhor nunca nos incomoda, dá-nos* ***sempre*** *muito prazer* 'the sir never CL bother.PRS.3SG, give.PRS.3SG=CL always great pleasure'/ *Se eu soubesse que havia de ser* ***sempre*** *tão feliz, casar-me-ia cinquenta* 'If I know.PST.SBJ.1SG that have.PST.3SG of be.INF always so happy marry=CL=FUT.PST.3SG fifty'), suggests that writers born in the 18th and 19th centuries had sufficient indications that T bore [*u*T] features in relation to all [+V] elements, encompassing both lexical and auxiliary verbs. Consequently, the value YES was marked for (37f), meaning that all finite verbs were probed by T *via* [*u*T].

Another variation across the different periods observed in Brazilian Portuguese originates from the reanalysis conducted by children in the 18th and 19th centuries regarding the [*u*T] feature of T. Up until the 19th century, Brazilian children had a stable grammar, with their primary linguistic data pointing towards a YES response for the parameter (37f), producing sentences in which both auxiliary and lexical verbs preceded certain adverbs. For external reasons[12] – whether due to principles of economy or language contact – lexical verbs in Brazilian Portuguese lost their temporal interpretations (cf. Cyrino, 2013; Tescari Neto *et al.*, 2021; Araújo-Adriano, 2023). This may account for why children lost evidence of [*u*T] in T concerning lexical verbs. In other words, lexical verbs – but not auxiliaries – ceased to establish a relationship with T, instead becoming associated solely with lower projections, such as Aspect (cf. Cyrino, 2013; Araújo-Adriano,

[12] These are issues that, due to scope limitations, prevent us from engaging in a more in-depth discussion.

2023). Thus, faced with an unstable grammar, children in the 20th and 21st centuries had to process their linguistic data and "move down" the Parameter Hierarchy to determine the optimal parametric solution for their primary linguistic data.

Consequently, data featuring only auxiliary verbs, but not lexical verbs, preceding adverbs such as *ainda* ('still'), *sempre* ('always'), and *quase* ('almost') no longer conform to the YES response for (37f). As a result, children in the 20th and 21st centuries restricted the elements probed by T to auxiliary verbs alone, stabilising their grammar with a NO response for (37f). Over time, we thus observe distinct parametric settings giving rise to distinct linguistic systems:

(38) Classical Colonial Portuguese (YES > 38d) >
Colonial Brazilian Portuguese (YES > 38f) >
Modern Brazilian Portuguese (NO > 38f).

Despite being initially developed for synchronic cross-linguistic comparison (Roberts, 2012; Biberauer & Roberts, 2012, 2015; Biberauer & Walkden, 2015; Biberauer, 2017; Roberts, 2019), it is evident that an approach such as the Parameter Hierarchy is also a fruitful framework for diachronic analysis. Just as each parametric setting leads a child to acquire a different target language, in the context of language change, each parametric setting would also lead a child to a different target language – despite these distinct target languages corresponding to different historical stages of the same language.

How, then, does the analysis presented here contribute to re-evaluating existing periodisation proposals? The data provide evidence that verb raising underwent three distinct periods of differentiation. **Period I** – *Classical Colonial Portuguese*: The language transplanted to the colony retains the same pattern as Classical Portuguese. We found evidence – such as the raising of both auxiliary and lexical verbs over the adverb *já* ('already') and the occurrence of OVS word order – that texts written by Brazilians in the 17th century exhibit a V2 grammar. **Period II** – *Colonial Brazilian Portuguese*[13]: Lexical verbs lose movement, ceasing to raise over *já*, whilst auxiliary verb raising persists and remnants of OVS order remain. **Period III** – *Modern Brazilian Portuguese*: The loss of verb movement becomes even more pronounced in lexical verbs, though not in auxiliary verbs. Based upon this, we propose the following periodisation.

[13] We emphasise that the terminology adopted here, although similar, diverges from the proposal put forward by Cardoso, Andrade, and Carneiro (2023). This is because we consider the grammar of what we term Colonial Brazilian Portuguese to bear the hallmarks of a Brazilian vernacular grammar, in contrast to the claims made by those authors.

Nova proposta de periodização para o português brasileiro		
Nossa proposta (2025)		
1ª fase Português Clássico Colonial (até 1700)	2ª fase Português Brasileiro Colonial (1700-1899)	3ª fase Português Brasileiro Moderno (1900-atual)

The periodisation proposed here is based on linguistic evidence, that is, on the internal history of the language. We reaffirm previous proposals that (i) the roots of Brazilian Portuguese emerged in the 18th century, and (ii) there has been a transitional phase between a classical grammar – the language of the caravels (Galves, 2007) – and the grammar of modern BP. Colonial Brazilian Portuguese, as this transitional phase between classical grammar and modern BP grammar, is not merely a reflection of the change from Classical Portuguese to European Portuguese (EP). Instead, it is only superficially influenced by this shift through the process of grammatisation/standardisation, which took place in Brazil, particularly from the second half of the 19th century onwards (cf. Pagotto, 1998). We further argue, following Corôa (2022), that Colonial Brazilian Portuguese bears little resemblance to EP grammar, as, by the 18th century, there were already indications of idiosyncratic phenomena in BP grammar, such as: the restriction of subject inversion to unaccusative verbs, the generalisation of proclisis in matrix sentences, the presence of indefinite subjects, and instances of referential null objects.

Although our proposal is grounded in internal linguistic factors, it also reflects aspects identified in previous periodisation proposals, particularly:

(i) Marlos Pessoa de Barros (2003), as our second phase aligns, to some extent, with what Barros terms *the stage of the formation of common Brazilian Portuguese*, which we refer to as *Colonial Brazilian Portuguese*;

(ii) Volker Noll (2008), whose *differentiation phase*, in which Brazilian Portuguese is formed and diverges from the European variety, corresponds to what we call *Colonial Brazilian Portuguese*.

Finally, we highlight the importance of integrating the findings of this study with the periodisation proposals outlined in the introduction, in order to gain a deeper understanding of the sociohistorical and linguistic factors shaping the diachrony of Portuguese in Brazil.

5 Final Remarks

In this paper, we have discussed the periodisation of Brazilian Portuguese based on evidence from the changing position of the verb in relation to adverbs throughout history. We analysed the positioning of the verb relative to the adverbs *bem* ('well'), *quase* ('almost'), *sempre* ('always'), *ainda* ('still'), and *já* ('already') across the 17th to 21st centuries. Based on this diachronic data, we observed that verb placement underwent three significant shifts over time.

Initially, both lexical and auxiliary verbs could precede *já* and all other lower adverbs. Additionally, we found evidence of object fronting, with OVS word order pointing to a V2-type grammar (Period I). This ordering ceased to be the preferred structure for lexical verbs from the 18th century onwards, although some remnants of the previous grammar persisted. However, in contexts with auxiliary verbs, the earlier grammar appeared to remain stable, as auxiliaries continued to precede *já* in the 18th century, just as in the 17th century. By the 19th century (Period II), however, no traces of the 17th-century grammar remained: both lexical and auxiliary verbs were now required to follow *já*, though they still preceded the lower adverbs *ainda*, *sempre*, *quase*, and *bem*. In the 20th and 21st centuries (Period III), yet another change in verb ordering emerged. During this period, lexical verbs no longer preceded *ainda*, *quase*, or *sempre*, but continued to appear before *bem*. By contrast, auxiliary verbs maintained their tendency to precede all these adverbs except *já*, retaining the pattern observed in the grammars of the 18th and 19th centuries.

We formalised this change within the Parameter Hierarchy framework, proposing that in Period I, uninterpretable features in Fin [*u*Fin] triggered verb movement to Fin, yielding word orders in which the verb preceded all analysed adverbs, consistent with a V2 grammar. In Period II, these features were lost, and Fin came to bear interpretable features [*i*Fin]. During this phase, T bore uninterpretable features [*u*T] for all verbal elements, triggering movement not only of auxiliary verbs but also of lexical verbs. Finally, in Period III, T retained its features in relation to auxiliary verbs but restricted them for lexical verbs, bearing interpretable features [*i*T]. As a corollary, a new pattern of verb movement arose: only auxiliary verbs moved to the T field, whereas lexical verbs moved to a position lower than T.

The data presented here highlight the need to reassess existing periodisation proposals for the history of Portuguese in Brazil. Accordingly, we propose a periodisation that aligns with the significant word order changes identified in this study. The pro-

posed framework is divided into the following phases: Period I, which we term Classical Colonial Portuguese; Period II, referred to as Colonial Brazilian Portuguese; and Period III, designated Modern Brazilian Portuguese.

We acknowledge that further analyses, including investigations of other syntactic phenomena, are necessary to either confirm or challenge the findings presented here. For now, we offer this contribution as a call to re-examine the syntax of past periods, in order to better understand when Brazilian Portuguese grammar actually emerged.

References

ANDRIANI, L. The modal value of ancora/angórə in Barese and northern Apulian varieties. *Cambridge Occasional Papers in Linguistics*, Cambridge, UK, v. 9, p. 235–245, 2016.

ANTONELLI, A. L.. Posição do verbo no português clássico: evidências de um sistema V2. Alfa: Revista de Linguística (São José do Rio Preto), v. 55, n. 2, p. 501–522, jul. 2011.

ANTONELLI, A.L. Sintaxe da posição do verbo e mudança gramatical na história do português europeu. 230f.p. Tese (Doutorado em Linguística)—Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, 2011.

ARAÚJO-ADRIANO, P. Â. The position of the verb with respect to the adverb `sempre' over four centuries: diagnosis for the (loss of high) verb movement in Brazilian Portuguese. *Journal of Historical Syntax*, [*s. l.*], v. 6, n. 3, p. 1–50, 2022.

ARAÚJO-ADRIANO, P. Â.; CORÔA, W. Reconstruindo a história do português do brasil pelo corpus Tycho Brahe Brasil: novos dados, novos olhares. *Revista Linguíftica*, [*s. l.*], v. 17, n. 3, p. 202–227, 2022.

ARAÚJO-ADRIANO, P. Â.; CYRINO, S. The looseness of auxiliary constructions in Brazilian Portuguese. Paper presented at the III Encontro de Gramática Gerativa, Salvador, 2022.

ARAÚJO-ADRIANO, P. Â. The Present Tense Analyticisation Process in Brazilian Portuguese: A Diachronic Approach. 2023. PhD Dissertation - Universidade Estadual de Campinas, Campinas, 2023.

ARAÚJO-ADRIANO, P. Â.; CYRINO, S. On the lack of adjacency in auxiliary constructions in Brazilian Portuguesew. *Cadernos de Estudos Linguísticos*, 2025.

BAKER, M. The macroparameter in a microparametric world. *In*: BIBERAUER, T. (ed.). *The Limits of Syntactic Variation*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2008. p. 351–374.

BELLETTI, A. *Generalized verb movement: aspects of verb syntax*. Torino: Rosenberg & Sellier, 1990.

BIBERAUER, T. & ROBERTS, I. Towards a parameter hierarchy for auxiliaries: diachronic considerations. *Cambridge Occasional Papers in Linguistics* 6(9). 267–294. Place: Cambridge, UK, 2012.

BIBERAUER, T. & ROBERTS, I. Rethinking formal hierarchies: a proposed unification. *Cambridge Occasional Papers in Linguistics* 7(1). 1–31. Place: Cambridge, UK, 2015b.

BIBERAUER, T & WALKDEN, G. (eds.). *Syntax over time: lexical, morphological, and information-structural interactions* (Oxford studies in diachronic and historical linguistics 15). Oxford, United Kingdom: Oxford University Press, 2015.

BIBERAUER, T. Factors 2 and 3: A principled approach. *Cambridge occasional papers in linguistics* 10. 38–65. Place: Cambridge, UK, 2017.

BIBERAUER, T. & ROBERTS, I. Parameter setting. In Adam Ledgeway & Ian G. Roberts (eds.), *The Cambridge Handbook of historical syntax*/Cambridge Handbooks in Language and Linguistics, 234–162. Cambridge: Cambridge University Press, 2017.

CARDOSO, L.; ANDRADE, A. L. D.; CARNEIRO, Z. O português colonial brasileiro: uma nova agenda de pesquisas entre o português clássico e o português brasileiro moderno. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada,* [*s. l.*], v. 39, n. 2, p. 1–27, 2023.

CHOMSKY, N. *Lectures on government and binding: the Pisa Lectures*. Dordrecht, Holland ; Cinnaminson, [N.J.]: Foris Publications, 1981. (Studies in generative grammar, v. 9).

CHOMSKY, N. Three Factors in Language Design. *Linguistic Inquiry*, Massachusetts, v. 36, n. 1, p. 1–22, 2005.

CHOMSKY, N. On phases. *In*: FREIDIN, R.; OTERO, C. P.; ZUBIZARRETA, M. L. (ed.). *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*. [*S. l.: s. n.*], 2008. (Current studies in linguistics). p. 133–166.

CHOMSKY, N. *The Minimalist Program*. Massachusetts: The MIT Press, 1995.

CINQUE, G. (org.). *Restructuring and Functional Heads. The Cartography of Syntactic Structures*. [*S. l.*]: Oxford University Press, 2006. v. 4

CINQUE, G. *Adverbs and functional heads.* Oxford ; New York: Oxford University Press, 1999.

CORÔA, W. *Rastreando as origens do Português Brasileiro: a dinâmica da mudança na escrita de "Homens Bons" na Bahia Colonial.* 2022. 406 f. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, Campinas, SP, 2022.

CYRINO, S. On richness of tense and verb movement in Brazilian Portuguese. *In*: CAMACHO-TABOADA, V. *et al.* (ed.). *Information Structure and Agreement*. Amsterdam: John Benjamins Publishing Company, 2013. p. 297–317.

EMONDS, J. The verbal complex V'-V in French. *Linguistic Inquiry*, Massachusetts, v. 9, n. 2, p. 151–175, 1978.

GALVES, C. A língua das caravelas: periodização do português europeu e origem do português brasileiro. In: CASTILHO, A. d.; MORAES, M. A. T.; LOPES, R.; CYRINO, S.(Ed.). *Descrição, história e aquisição do português brasileiro*. Campinas: Pontes, 2007. p.513–528.

GALVES, C. Relaxed verb second in classical portuguese. In: WOODS, R.; WOLFE, S. (Ed.). *Rethinking verb second.* Oxford: Oxford University Press, 2020. p. 368–395.

GALVES, C.; PAIXÃO DE SOUSA, M. C. The change in the position of the verb in the history of portuguese: Subject realization, clitic placement, and prosody. *Language, Linguistic Society of America*, v. 93, n. 3, p. e152–e180, 2017. ISSN1535-0665.

GALVES, C.; GIBRAIL, A. Subject inversion intransitive sentences from classical to modern european portuguese: A corpus-based study. In: MARTINS, A. M.; CARDOSO, A. (Ed.). *Word Order Change*. Oxford: Oxford University Press,2018.

GIBRAIL, A. *Contextos de formação de estruturas tópico foco no Português Clássico*. 2010. 298 f. Tese (Doutorado em Linguística) — Universidade Estadual de Campinas (UNICAMP). Instituto de Estudos da Linguagem, Campinas, 2010.

HAEGEMAN, L. Verb-second, the split cp and null subjects in early Dutch finite clauses. *GenGenP* (Geneva Generative Papers), v. 4, p. 133–175, 1996.

HOLMBERG, A. Verb second. In: KISS, T.; ALEXIADOU, A. (Ed.). *Syntax theory and analysis*: an international handbook of contemporary syntactic research. Berlin:[s.n.], 2015. p. 342–383.

KATO, M. O português são dois… ou três?. *In*: LOBO, T. *et al.* (org.). *Rosae: linguística histórica, história das línguas e outras histórias*. Salvador: EDUFBA., 2012. p. 93–108.

KROCH, A. S. Syntactic Change. *In*: BALTIN, M.; COLLINS, C. (ed.). *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell Publishers, 2001. p. 699–729.

KROCH, Anthony. Morpho-syntactic variation. In: BEALS, Kenneth et al. (Orgs.). *Papers from the 30th Regional Meeting of the Chicago Linguistics Society:* Parasession on Variation e Linguistic Theory, Chicago, v. 2, p. 180-201, 1994.

LEDGEWAY, A.; LOMBARDI, A. Verb Movement, Adverbs and Clitic Positions in Romance. *Probus*, [*s. l.*], v. 17, n. 1, 2005. Disponível em: Acesso em: 20 mar. 2022.

LIGHTFOOT, D. *Principles of diachronic syntax*. Cambridge, UK: Cambridge University Press, 1979.

LUNGUINHO, M. V. da S. Verbos auxiliares e a sintaxe dos domínios não-finitos. 2011. 215 f. Tese – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2011.

LOBO, T. A questão da periodização da história linguística do Brasil. In: CASTRO, I.; DUARTE, I. (Ed.). *Razões e emoção: miscelânea de estudos em homenagem a Maria Helena Mateus.* [S.l.]: Imprensa Nacional-Casa da Moeda, 2003. p. 395–409.

LOPES, A. C. M. Contribuição para o estudo dos valores discursivos de "sempre". *In*: *ACTAS DO XIII ENCONTRO NACIONAL DA APL* 1997. Lisboa: Associação Portuguesa de Linguística, 1998. v. II, p. 3–14.

LUCCHESI, D. A periodização da história sociolinguística do Brasil. *DELTA: Documentação de Estudos em Lingística Teórica e Aplicada*, Fap UNIFESP (SciELO), v. 33, n. 2, p. 347–382, aug 2017.

MARTINS, A. M. Clíticos na história do português. Tese (Doutorado)—Universidade de Lisboa, 1994.

MATOS, G.; CYRINO, S. *Elipse de VP no Português Europeu e no Português Brasileiro*. Universidade Federal do Ceará: [*s. n.*], 2001.

PAIXÃO DE SOUSA, M. C. *Língua Barroca: Sintaxe e História do português nos 1600*. Tese (Doutorado emLinguística)—Universidade Estadual de Campinas (UNICAMP), Campinas, 2004.

PESSOA, M. d. B. *Formação de uma variedade urbana e semi-oralidade*: o caso do Recife, Brasil. Tubigen: Niemeyer, 2003.

POLLOCK, J.-Y. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry*, [*s. l.*], v. 20, n. 3, p. 365–424, 1989.

RIBEIRO, I. A sintaxe da ordem no português arcaico: o efeito V2. Unicamp. 286f. Tese (Doutorado emLinguística)—Universidade Estadual de Campinas (UNICAMP). Instituto de Estudos da Linguagem, Campinas, 1995.

RIBEIRO, I. A mudança sintática do português brasileiro é mudança em relação a que gramática. In: CASTILHO, A. T. d. (Ed.). *Para História do Português Brasileiro*. São Paulo: Humanitas, 1998. p. 101–119.

RIZZI, L. The fine structure of the left periphery. In: HAEGEMAN, L. (Ed.). *Elements of grammar*. Dordrecht: Springer, 1997. p.281–337.

ROBERTS, I. Macroparameters and minimalism: A programme for comparative research. *In*: GALVES, C. *et al.* (ed.). *Parameter theory and linguistic change*. Oxford: Oxford University Press, 2012. (Oxford studies in diachronic and historical linguistics, v. s). p. 319–334.

ROBERTS, I. On some languages lacking V-to-I movement. *A Schrift to Fest Kyle Johnson*, Linguistics Open Access Publications, p. 321–335, 2017.

ROBERTS, I. *Parameter hierarchies and Universal Grammar.* Oxford: Oxford University Press, 2019.

ROBERTS, I.; ROUSSOU, A. *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge, UK: Cambridge University Press, 2003. (Cambridge Studies in Linguistics, v. 100).

ROBERTS, I. The C-System in Brythonic Celtic Languages, V2, and the EPP, *In*: Luigi Rizzi (ed.), *The Structure of CP and IP*: The Cartography of Syntactic Structures, Volume 2, New York, NY, 2004; online edn, Oxford Academic, 2023.

ROBERTS, I. *Parameter hierarchies and Universal Grammar*. Oxford: Oxford University Press, 2019.

SAPIR, Edward. *Language*: an introduction to the study of speech San Diego: HJB Books, [1921] 1949.

SILVA NETO, Serafim da. Introdução ao estudo da língua portuguesa no Brasil. 2 ed. Rio de Janeiro: INL, 1986[1951].

TARALLO, F. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX. *In*: ROBERTS, I.; KATO, M. (org.). *Português brasileiro: uma viagem diacrônica*. 3. ed. São Paulo: Editora Contexto, 1993. p. 54–82.

TESCARI NETO, A. On the Movement of Verbal Forms in Romance and English. *Language at the University of Essex (LangUE) Proceedings 2011*, [*s. l.*], p. 54–67, 2012.

TESCARI NETO, A. *On Verb Movement in Brazilian Portuguese: A Cartographic Study*. 2013. 392 f. Tese (Phd em Filosofia) - Università Ca' Foscari di Venezia, Veneza, 2013.

TESCARI NETO, A. Advérbios e o movimento do verbo. *Fórum Linguístico*, [*s. l.*], v. 16, n. 1, p. 3563–3578, 2019.

TESCARI NETO, A.; FORERO PATAQUIVA, F. de P. Do movimento do verbo finito e infinitivo em português brasileiro e espanhol colombiano: microvariação e cartografias. *Cuadernos de la ALFAL*, [*s. l.*], v. 12, n. 2, p. 491–511, 2020.

TESCARI NETO, A.; FORERO PATAQUIVA, F. de P.; WECHSLER, A. L. R. Hipóteses sobre o que desencadearia o movimento do verbo. *In*: III *Encontro de gramática gerativa homenagem a Sonia Cyrino e Maria Eugênia Duarte*, 2021, Online. Anais [...]. Online: [*s. n.*], 2021.

TEYSSIER, P. *História da língua portuguesa*. São Paulo: Martins Fontes,1984.

TORRES MORAIS, M. A. *Do português clássico ao português europeu moderno: um estudo diacrônico da cliticização e do movimento do verbo*. 1995. 374 f. Thesis (PhD in Linguistics) - Universidade Estadual de Campinas, Campinas, SP, 1995.

NOLL, V. *O português brasileiro: formação e contrastes*. São Paulo: Globo, 2008.

WOLFE, S. A comparative perspective on the evolution of Romance clausal structure. *Diachronica*, JohnBenjamins, v. 33, n. 4, p. 461–502, 2016.

WOLFE, S. *Verb second in medieval Romance*. Oxford: Oxford Studies in Diachronic, 2018.

WOLFE, S. Redefining the typology of v2 languages: *The view from medieval Romance and beyond*. Linguistic Variation, JohnBenjamins, Amsterdam; NewYork, v. 19, n. 1, p. 16–46, 2019.

http://doi.org/10.5007/1984-8420.2025.e100843

A EXPANSÃO DA PREPOSIÇÃO *PARA* NAS DITRANSITIVAS DO PORTUGUÊS BRASILEIRO: A PERSPECTIVA DO MODELO DE TRÊS FATORES

THE EXPANSION OF THE PREPOSITION *PARA* IN DITRANSITIVE SENTENCES IN BRAZILIAN PORTUGUESE: THE PERSPECTIVE OF THE THREE FACTORS MODEL

Ana Calindro | Lattes | anacalindro@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Maria Aparecida Torres Morais | Lattes | torres.mariacida@gmail.com
Universidade de São Paulo/CNPq

Resumo: O percurso histórico do português brasileiro (PB), comparado ao do português europeu (PE), revela inovações na expressão morfossintática do objeto indireto (OI) nas estruturas ditransitivas. (i) Substituição do *a* - marcador de caso dativo- pela preposição *para*- 'O Pedro deu um celular à namorada/*para a namorada*. (ii) Substituição dos clíticos dativos de 3ª pessoa, *lhe/lhes*, pelas formas oblíquas *ele(s), ela(s),* introduzidos por *a/para* - 'O Pedro deu-*lhe* um celular/ deu um celular *a/para ela*'. O objetivo principal deste estudo é apresentar uma abordagem dessas inovações com base no quadro formal da sintaxe, delineado no Modelo dos Três Fatores (Chomsky, 2005), o qual implica uma visão minimalista da Faculdade da Linguagem (Fator 1), o *input* (Fator 2), e aspectos cognitivos gerais (Fator 3). Em particular, assumimos o modelo de aquisição denominado *Maximize Meios Mínimos* (MMM) (Biberauer, 2018, 2019a, 2019b), com os seguintes efeitos linguísticos na abordagem do *input:* (i) Economia de traços – postule o menor número possível de traços formais; (ii) Generalização do Input - aproveite ao máximo os traços formais disponíveis (Roberts, 2007; Roberts e Roussou 2003; Biberauer e Roberts (2017). Nessa perspectiva, a nossa hipótese é a de que, no percurso de aquisição das variedades do PB, os falantes manifestaram um "viés cognitivo", ou seja, um padrão de *sobregeneralização* do *input,* o qual, sem o efeito da retração, impossibilitada pela perda das marcas casuais dativas, resultou no uso generalizado da preposição *para* como introdutora do argumento recipiente/beneficiário nas estruturas ditransitivas.

Palavras-chave: Objeto indireto. Estrutura ditransitiva. Português brasileiro. Modelo dos três fatores. Aquisição, variação e mudança.

http://doi.org/10.5007/1984-8420.2025.e100843

Abstract: Synchronic and diachronic corpora show that, when compared to European Portuguese, Brazilian Portuguese displays two innovative properties regarding the morphosyntactic realization of the indirect object in ditransitive sentences. In its nominal expression, the preposition *a* - a case marker - is replaced by the preposition *para* – O Pedro deu um celular *à namorada / para a namorada.* 'Pedro gave *his girlfriend* a cell phone / a cell phone *to his girlfriend*'. In its 3rd person pronominal expression, the dative clitics *lhe/lhes* are replaced by the oblique pronouns *ele(s), ela(s)*, introduced by prepositions *a/para* – O Pedro deu-*lhe* um celular/ deu um celular *a/para ela.* 'Pedro gave *her* a cell phone/gave a cell phone *to her*'. The aim of this paper is to present a proposal for understanding the following question: Why BP speakers, during the centuries of its expansion in Brazil, have converged to a grammatical system divergent from the linguistic input provided by the Portuguese speakers? Our answer is based on a minimalist approach to acquisition, variation and change as outlined in the Three Factors model (Chomsky, 2005), highlighting the input (Factor 2), and the general cognitive aspects related to the speaker's interaction with the linguistic input (Factor 3). We assume, based on Roberts (2007), Roberts and Roussou (2003), and Biberauer and Roberts (2017), two linguistic effects of a cognitive bias, named Maximize Minimal Means (MMM) (Biberauer (2018, 2019a, 2019b): (i) *Feature economy* - Postulate the fewest number of formal features. (ii) *Input Generalization* - Make the most out of the available features. In this perspective, our hypothesis is that, during the acquisition of the Brazilian Portuguese, occurs a process of overgeneralization in the use of the preposition *para* to introduce the indirect object recipient/benefactive in the context of ditransitive verbs.

Keywords: Indirect Object. Ditransitive structures. Brazilian Portuguese. Three Factors Model. Acquisition, variation and change.

Introdução

Estudos sobre a sintaxe diacrônica do português europeu (PE) e do português brasileiro (PB) identificam, ao longo do século XVIII, "pontos de inflexão" nos seus sistemas gramaticais. Com relação ao PE, observou-se uma maior produtividade de certos fenômenos sintáticos, tais como a colocação enclítica dos pronomes átonos na ausência de atratores de próclise (cf. Martins, 2002, 2016) e a posição pré-verbal do sujeito nas sentenças transitivas, configurando a ordem SVO não-marcada. Segundo alguns autores, o século XVIII representaria o final do português clássico e o início do PE (Galves, 2007, 2020, entre outros).

Da mesma forma, os dados provenientes dos *corpora* produzidos por autores brasileiros do século XVIII e, mais especificamente, do século XIX, permitem, entrever rumos distintos da gramática brasileira, desenvolvendo características próprias. Em particular, tem-se enfatizado as marcantes alterações no quadro dos pronomes pessoais e dos pronomes possessivos, com profundas consequências de natureza morfossintática e semântica. Destacam-se igualmente as mudanças na colocação dos pronomes clíticos, uma vez que a gramática brasileira passa a revelar nos textos uma próclise generalizada, independente de condicionamentos sintáticos (cf. entre muitas outras obras, artigos, dissertações e teses, Wetzels, Costa e Menuzzi, 2016; Roberts e Kato, 2018; Galves, Torres Morais, Ribeiro, 2005; Galves, Kato e Roberts, 2019; Kato, Martins e Nunes, 2023; Calindro, Alves, Martins, 2024). Vale ressaltar ainda que, durante o período colonial, entre os séculos XVI e XIX, o Brasil era um país acentuadamente plurilíngue pelo contato entre os portugueses, povos originários e africanos.

Considerando um cenário externo de efervescência multicultural e multilinguística, este estudo se dedica à discussão de estratégias gramaticais inovadoras na codificação do argumento dativo/objeto indireto (OI), as quais revelam um interessante processo diacrônico na história do PB, amplamente documentado nos bancos de dados, durante os séculos XVIII, XIX e XX, até chegar às variedades sincrônicas, coloquiais e formais, da fala e da escrita. Destacaremos duas delas, expressas na realização morfossintática do OI, na sua expressão nominal e pronominal de 3ª pessoa nas estruturas ditransitivas*:* (i) substituição da preposição *a* - marcador de caso dativo - pela preposição *para* (1a)[1]; (ii) substituição das formas clíticas dativas *lhe, lhes,* pelas formas pronominais oblíquas *ele(s), ela(s),* introduzidas pelas preposições *a/para* (1b).

(1) a. O Pedro deu o celular *para a namorada*.
b. O Pedro deu o celular *a/para ela*.

Observe-se que o reconhecimento do caráter inovador atribuído a essas propriedades decorre igualmente de uma abordagem comparativa com dados provenientes do português arcaico, do português medieval e do português europeu moderno, nos quais a identificação do objeto indireto, nos contextos ditransitivos com verbos de transferência

[1] Vale ressaltar que a substituição do clítico dativo de 3ª pessoa pelas formas tônicas preposicionadas *a/para ele(s), a/para ela(s)*; *dele/dela(s) atinge o paradigma completo dos clíticos de 1ª e 2ª pessoas que se mantém na gramática pronominal, mas alternam* com sintagmas preposicionados *para mim*, *para você.* Ressalte-se que, em algumas regiões do Brasil, em que permaneceu o pronome *tu*, ainda se usa *para ti* (cf. Lopes, 2007). Ademais, o clítico *lhe* foi dialetalmente recategorizado como dativo e acusativo de 2ª pessoa (cf. Sales, 2007; Araújo e Carvalho, 2015; Martins; Moura; Costa da Silva, 2019).

e movimento, e verbos de criação, envolve categoricamente a sua pronominalização pelo clítico dativo *lhe(s)* (2b) e a presença da preposição casual *a* (2a) Vale notar que a alternância com formas pronominais introduzidas por preposições plenas é agramatical (2c).

(2) a. O Pedro deu o celular à professora.
b. O Pedro deu-*lhe* o celular.
c. *O Pedro deu o celular *a ela.*

Os fatos descritos em (1) e (2) motivam, pois, o objetivo central deste estudo que é o de apresentar uma abordagem formal da aquisição, variação e mudança para os novos alinhamentos preposicionais na expressão gramatical do OI no PB. Tal objetivo é permeado pela seguinte questão: Assumindo que a gramática dativa é um traço conservador no percurso histórico do PE, o que teria motivado a sua perda no PB? A nossa proposta para a expansão da preposição *para* no PB considera tanto a sua relação com as reanálises casuais no quadro do sistema pronominal, quanto o fato de que os predicados ditransitivos, no PE, podem projetar dois tipos distintos de estrutura sintática. Uma delas implica uma leitura de transferência direta de posse entre dois argumentos, o objeto direto e o OI dativo, interpretado como recipiente/destinatário (cf. 2a e 2b). A outra estrutura envolve um complemento oblíquo (OBL), introduzido pela preposição *para*, configurando um tipo de "alternância beneficiária". A contraparte preposicionada implica um movimento espacial em direção a um participante [+humano], interpretado como beneficiário do evento expresso pelo verbo (cf.3).

(3) a. O Pedro deu o celular *para a professora/para ela.*

O texto está estruturado da seguinte forma: a seção 2 descreve as propriedades codificadoras do objeto indireto nas estruturas ditransitivas do PE, com ênfase na sua expressão morfossintática e no fenômeno denominado "alternância beneficiária". Ambos são relevantes para a reanálises preposicionais que se verificam na expressão do OI no PB. A seção 3 traz uma descrição dos novos alinhamentos das preposições *a* e *para* na expressão do OI de 3ª pessoa pronominal e não pronominal, no percurso histórico do PB; por sua vez a seção 4 traz uma abordagem dessas inovações com base no quadro formal da sintaxe, delineado no Modelo dos Três Fatores (Chomsky, 2005).

Em particular, discutimos o modelo de aquisição denominado *Maximize Meios Mínimos* (MMM) (Biberauer, 2018, 2019a, 2019b), e os seus efeitos linguísticos na abor-

dagem do *input:* (i) *Economia de traços* – postule o menor número possível de traços formais; (ii) *Generalização do Input* - aproveite ao máximo os traços formais disponíveis, nos termos de Roberts (2007), Roberts e Roussou (2003), Biberauer e Roberts (2017).

A nossa hipótese é a de que, os falantes adquirindo o PB manifestaram um padrão de *sobregeneralização* do *input,* o qual, sem o efeito da retração, resultou no uso generalizado da preposição *para* como introdutora do argumento recipiente/beneficiário nas estruturas ditransitivas. Na seção 5, trazemos as nossas considerações finais.

2. Dados do *Input*- A codificação dativa do objeto indireto e sua contraparte preposicionada nas ditransitivas do PE

Nesta seção, apresentamos um panorama descritivo das propriedades morfossintáticas e semânticas que identificam o objeto indireto (OI), em oposição ao argumento oblíquo, nas estruturas ditransitivas da variedade padrão (culta) do português europeu contemporâneo. Como dissemos, essa gramática fazia parte da competência do falante do PE, ou seja, estava na fala dos portugueses, ao longo da história colonizadora do Brasil, servindo potencialmente de *input* para a aquisição do que viria a constituir o PB, seja como língua nativa, seja como segunda língua[2].

2.1. Propriedades identificadoras do objeto indireto no português europeu contemporâneo

O argumento que exerce a função de complemento/objeto indireto no PE no contexto dos predicados ditransitivos é identificado pelos linguistas e nas gramáticas contemporâneas a partir de critérios morfossintáticos e semânticos que não se alinham completamente com os adotados pela tradição gramatical.

Um desses critérios refere-se ao seu estatuto categorial como sintagma nominal definido (DP), precedido obrigatoriamente pelo morfema *a* – um marcador de Caso dativo (cf. Duarte, 1987, 2003) –, e tendo sua pronominalização restrita aos clíticos dativos de

[2] A caracterização das gramáticas dos primeiros portugueses que aqui aportaram tem sido um grande tema da pesquisa recente (cf. Ribeiro, 1998; Moraes de Castilho, 2001, Paixão de Sousa e Kewitz (2011), entre outros). Em particular, Galves (2007) propõe uma periodização do português europeu associada à vinda dos portugueses para o Brasil, considerando um fenômeno linguístico que hoje distingue de forma marcante as duas variedades do português, a saber, a colocação dos pronomes clíticos. É interessante notar que a sintaxe da colocação dos clíticos sofre alterações tanto na história do PE (Martins, 1994) quanto na história do PB, cuja trajetória é marcada por uma renovação no inventário dos clíticos, com perdas e ganhos no quadro pronominal. Consequentemente, o *input* relacionado aos clíticos foi "instável" ao longo dos séculos. Entretanto, as inovações na identificação do OI nas ditransitivas teriam tido uma natureza distinta, uma vez que se manifestaram na diacronia do PB, mas não na diacronia do PE.

3ª pessoa *lhe, lhes*. Por sua vez, em termos semânticos, o argumento está associado a uma restrição de animacidade de seu referente, no contexto de verbos de transferência (4), verbos de movimento (5) e verbos de criação (6).

(4) a. O João deu um colar de pérolas à Maria
b. O João deu-*lhe* um colar de pérolas.

(5) a. O João enviou/ entregou uma carta à Maria.
b. O João enviou/ entregou-*lhe* uma carta

(6) a. A Maria preparou um jantar *ao João.*
b. A Maria preparou-*lhe* um jantar.

Como observam Gonçalves e Raposo (2013), o objeto indireto é um argumento adicionado aos predicados ditransitivos que expressam semanticamente uma relação de transferência possessiva entre dois objetos: o objeto direto (acusativo) e o objeto indireto (dativo). A transferência entre duas entidades e/ou mudança de posse é concebida tanto literal, quanto metaforicamente, por envolver entidades abstratas. O objeto indireto se alinha, portanto, a uma interpretação de recipiente/destinatário/origem, como exemplificado em (6):

(7) a. O Pedro [deu/entregou/vendeu] o livro *ao António.* [destinatário]
b. O Pedro [roubou/tirou/comprou] o livro *ao António.* [origem]
c. A Maria [sorriu/falou/escreveu] (uma carta) *ao Luís*. [destinatário]
d. O professor explicou o teorema *aos alunos.* [destinatário]
(Gonçalves; Raposo, 2013, p.1173)

Na gramática ditransitiva do PE, destacamos outra propriedade relevante, quando pensamos numa abordagem comparativa com o PB: o argumento recipiente é categoricamente pronominalizado pelas formas clíticas, nunca pelas formas oblíquas preposicionadas. Tal restrição é suspensa nas construções de redobro de clítico, nas quais o clítico tem a mesma referência do pronome forte, obtendo-se uma interpretação contrastiva. Na sentença (8), entende-se que Miguel mostrou as fotografias *a ela, Maria*, e não *a ele, João*, por exemplo.

(8) a. *O Miguel mostrou as fotografias *a ela*.
b. O Miguel mostrou-*lhe* as fotografias *a ela*.
(Gonçalves; Raposo, 2013, p. 1172)

Outro contexto licenciador da forma *a + pronomes tônicos* são as respostas curtas a questões diretas.

(9) a. *A quem* mostrastes as fotografias?
b. A ela.
(Gonçalves; Raposo, 2013, p. 1172)

Por sua vez, a sequência *a + pronome tônico* é obrigatória nos contextos verbais que não licenciam um OI. Como exemplificado em (9), o constituinte *ao pai* não é um OI porque não alterna com o clítico *lhe*. Trata-se, portanto, de um complemento preposicionado/oblíquo.

(10) a. O José recorreu *ao pai*, quando estava endividado.
b. *O José recorreu-*lhe.*
c. O José recorreu *a ele*.

O quadro descritivo apresentado nos permite concluir que o dinamismo preposicional da gramática portuguesa se manifesta na ambiguidade da preposição *a*: como marcador de caso e como preposição transitiva. Este ponto também se evidencia no seguinte contraste: o OI dativo pode ser licenciado no contexto dos 'locativos metafóricos' (11a). Nesse caso, a marca de plural da contraparte clítica (*lhes*) indica a correferência com a expressão 'funcionários do Banco de Portugal'. No entanto, na introdução dos 'locativos puros' (11b), a preposição *a* se comporta como uma verdadeira preposição restringindo a pronominalização dativa. Neste caso, o locativo é categoricamente um sintagma preposicionado.

(11) a. Enviei um documento *ao Banco de Portugal* / enviei-*lhes* um documento.
b. O médico enviou o paciente *ao hospital*/ *enviou-*lhe/lhes* o paciente.

Observe-se que, em paralelo ao estatuto locativo da preposição *a*, também a preposição *para* se apresenta como introdutora de locativos, intercambiável com advérbios locativos, mas não com pronominais clíticos (12).

(12) O José enviou uma carta *para Lisboa* / enviou uma carta *para lá* / *enviou-*lhes* uma carta.

2.2. Alternância nas estruturas ditransitivas: argumentos dativos e argumentos oblíquos

As gramáticas contemporâneas do PE estabelecem ainda uma distinção relevante, não reconhecida unanimemente nas gramáticas tradicionais, entre o complemento dativo (OI) das estruturas ditransitivas, interpretado como recipiente/destinatário, e o argumento introduzido pela preposição *para*, denominado argumento oblíquo (OBL), e interpretado como beneficiário do evento expresso pelo verbo[3]. No que se segue enfatizaremos essa distinção.

Comecemos por destacar que, no contexto dos verbos ditransitivos de transferência/movimento, como *dar*, *entregar*, etc, o argumento introduzido por *a* é interpretado como recipiente/destinatário, expressando uma relação possessiva direta entre os dois objetos. Por sua vez, o constituinte introduzido por *para* é interpretado como beneficiário do evento expresso pelo verbo, comportando-se sintaticamente como um sintagma independente, ou seja, como um complemento oblíquo (13a) (cf. Gonçalves; Raposo, 2013).

Consequentemente, é agramatical a coocorrência de dois sintagmas introduzidos por *a*, como exemplificado em (13b)

(13) a. O Marco entregou os documentos [*ao contínuo*] [*para a professora*].
b. *O Marco entregou os documentos [*ao contínuo*] [à professora].
(Gonçalves; Raposo, 2013, p.1177)

[3] Gonçalves e Raposo (2013, p. 904) seguem as gramáticas contemporâneas na sua acepção de que o termo *oblíquo* é usado em dois sentidos distintos, embora relacionados: um morfológico e outro sintático. Na acepção morfológica, o pronome oblíquo toma uma forma casual particular, quando é complemento da preposição (*mim, ti, ele(s), ela(s)*). No PE, as formas oblíquas são sempre fortes (tônicas), enquanto as formas acusativas e dativas são sempre fracas, ou seja, clíticas. Na acepção sintática, o termo é usado para se referir a um complemento do verbo, que não é nem complemento direto, nem indireto, e obrigatoriamente preposicionado. Como os autores esclarecem, a noção de "Caso oblíquo", das gramáticas descritivas contemporâneas difere da noção tradicional, onde o termo "oblíquo" é usado para os Casos acusativo, dativo, genitivo, ablativo, em oposição ao termo "reto" para o Caso nominativo. Brito (2008) destaca ainda que, embora o PE, como as demais línguas românicas, não manifestem a flexão de caso acusativo, dativo e nominativo nos sintagmas nominais, conservam um sistema flexional rico nos pronomes pessoais, distinguindo as formas nominativas (sujeito), acusativas (objeto direto), dativas (objeto indireto) e oblíquas introduzidas por preposições. Para uma descrição dos fatos do PB, confira Castilho (2010).

No emaranhado das sutilezas gramaticais do PE, outro ponto a ser destacado é o que se refere à afirmação dos autores acima citados de que, com certos verbos ditransitivos, o OI pode obter uma interpretação secundária (*itálico nosso*) de beneficiário, dependendo do envolvimento pessoal do seu referente no ato da transferência, em particular se este tem algo a ganhar ou a perder. Assim, na interpretação mais natural de *dei-lhe um livro no seu aniversário*, o pronome clítico dativo é simultaneamente destinatário e beneficiário (Gonçalves e Raposo, 2013, p. 1174).

Ainda relevante é a constatação de que os verbos transitivos como *comprar*, *vender*, *ler*, *escrever* e verbos de criação, como *desenhar*, *preparar*, *pintar*, licenciam uma alternância entre o OI dativo e um OBL, introduzido pela preposição *para*. De fato, em ambos os casos, os complementos podem ser interpretados como beneficiário.

(14) a. O Marco comprou um carro *ao filho/para o filho*.
b. Construí uma casa *aos meus pais/para os meus pais*.
c. A bruxa preparou uma armadilha à princesa/para a princesa.
(Gonçalves; Raposo, 2013, p.1176-1178)

No entanto, o sintagma preposicionado introduzido por *para* não se assemelha sintaticamente a um OI. Essa afirmação é evidenciada no contraste da realização pronominal dos argumentos dativos e argumentos oblíquos, ou seja, na distinção entre as formas clíticas, *lhe/lhes,* e as formas pronominais tônicas, *ele(s), ela(s),* introduzidas por *para.* Nas sentenças exemplificadas em (15a-b), a pronominalização do OI destinatário/beneficiário está restrita às formas clíticas dativas. Daí o estatuto agramatical da sentença (15c). Nesse contexto, os clíticos dativos são substituídos pelas formas pronominais tônicas, na pronominalização do complemento oblíquo.

(15) a. O Marco comprou um carro *ao filho*[4].
b. O Marco comprou-*lhe* um carro.
c. Marco comprou (**lhe*) um carro *para o filho / para ele*.
(Gonçalves; Raposo, 2013, p.1176-1178)

2.3. "Alternância beneficiária" na estrutura argumental dos ditransitivos no PE

Na subseção anterior, destacamos o fato de que, no PE, os verbos ditransitivos de transferência/movimento, verbos transitivos e verbos de criação realizam dois tipos dis-

[4] Gonçalves e Raposo (2013, p.1176) apontam que o OI tem aqui interpretação ambígua, pois o filho de Marco também poderia ser o vendedor do carro e, portanto, terá a leitura de fonte.

tintos de estrutura argumental, representando estratégias condicionadas pela relação direta ou indireta entre dois argumentos. Assim, com verbos como *dar*, por exemplo, *dar um presente*, *uma carta*, *dinheiro*, a relação entre o objeto direto (OD) e o OI não implica movimento espacial, mas transferência direta de posse. Nesses predicados ditransitivos, o argumento recipiente/destinatário é um dativo, não um complemento oblíquo (OBL). Por sua vez, a contraparte preposicionada implica um movimento espacial em direção a um recipiente [+humano], embora não envolva uma transferência direta de posse entre dois constituintes. Nessa estrutura, o argumento oblíquo é interpretado como beneficiário do evento expresso pelo verbo.

Tal dicotomia fica claramente evidenciada no contexto de verbos como *atribuir/devolver/distribuir/emprestar*, nos quais não se tem uma contraparte preposicionada. A única interpretação possível é a de transferência possessiva entre os dois objetos. Ou seja, o argumento recipiente/destinatário comporta-se como um OI dativo.

(16) O José atribuiu uma tarefa *aos alunos*/atribuiu-*lhes* uma tarefa.

O fato de os falantes portugueses distinguirem o OI e o OBL leva a diferentes interpretações das ditransitivas, nas quais estão em jogo as preposições *a* e *para*. Com verbos como *atribuir*, *devolver*, *distribuir*, *emprestar*, não há a contraparte preposicionada.

(17) a. O José atribuiu uma tarefa *aos alunos*/atribuiu-*lhes* uma tarefa.
b. *O José atribuiu uma tarefa *para os alunos.*

Por sua vez, verbos como *fornecer/oferecer* licenciam o que temos denominado de "alternância beneficiária", envolvendo uma estrutura dativa (18a) e uma estrutura oblíqua, preposicionada (18b). Na variante dativa, o OI é simultaneamente recipiente e beneficiário, enquanto na variante preposicionada, o argumento introduzido pela preposição *para* é interpretado como beneficiário do evento expresso pelo verbo. Nesses termos, a contraparte preposicionada implica uma interpretação, na qual a ajuda dada aos flagelados pode não ter sido direta.

(18) a. O José forneceu/ofereceu/ ajuda *aos flagelados da seca*/ofereceu-*lhes* ajuda.
b. O José forneceu/ofereceu ajuda *para os flagelados da seca*.

Em outras palavras, na gramática do PE, os diferentes usos ou ocorrências da preposição *para* são reconhecidos somente dentro do seu contexto/configuração gramatical. As diferentes configurações determinam o significado. Como discutiremos posteriormente, o uso inovador da preposição *para* no PB, para contemplar os dativos, será mediada pela organização gramatical da língua, ou seja, não se trata lexicalmente de uma outra preposição *para.*

3. A expansão da preposição *para* nas ditransitivas do PB

3.1 O percurso diacrônico das inovações

Estudos quantitativos sincrônicos e diacrônicos, em particular, dentro do cenário comparativo com o português europeu (PE), revelam que, na identificação do objeto indireto, os clíticos dativos de 3ª pessoa, *lhe(s)*, e os sintagmas nominais introduzidos por *a* – um marcador dativo – não mais fazem parte da gramática nuclear do PB contemporâneo[5] (cf. Gomes, 2003; Freire, 2005; Torres Morais; Berlinck, 2006, 2007, 2018; Torres Morais e Salles, 2010; Salles; Torres Morais, 2020; Ramos; Duarte, 2007; Figueiredo Silva, 2007; Armelin, 2011, Calindro, 2015, 2016, 2020, entre muitos outros, sob diferentes perspectivas teóricas). Os clíticos dativos *lhe/lhes* são substituídos pelas estratégias pronominais oblíquas *a/para ele(s); a/para ela(s)*. Por sua vez, ocorre a generalização da preposição *para* como introdutora do OI nominal nas estruturas ditransitivas.

Torres Morais e Berlinck (2018), retomando estudos anteriores, apresentam um panorama da expansão das formas inovadoras na expressão do argumento dativo/objeto indireto (OI), nas construções ditransitivas, com um vasto e diversificado material histórico, a partir do século XVIII, coletado no *Projeto para a História do Português Paulista* (PHPP)[6].

No que se refere à expressão pronominal do OI, o que as autoras destacam, a partir dos dados coletados, é que o clítico dativo constitui a estratégia quase exclusiva até o final do século XIX, independente da pessoa gramatical. Fica evidente pelos resultados quantitativos obtidos nas primeiras décadas do século XX que o processo que deu origem ao quadro atual (com o predomínio do OI oblíquo, introduzido por preposição, sobretudo na 3ª pessoa) deve ter se manifestado tardiamente.

[5] Lembramos que o qualificativo *nuclear* se refere à gramática adquirida pela criança na aquisição da língua materna, ou Língua -I(nterna), na perspectiva gerativista. Os clíticos dativos na referência de 3ª pessoa, mesmo sendo adquiridos ao longo da escolaridade, são muito pouco usados na fala e escrita, excetuando-se as situações de maior monitoramento e nos textos literários.

[6] Projeto que faz parte do *Projeto para a História do Português Brasileiro* (https://sites.google.com/site/corporaphpb/home).

O mesmo percurso histórico pode ser desenhado na expansão das preposições *a* e *para* nas estruturas ditransitivas com verbos de transferência/movimento, ao longo dos séculos XVIII e XIX. Dados obtidos em anúncios de jornais da primeira metade do século XIX, por exemplo, revelam que a preposição *a* se associava predominantemente com o OI interpretado como recipiente/destinatário (19) (cf. Torres Morais; Berlinck, 2018). Por seu lado, as ocorrências de *para* se restringem a introduzir argumentos interpretados como beneficiário. Ademais, no mesmo conjunto de verbos, a introdução do argumento locativo se mostrou um contexto de predomínio da preposição *para* (20)[7].

(19) [19,1 A SP] Foi remettido *a um senhor Negociante desta Cidade* um caxóte de livros…

(20) 19,1 A SP] Remette-se *para qualquer ponto de Estrada de ferro* Cal, cimento, telhas...

(Torres Morais; Berlinck, 2018, p.279)

Finalmente, consideremos o fato de que, na diacronia do PB, constatou-se o uso da preposição *a* na introdução de argumentos locativos, configurando o seu estatuto ambíguo como marcador de caso dativo e preposição plena. No exemplo (19), referente à segunda metade do século XIX, temos a preposição *a* introduzindo um argumento locativo (Taubaté) e um OI (Antonio Joaquim da Silva) .

(21) [19,2 A SP] Quem o prender e levar *a* Taubaté á Antonio Joaquim da Silva ou á João Evangelista da Costa Marcondes, será generosamente gratificado. Taubaté, 6 de Outubro de 1879. Antonio Joaquim da Silva.

(Torres Morais; Berlinck, 2018, p.279)

Tais resultados empíricos aqui comentados revelaram, através dos dados coletados ao longo século XX (Calindro, 2015), que a mudança que se implementa na gramática brasileira é a de uma estrutura ditransitiva preposicionada, na qual o argumento recipiente/beneficiário é morfossintaticamente um argumento oblíquo. Note-se que há uma distinção entre formas pronominais e nominais na realização do OI: as formas pronominais *ele(s), ela(s)*s são introduzidas por *a* e *para.* As formas nominais são majoritariamente

[7] Um ponto a ser considerado no percurso de implementação da preposição *para* é o de que ele se inicia em contextos mais concretos e se difunde para contextos mais abstratos (com verbos de transferência verbal/perceptual e OD [abstrato]), por exemplo, ao longo do século XX. (cf. Gomes, 2003; Torres Morais; Berlinck, 2007, 2018).

introduzidas por *para* no contexto dos verbos ditransitivos de transferência material (22) e verbos de movimento (23)[8]. Entretanto, com verbos de criação e verbos transitivos a preposição *para* é categórica (24) (cf. Armelin, 2011).

(22) a. O João deu um colar de pérolas *para a Maria*
b. O João deu um colar de pérolas *para ela.*
(23) a. O João enviou/ entregou uma carta *para a Maria.*
b. O João enviou/ entregou uma carta *para ela.*
(24) a. A Maria preparou um jantar *para o João.*
b. A Maria preparou um jantar *para ele.*

Ao incorporar dados de língua escrita, em diferentes tipos de textos, o percurso diacrônico que discutimos nesta seção é relevante por nos permitir estabelecer um aspecto comparativo entre o PB e o PE, no que se refere à expressão do OI nas ditransitivas. O PE conserva historicamente as propriedades morfossintáticas e semânticas que distinguem o OI e o OBL. Na diacronia do PB, porém, inovações no sistema pronominal e preposicional resultam na perda das propriedades identificadoras do OI dativo.

3.2 Conclusões parciais

Na seção 2 deste estudo, detalhamos as propriedades morfossintáticas que identificam o OI dativo, introduzido pela preposição *a,* semanticamente interpretado como recipiente/destinatário nas ditransitivas de objeto duplo do PE.

Observamos ainda uma contraparte preposicionada, configurando uma "alternância beneficiária", na qual o argumento introduzido por *para* é interpretado como beneficiário, e apresenta comportamento de complemento oblíquo (OBL). Nesse contexto, não se infere a transferência possessiva direta entre os dois complementos da estrutura ditransitiva. Destacamos ainda o fato de que, na gramática portuguesa, a preposição *para* introduz tanto o complemento oblíquo beneficiário, como um locativo "puro".

Por sua vez, a seção 3 descreve as inovações implementadas na diacronia do PB, envolvendo a perda da preposição *a* como marcador dativo (*a*-DP) e sua substituição generalizada pela preposição *para,* assim como a perda dos clíticos dativos de 3ª pessoa *lhe(s)*, sendo substituídos pelas formas pronominais plenas, *ele(s), ela(s),* introduzidas por *a/para*.

[8] Durante a escolarização, a preposição *a* é associada ao OI e passa a ser utilizada pelos falantes escolarizados na fala monitorada e na escrita formal e literária. No entanto, a nossa hipótese é de que não se trata mais de uma preposição dativa, mas de uma preposição oblíqua (cf. Calindro, 2015).

A hipótese que assumimos, e que será incorporada na nossa análise formal, é a de que os falantes adquirindo o português no Brasil dos primeiros séculos entravam em contato com a preposição *a*, na sua natureza polissêmica, como preposição gramatical e lexical, neste último caso, introduzindo locativos puros. No entanto, apesar do *input,* os falantes desenvolveram na sua fala novas estratégias na expressão do OI, ou seja, não mais reconhecem a preposição *a* como marca gramatical identificadora do OI, associada à pronominalização necessária com os pronomes dativos, *lhe(s)*[9]. Ao mesmo tempo, expressam uma generalização no uso da preposição *para,* com significado de beneficiário, para os significados de recipiente/destinatário no contexto das sentenças ditransitivas. Essa generalização também afeta os locativos puros, nos quais a preposição *para* alterna com a preposição *a* no PB.

Por fim, é importante destacar o papel da queda do clítico dativo *lhe(s)* de terceira pessoa no PB e sua substituição pelos pronominais tônico, *ele(s);ela(s)* introduzidos por *a* e *para* como ilustrado nos exemplos (22-24). Note-se que sentenças em que este tipo de uso pronominal ocorre, sem o redobro de um clítico e introduzindo um OI recipiente, são agramaticais em PE (25a-b).

(25) a. O João enviou *enviou-lhe* uma carta.
b. * O João enviou uma carta *a ela*.

Segundo Torres Morais e Berlinck (2018), a substituição do clítico *lhe(s)* pelo sintagma preposicionado a/*para* + *ele(a)(s)* representou o contexto inicial de mudança na expressão do OI, coletado no corpus do PB.

4. Reanálise da expressão do OI em PB na perspectiva do modelo Maximize Meios Mínimos (MMM)

4.1. Introduzindo o modelo MMM

Nesta seção, trazemos uma discussão mais detalhada da aquisição e mudança sintática do *Modelo Maximize Meios Mínimos* (MMM), o qual se insere numa perspectiva da Faculdade da Linguagem, proposta no modelo dos Três Fatores de Chomsky (2005, p.6), como expresso em (26):

[9] Muito importante lembrar que a perda da preposição *a* como marcador casual do OI abrange contextos verbais muito mais amplos, alterando inclusive o seu estatuto funcional de OI para OD. Verbos psicológicos são um exemplo que podem ilustrar este ponto: (i) O José agradeceu *o professor/ao professor*; (ii) O filho deve obedecer *aos pais/os pais*; (iii) O jantar agradou/desagradou *aos convidados/os convidados.*

(26) 1. Gramática Universal (GU) + 2. Dados Linguísticos Primários (DLP) + 3. Princípios Cognitivos Gerais → Gramática do Adulto (= uma Língua-I)

Segundo Chomsky, a aquisição de parâmetros evidentes na gramática dos adultos não é moldada apenas pela GU, mas através da interação com os fator 2 - dados linguísticos primários e com o fator 3 que inclui dois tipos de princípios cognitivos gerais:

I. princípios de análise dos dados, que devem ser usados na aquisição da linguagem e em outros domínios;

II. princípios de arquitetura estrutural e restrições de desenvolvimento, incluindo-se princípios de computação eficiente.

Os três fatores são posteriormente trabalhados no modelo MMM proposto por Biberauer (2018, 2019a e 2019b), com base nas seguintes questões: (i) qual seria a contribuição da GU para direcionar a criança na aquisição da Língua-I, sua gramática internalizada? (ii) quais aspectos dos DLP constituem dados cruciais para a construção da gramática na fase da aquisição? (iii) Como caracterizar os princípios gerais de computação e cognição? Como a autora declara " (...) o modelo MMM foi concebido como um modelo minimalista que explicitamente busca entender a variação linguística – e também o processo de aquisição que deu origem a essa variação – como consequência da interação entre os Três Fatores delineados em Chomsky (2005)" (Biberauer, 2019b, p.213)[10].

O modelo se configura, portanto, no seguinte esquema:

(27) GU + Input + Maximise Minimal Means (MMM) → Gramática do Adulto
Biberauer (2019b, p.213)

Com relação ao Fator 1, Biberauer afirma que o modelo MMM pretende explicar os processos de aquisição, variação e mudança, ao assumir uma GU *minimamente especificada*. A hipótese é a de que a GU não desenha um mapa genético de princípios e parâmetros que guiam a criança na formatação da gramática materna. Em seu lugar são os mecanismos cognitivos gerais, incluindo vieses de aprendizagem geral, que terão o papel central em ativar a aquisição e formatar as gramáticas adultas.

[10] Tradução nossa: "… the MMM model was conceived as a minimalist model which explicitly seeks to understand language variation – and thus also the acquisition process that gave rise to that variation – as the consequence of the interaction of all Three Factors outlined in Chomsky (2005)."

Com relação ao Fator 2, a autora assume que os DLP não contemplam a totalidade dos dados do *input* recebidos pelo falante. Essa parte do *input*, denominada na literatura corrente de *intake*, se refere apenas aos dados *accessíveis* à criança, de modo que o processo de aquisição decorre de uma identificação de traços morfossintáticos dos DLP regulados pelo *intake*. Ou seja, os DLP são constituídos por certos aspectos do *input* destacados por terem uma importância particular, tais como *redobro morfológico*, *variação na ordem de palavras*, e *funções múltiplas* (cf. Biberauer 2019a, p.54-59 para uma discussão mais detalhada).

Finalmente, em relação ao Fator 3, referentes aos princípios cognitivos gerais, o modelo MMM propõe viéses de aprendizagem com um papel central na aquisição e formatação das gramáticas adultas atuando tanto para o processo de aquisição como para o processo de mudança linguística. O esquema em (28b) ilustra essa reinterpretação de Biberauer para o modelo dos Três Fatores de Chomsky em (28a):

(28) a. GU + DLP + Fatores Cognitivos Gerais→ Gramática do Adulto
b. GU + DLP + Maximise Meios Mínimos (MMM) → Gramática do Adulto

Vale ressaltar que o modelo MMM apresenta os seguintes reflexos linguísticos na aquisição (Biberauer, 2019b, p. 212,213).

(29) Economia dos traços – *Feature Economy* (FE) (cf. Roberts e Roussou, 2003)
Postule o menor número possível de traços formais na abordagem do *input/intake*.

(30) Generalização do Input – *Input Generalization* (IG) (cf. Roberts, 2007)
Aproveite ao máximo os traços disponíveis na abordagem do *input/intake*.

Portanto, no processo de aquisição regulado pelo modelo MMM, a fixação dos parâmetros é engatilhada pelo *intake*. E, ao postular novos traços formais, as crianças priorizam a reutilização de traços já postulados, no lugar de novas elaborações. Esses traços gramaticais serão, portanto, refinados por novos *intakes* até que uma realidade sistemática seja alcançada.

Este é um processo que pode ser entendido como uma forma de *sobregeneralização* (*overgeneralization*), podendo levar a criança a produzir sentenças ou formar palavras que

não são gramaticais nos dados do *input*. Finalmente, se a criança não se retrata, a variação e mudança se instalam. Ou seja, para o processo de aquisição de primeira língua, quando há duas formas, formas dublês nos termos de Kroch (2001), se a criança não retrai para a forma mais antiga, e passa somente a usar a nova variante, um novo postulado é feito (Biberauer; Roberts 2017; Biberauer, 2019a).

Nas reflexões sobre o *input*, seguindo o Princípio de Tolerância de Yang (2018), Biberauer (2018, p.216) afirma que menos vocabulário faz com que a criança tenha mais facilidade para adquirir regras. A autora argumenta que, no processo de aquisição, 'menos é mais' (Newport, 1990).

Nesses termos, o modelo MMM prevê que as crianças tirem o maior proveito do *input* ao qual têm acesso. O sistema dá pistas para elas selecionarem blocos de acordo com o *input* que recebem, muito antes de ter um conhecimento lexical complexo. O acesso a apenas uma parte do *input*, ou seja, somente ao *intake,* faz com que a criança seja capaz de se focar em sub-componentes linguísticos.

Um outro aspecto interessante se refere à ação de um Princípio de otimização ligado ao Fator 3, denominado *Condição dos Subconjuntos,* o qual é impulsionado pela *Economia de Traços* e pela *Generalização do Input*. A sua atuação é evidenciada por estudos de aquisição, nos quais as crianças começam com grandes categorias e tem que recuar para categorias menores (Biberauer; Roberts, 2017). Consequentemente, ao ampliar o aprendizado para outros domínios apropriados, o falante maximiza o uso dos traços já disponíveis.

A seguir discutiremos a nossa hipótese de que a mudança na expressão do OI nas ditransitivas no PB expressa um processo de *sobregeneralização* do *input* no uso da preposição *para*. Lembrando que essa hipótese leva a uma outra discussão: por que teria ocorrido este processo ocorrido, diante do uso multifuncional da preposição *a*?

4.2 MMM e a expansão da preposição para nas ditransitivas do PB

Nas nossas reflexões a seguir teremos como objetivo lidar com uma questão que se tornou inevitável, uma vez assumida a perspectiva gerativista de que a variação e a mudança ocorrem na fase da aquisição da linguagem: o que teria ativado a criança brasileira, na fase da aquisição, a desconsiderar, nos dados do *input,* a preposição *a* como marcadora gramatical de caso dativo? Relacionada a esta questão, há uma outra: O que teria levado a criança a generalizar as formas pronominais oblíquas de 3ª pessoa, em detrimento dos clíticos dativos?

Para respondê-las, consideraremos a hipótese de que as formas inovadoras na expressão do OI no PB contemporâneo resultaram de um efeito linguístico de *generalização*

do input na fase da aquisição da linguagem, com base na estrutura argumental dos verbos de criação, entre eles, *bater* (um bolo); *construir; pintar, preparar, desenhar* e verbos transitivos, entre eles, *comprar, vender ler, escrever,* incorporando ainda os verbos ditransitivos de transferência/movimento, entre eles*, dar, enviar, mandar*.

Como vimos na seção 2, verifica-se no PE contemporâneo uma "alternância beneficiária", manifestada nos contextos verbais licenciadores de um argumento morfossintaticamente identificado como OI dativo e uma contraparte preposicionada, na qual a preposição *para* serve para expressar o beneficiário.

Os exemplos (31a), (32a) e (33a) ilustram essa dinâmica das gramáticas do PE e PB, uma vez que tal restrição interpretativa imposta gramaticalmente à preposição *para* no PE não se manifesta no PB.

(31) a. A Maria preparou um elegante jantar *para o João / para ele.* (PE; PB).
b. A Maria preparou um elegante jantar *ao João/* preparou-*lhe* um elegante jantar. (PE)

(32) a. O João comprou um colar de pérolas *para a Maria/ para ela.* (PE; PB)
b. O João comprou um colar de pérolas à Maria/comprou-*lhe* um colar de pérolas. (PE)

(33) a. O João enviou uma carta *para o Pedro.* (PE; PB)
b. O João enviou uma carta *ao Pedro.* (PE)
c. O João enviou *enviou-lhe* uma carta. (PE)
d. O João enviou uma carta *para/a ele.* (PB)

Tais desenvolvimentos nos levam a considerar que, na aquisição do PB, os falantes não teriam reconhecido como *input* linguístico a preposição marcadora de caso dativo *a* alternando com os clíticos dativos *lhe(s)* e introduzindo essencialmente o argumento recipiente, semanticamente associado ao traço [+humano]. Simultaneamente, teria se efetivado o processo de sobregeneralização (*overgeneralization*) da preposição *para,* a partir do argumento beneficiário para o argumento recipiente em sentenças ditransitivas, nas quais se expressa a transferência direta de posse entre os dois argumentos: o objeto direto (OD-tema) e o OI recipiente, ilustrado em (26) com o verbo *dar.* Observe-se ainda que a contraparte pronominal licencia a preposição *para* nas formas *para ele(s), para ela(s).*

(34) O João deu um colar de pérolas *para a Maria/para ela.*

Observe-se que o mesmo cenário diacrônico de variação e mudança se manifesta na queda dos clíticos dativos de 3ª pessoa *lhe(s),* decorrentes de sua substituição pelos pronomes plenos, cuja expansão no PB tem sido amplamente documentada nos resultados quantitativos. Lembramos que este tipo de perífrase fora do contexto do redobro do clítico é agramatical em PE (cf. 8b).

(35) a. O João *enviou-lhe* uma carta. (PE)
b. O João enviou uma carta *para/a ele.* (PB)

Em outras palavras, a mudança que se consolidou na gramática do PB, desconsiderando o *input* dativo do PE, pode ser definida nos seguintes termos: a preposição *para* se generalizou para os demais contextos ditransitivos, relacionando dois argumentos DPs, como veremos com detalhes na seção 4.2.1, a proposta de análise para a nova gramática supõe um sintagma preposicionado em camadas (pP seguindo Wood, 2012). Na camada interna, o DP introduzido pela preposição é licenciado com caso oblíquo e interpretado como destinatário/beneficiário, ambos condicionados pelo traço [+animado] do referente.

Dessa forma, da perspectiva do modelo MMM, esse desenvolvimento e substituição da preposição *a* por *para* em PB pode ser entendido como a consequência da *Generalização do Input*: a presença de um item multifuncional como o *a*, levou à sua substituição por diversos itens como *de*, *em*, além do *para*. Como mencionado anteriormente, através da *Condição de Subconjuntos,* que atua na otimização do sistema linguístico, uma vez que a preposição *a* tinha um estatuto ambíguo por ser responsável pela introdução de elementos com múltiplas leituras semânticas, o sistema generalizou outras preposições presentes no sistema, criando subconjuntos. Neste texto, nosso foco é a preposição *para*, cujo uso foi ampliado para outros contextos dentro de um domínio do qual já fazia parte. Esse padrão ampliado também denota a perda do clítico [+humano] *lhe*, que foi consistentemente substituído por estruturas em que a preposição lexical *para* é acompanhada por um pronome pleno. A perda da contraparte clítica associada ao traço [+humano] também pode ser vista como uma simplificação gramatical, pois menos micro-parâmetros (cf. Roberts, 2012) precisam ser adquiridos pelo falante durante o processo de aquisição.

4.2.1 Representação das estruturas ditransitivas no PB

Neste ponto de nossa discussão, vale destacar que as inovações na identificação morfossintática do OI, as quais se verificam na diacronia e sincronia do PB, não se manifestam nas demais línguas românicas. Partindo dessa constatação, alguns autores, assumiram a proposta do núcleo aplicativo baixo (ApplP) (cf. Pylkkänen, 2002), o qual introduz o OI, nominal e pronominal, nas estruturas ditransitivas, atribuindo a ele Caso inerente dativo, associado ao papel temático recipiente/destinatário. Segundo essa análise, o núcleo aplicativo denota semanticamente uma relação direta de posse entre o OD e o OI [+animado]. Nesse contexto, a preposição *a* não é uma preposição lexical, mas um marcador de Caso dativo, como mencionado anteriormente. Com relação ao PE, a estrutura aplicativa (36) foi assumida por alguns autores, entre eles, Torres Morais (2007); Torres Morais e Salles (2010):

(36) $[_{vP}$ A Maria $[_{v'}v$ $[_{VP}$ enviou $[_{ApplP}$ ao João/lhe $[_{APPL'}$ Ø $[_{DP}$ uma mensagem]]]]]]]

Já em relação à preposição *para*, pode-se afirmar que os seus diferentes usos e ocorrências são reconhecidos somente dentro do seu contexto gramatical, pois as diferentes configurações determinam o significado da forma. Nessa perspectiva, o uso inovador da preposição *para* em PB é mediado pela organização gramatical da língua, ou seja, não se trata de uma outra preposição *para*, diferente da que faz parte do PE. A mudança que se consolidou na gramática do PB, que desconsiderou o *input* dativo do PE, pode ser definida nos seguintes termos: a preposição *para* se generalizou dos contextos de beneficário/locativo para os contextos ditransitivos, sendo responsável, portanto, por relacionar os dois argumentos presentes nessas construções.

Esse processo pode ser explicado pelo modelo da MMM da seguinte forma: uma vez que o aprendiz da L1 (das novas gerações) tenha evidência suficiente por meio do *intake* de que o elemento *para* também pode ser usado nos contextos de recipiente/destinatário, acrescido do fato de que não há evidências morfofonológicas de ApplP (como no PE), devido à perda da expressão morfológica do ApplP na forma do clítico dativo *lhe,* como expresso na representação em (36), o mecanismo da *Economia de Traços* considera não apenas o traço lexical de *para*, mas também a ordem V ACC PP, além do traço do item categorial de uma projeção funcional encabeçada pela preposição (pP, nos termos de Wood, 2012) para introduzir o OI em uma construção ditransitiva preposicionada, em que é um complemento oblíquo, conforme explica Calindro (2020). De acordo com

a autora, os OIs introduzidos por *para* em PB são realizados na sua estrutura argumental através de uma projeção pP, não por um ApplP, como no PE.

Na proposta de um pP em camadas, na camada interna PP, o DP selecionado pela preposição é licenciado com Caso oblíquo inerente, associado aos papéis temáticos de recipiente/destinatário e beneficiário, além de expressar uma condição de referencialidade do DP, ou seja, o seu traço [+animado]. O mesmo se verifica com a 3ª pessoa que se realiza na sua forma forte introduzida pela preposição.

(37) a. A Maria enviou uma mensagem *para mim, para você, para ele.*

O uso de *para* no contexto de verbos de transferência/movimento é, portanto, mediado pela organização gramatical da língua. Uma vez que o PB perdeu a preposição a marcadora de caso dativo responsável por introduzir o elemento dativo no ApplP em PE e em outras línguas românicas, esse núcleo funcional foi reanalisado em PB. Segundo Calindro (2020), uma projeção funcional pP em camadas introduz a relação entre o OD-tema e o OI recipiente/destinatário/beneficiário licenciado no PP pela preposição transitiva *para*, como verificamos nas representação a seguir:

(38) $[_{vP}$ A Maria $[_{v'v}$ $[_{VP}$ enviou $[_{pP}$ uma mensagem $[_{p'}$ Ø $[_{PP}$ *para* o João]]]]]]]

5. Considerações Finais

No presente artigo, vimos que em PB a realização do objeto indireto pronominal de 3ª pessoa, interpretado como recipiente/beneficiário nas estruturas ditransitivas, apresenta dois aspectos inovadores: (i) a substituição da preposição *a* por *para*; (ii) a substituição do clítico dativo *lhe(s)* pelas perífrases preposicionadas *a/para ele(a)(s)*. Como discutido ao longo do texto, não se trata de uma inovação trivial quando se pensa que no português histórico/clássico e no PE contemporâneo, em que a identificação do OI envolve, obrigatoriamente, o uso da preposição *a* na introdução dos OIs recipientes, e a pronominalização por *lhe(s)*. Em paralelo à presença da preposição *a* na introdução de locativos puros e metafóricos, assim como da preposição *para* na introdução de OIs beneficiários e locativos puros.

Após a discussão dos dados do PE e do PB, apresentamos nossa proposta que se apoia numa abordagem minimalista formal da aquisição, variação e mudança gramatical, delineada no modelo dos Três Fatores (Chomsky, 2005) e reinterpretada dentro do

Modelo de Maximizar Meios Máximos, proposto por Biberaruer (2018, 2019a, 2019b), diretamente relacionado ao Fator 3. As principais manifestações do Fator 3 relevantes para a nossa discussão são a *Economia dos traços* e a *Generalização do Input*, pois assumimos que a interação do falante com o *intake* recebido durante a fase de aquisição dá origem à variação e mudança linguística.

Dessa forma, discutimos neste trabalho que a presença da preposição *para* nos dados recebidos pelos falantes do que viria a ser o PB se deu através de generalizações e uso econômico de traços formais, o que leva à utilização ao máximo de recursos mínimos. Esse fato acarretou, portanto, na expansão dos usos de *para* nos contextos em que somente o *a* é aceito em PE e era aceito no português clássico que aportou no Brasil. Dessa forma, a preposição *para* em PB passa a introduzir, além de locativos puros e beneficiários, OIs recipiente/destinatário, no lugar da preposição *a*.

Referências

ARAÚJO, F.; CARVALHO, H. Te e Lhe como clíticos acusativos de 2ª pessoa em cartas pessoais cearenses. *Labor Histórico*, Rio de Janeiro, 1 (1): 62-80. 2015. http://dx.doi.org/10.17074/lh.v1i1.176

ARMELIN, P. *Sentenças Bitransitivas do Português Brasileiro revisitadas à luz da teoria de núcleos funcionais aplicativos.* Dissertação de Mestrado. Universidade de São Paulo, São Paulo, 2011.

BIBERAUER, T.; ROBERTS, I. Parameter setting. *In:* LEDGEWAY, A.; ROBERTS, I. (eds.), *The Cambridge handbook of historical syntax*, 134–162. Cambridge: CUP. 2017.

BIBERAUER, T. Less IS more: Some thoughts on the Tolerance Principle in the context of the Maximise Minimal Means model. *Cambridge Occasional Papers in Linguistics* 11, 6, p.131–145. 2018.
BIBERAUER, T. Factors 2 and 3: Towards a principled explanation. *In:* GALLEGO, A.; OTT, D. (eds.), *Catalan Journal of Linguistics* (Special issue: Generative Syntax. Questions, Crossroads, and Challenges), p.45-88. 2019a.
BIBERAUER, T. Children always go beyond the input: The Maximise Minimal Means perspective. *Theoretical Linguistics 45* (3-4), p. 211-224. 2019b.

BRITO, A. M. Grammar variation in the expression of verb arguments: the case of the Portuguese Indirect Object. *Phrasis*, 49 (2), p.31–58. 2008.

BRITO, A. M. Two base generated structures for ditransitives in European Portuguese. Simões, Barreiro, Santos, Sousa-Silva & Tagnin (Ed.) *Linguística, Informática e Tradução: Mundos que se Cruzam*, Oslo Studies in Language 7(1), 2015. 337–357.

CALINDRO, A. Introduzindo Argumentos: uma proposta para as sentenças ditransitivas do português brasileiro. Tese de Doutorado. Universidade de São Paulo, São Paulo, 2015.

CALINDRO, A. Introducing indirect arguments – the locus of a diachronic change**.** *Rivista di Grammatica Generativa*. Perugia, Itália, v.38, 2016.

CALINDRO, A. Ditransitive constructions: what sets Brazilian Portuguese apart from other Romance languages? *In* PINEDA, A.; MATEU, J. *In Dative constructions in Romance and Beyond.* Open Generative Syntax Series, Language Science Press, 2020.

CALINDRO, A.; ALVES, M.; MARTINS, A. Colocação pronominal em português: uma proposta para um contínuo luso-afro-brasileiro. *Matraga*, v.31, p.280-302. 2024

CHOMSKY, N. *Lectures on government and binding*. Dordrecht. Foris, 1981.

CHOMSKY, N. Three factors in Language Design. *Linguistic Inquiry* 36, p. 1-22. 2005.

CUERVO, C. *Datives at large*. PhD. Dissertation. MIT, 2003.

DIACONESCU, C.; RIVERO, M. L. An applicative analysis of double object constructions in Romanian. *Probus*, 19, p.171–195. 2007.

DUARTE, I. A construção de topicalização na gramática do português: regência, ligação e condições sobre movimento. Ph. D. Dissertation, Universidade de Lisboa, 1987.

DUARTE, I. Relações gramaticais, esquemas relacionais e ordem de palavras. *In:* Mateus et alii (2003): Gramática da Língua Portuguesa. Lisboa: Caminho, 2003, p. 275-321.

FIGUEIREDO SILVA, M. A perda do marcador dativo e algumas das suas consequências". *In:* CASTILHO, A.; TORRES MORAIS, M.; LOPES, R.; CYRINO, S. (Org.). *Descrição, história e aquisição do português brasileiro – Estudos dedicados a Mary Aizawa Kato.* São Paulo: Fapesp, Campinas: Pontes Editores, 2007, p. 85-110.

FREIRE, G. *Os clíticos de terceira pessoa e as estratégias para sua substituição na fala culta brasileira e lusitana*. Dissertação de Mestrado. UFRJ, 2000.

GALVES, C. O enfraquecimento da concordância no português brasileiro" *In*: KATO, M.; ROBERTS, I. (Org.) *Português brasileiro — uma viagem diacrônica*. Campinas, Ed. UNICAMP, 1993.

GALVES, C.; TORRES MORAIS, M; RIBEIRO, I. A. Syntax and morphology in the placement of clitics in European and Brazilian Portuguese. Journal of Portuguese Linguistics, Lisboa, v. 4, n.2, 2005, p. 143-177.

GALVES, C. A língua das caravelas: periodização do português europeu e origem do português brasileiro. *In:* CASTILHO, A.; TORRES MORAIS, M. A.; LOPES, R.; CYRINO, S. (eds.), *Descrição, história e aquisição do português brasileiro*, 513- 52. Campinas: Ed. Pontes, 2007, p.513-52.

GALVES, C. Mudança sintática no português brasileiro. *Cuadernos da Alfal*, n.12(2), p.17-43. 2020.

GALVES, C.; KATO, M. ; ROBERTS, I. Português Brasileiro: uma segunda viagem diacrônica. Campinas, SP: Editora da Unicamp, 2019.

GOMES, C. Variação e mudança na expressão do dativo no português brasileiro. *In:* PAIVA, M. C.; DUARTE, M. E. L. (eds.): Mudança Linguística em Tempo Real, Rio de Janeiro: FAPERJ, Contra Capa, 2003, p.81-96.

GONÇALVES, A.; RAPOSO, E. Verbo e sintagma verbal. *In:* RAPOSO, E. et alii. *Gramática do português.* Fundação Calouste Gulbenkian. Vol. II. Cap. 28, 2013, p. 1155-1218.

KATO, M.; MARTINS, A. M.; NUNES, J. Português brasileiro e português europeu: sintaxe comparada. São Paulo: Contexto, 2023.

KROCH, A. Syntactic Change. *In:* Mark Baltin e Chris Collins (Ed.) *The Handbook of Contemporary Syntactic Theory*. Blackwell, 2001.

LONGOBARDI, G. 2018. Principles, Parameters, and Schemata: a radically underspecified UG. *Linguistic Analysis* 41(3-4), p. 517-558. 2018.

LOPES, C. Pronomes pessoais. *In:* BRANDÃO, S.; VIEIRA, S. (Org.). *Ensino de gramática: descrição e uso*. 1 ed. São Paulo: Contexto, 2007, v. 1, p. 103-114.

LOPES, C. ; RUMEU, M. C. O quadro de pronomes pessoais do português: as mudanças na especificação dos traços intrínsecos. *In:* CASTILHO, A.; TORRES MORAIS, M. ; LOPES, R.; CYRINO, S. (Org.). *Descrição, história e aquisição do português brasileiro – Estudos dedicados a Mary Aizawa Kato.* São Paulo: Fapesp, Campinas: Pontes Editores, 2007, p.419-436.

MARTINS, A. M. *Clíticos na história do português*. Tese de Doutorado – Universidade de Lisboa, 1994.

MARTINS, A. M. Mudança sintáctica e História da Língua Portuguesa. In: B. F. Head, J. Teixeira, A. S. Lemos, A. L. de Barros e. Pereira, eds. *História da Língua e História da Gramática: Actas do Encontro*, Braga: Universidade do Minho /ILCH. 251-297, 2002.

MARTINS, A. M.Introdução: O português numa perspetiva diacrónica e comparativa. In: A. M. Martins e E. Carrilho, eds. *Manual de Linguística Portuguesa*. Berlin/Boston: De Gruyter. 3-39, 2016

MARTINS, M. A.; MOURA, K.; COSTA DA SILVA, F. Análise diatópico-diacrônica dos complementos pronominais de verbos na escrita brasileira dos séculos XIX e XX. *Working papers em linguística* (online), v. 20, n. 2, p. 195-216. 2019.

MORAES DE CASTILHO, C. M. (2001). Seria quatrocentista o português implantado no Brasil? Estruturas sintáticas duplicadas em textos portugueses do século XV. In R. V. Mattos e Silva (Org.), *Para a história do Português Brasileiro* (p. 57-89), São Paulo, SP: Humanitas / Fapesp, p. 57-89, 2001.

NEWPORT, E. Maturational constraints on language learning. *Cognitive Science* 14, p.11-28. 1990.

KEWITZ, V; PAIXÃO DE SOUSA, M. C. Vésperas Brasilianas: uma agenda para os estudos sintáticos do Português Brasileiro nos primeiros séculos. Revista Portuguesa de Humanidades, v. 15, p. 65-90, 2011.

PYLKKÄNEN, L. *Introducing Arguments*. PhD . Dissertation. MIT, 2002.

RAMOS, J. ; DUARTE, F. Ordem de constituintes, conteúdo de traços-phi e mudança gramatical no PB. *In:* CASTILHO, A.; TORRES MORAIS, M. ; LOPES, R.; CYRINO, S. (Org.). *Descrição, história e aquisição do português brasileiro – Estudos dedicados a Mary Aizawa Kato.* São Paulo: Fapesp, Campinas: Pontes Editores, 2007, p. 111-130.

RIBEIRO, I. A mudança sintática do português brasileiro é mudança em relação a que gramática? *In:* CASTILHO, A. (org.) *Para a história do português brasileiro: primeiras idéias*, São Paulo: Humanitas, 1998, p. 101-120.

ROBERTS, I.; ROUSSEAU, A. *Syntactic change: a Minimalist Approach to Grammaticalization*. Cambridge. Cambridge University Press, 2003.
ROBERTS, I. *Diachronic Syntax.* New York. Oxford University Press, 2007.

ROBERTS, I. On the nature of syntactic parameters: a programme for research. In: GALVES, C.; CYRINO, S.; LOPES, R.; SANDALO, F.; AVELAR, J. (eds.). *Parameter Theory and Linguistic Change*, 319-334. Oxford: Oxford University Press, 2012.

TORRES MORAIS, M. A. *Dativos*. Tese de Livre Docência. São Paulo, SP. Universidade de São Paulo, 2007.

TORRES MORAIS, M. A.; BERLINCK, R. A caracterização do objeto indireto no português: aspectos sincrônicos e diacrônicos *In:* LOBO T; RIBEIRO, I; ALMEIDA, Z. *Novos dados, novas análises***.** Vol. VI. Tomo I. Salvador. EDUFBA, 2006. p 73-106.

TORRES MORAIS, Maria Aparecida; BERLINCK, Rosane. 'Eu disse pra ele' ou 'Disse-lhe a ele': A expressão do dativo nas variedades brasileira e europeia do português. *In:* CASTILHO, A. *et al. Descrição, história e aquisição do português brasileiro*. Campinas, Ed. Pontes, 2007.

TORRES MORAIS, M. A.; BERLINCK, R. O objeto indireto: argumentos aplicados e preposicionado. *In:* CYRINO, S.; TORRES MORAIS, M. A. (Org.). *Mudança sintática do português brasileiro: perspectiva gerativista.* 1ed. São Paulo: Contexto, 2018. p. 252-307.

TORRES MORAIS, M. A.; SALLES, H. Parametric change in the grammatical encoding of indirect objects in Brazilian Portuguese. *Probus*, v.22, 2010.

SALES, I. *Aspectos linguísticos e sociais no uso de pronomes em cartas pessoais baianas*. Dissertação (Mestrado em Língua Portuguesa). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2007.

SALLES, H.; TORRES MORAIS, M. A. Estrutura Argumental no Português Brasileiro: perdas e alinhamentos inovadores nos sistemas pronominal e preposicional. Cuadernos de la Alfal, v. 12(2), p. 467-490, 2020.

WETZELS, L.; COSTA, J.; MENUZZI,S. *The handbook of Portuguese Linguistics*. USA: Wiley-Blackwell, 2016.

WOOD, J. *Icelandic Morphosyntax and Argument Structure*. New York University, New York City, 2012.

YANG, C. A formalist perspective on language acquisition. *Linguistic Approaches to Bilingualism* 8(6). p. 665–706. 2018.

http://doi.org/10.5007/1984-8420.2025.e106514

A TOPICALIZAÇÃO DE OBLÍQUOS NUCLEARES EM PEÇAS TEATRAIS BRASILEIRAS DOS SÉCULOS XIX E XX: EVIDÊNCIAS PARA A HIPÓTESE DE COMPETIÇÃO DE GRAMÁTICAS

THE TOPICALIZATION OF NUCLEAR OBLIQUES IN BRAZILIAN PLAYS FROM THE 19TH AND 20TH CENTURIES: EVIDENCE FOR THE HYPOTHESIS OF GRAMMAR COMPETITION

Mônica Tavares Orsini | Lattes | monica.orsini@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Marco Antonio Rocha Martins | Lattes | marcomartins.ufsc@gmail.com
Universidade Federal de Santa Catarina/CNPq

Resumo: Este artigo tem como objetivo investigar as construções de topicalização de oblíquo nuclear, em que o tópico é um constituinte projetado por um predicador verbal ou nominal e movido para a periferia esquerda da sentença. Partimos da descrição de Mateus *et al.* (2003) e Raposo *et al.* (2013), segundo a qual as construções em foco possuem, no PE, duas restrições, a saber: não licenciam supressão de preposição com mais conteúdo semântico e não ocorrem em contexto de ilha sintática. Os dados, coletados de 20 peças teatrais escritas por 8 autores brasileiros nascidos no decorrer dos séculos XIX e XX, apontam para o instanciamento da gramática do PE no século XIX e a competição entre as gramáticas do PE e do PB, nos termos de Kroch (1989, 2021[2001]), no século XX. É nos autores nascidos depois de 1920 que encontramos ocorrências com supressão de preposição com mais conteúdo semântico, uma evidência do instanciamento da gramática do PB.
Palavras-chave: sintaxe; topicalização de oblíquo nuclear; mudança linguística; competição de gramáticas.

Abstract: This paper aims to investigate nuclear oblique topicalization constructions, in which the topic is a constituent projected by a verbal or nominal predicator and moved to the left periphery of the sentence. We start from the description by Mateus *et al.* (2003) and Raposo *et al.* (2013), according to which the constructions in focus have, in EP, two restrictions, namely: they do not license the suppression of prepositions with more semantic content and they do not occur in the context of a syntactic island. The data, collec-

ted from 20 plays written by 8 Brazilian authors born during the 19th and 20th centuries, point to the instantiation of EP grammar in the 19th century and the competition between EP and BP grammars, in Kroch's terms (1989, 2021[2001), in the 20th centuries. It is in authors born after 1920 that we find occurrences with suppression of prepositions with more semantic content, an evidence of the instantiation of BP grammar.

Keywords: syntax; oblique topicalization; linguistic change; grammar competition.

1. O fenômeno da topicalização e sua caracterização no Português

As construções denominadas *topicalização* por Mateus *et al.* (2003) e Raposo *et. al.* (2013) são um tipo específico de construção de tópico marcado no qual o constituinte, movido para a periferia esquerda da sentença, exibe correferencialidade com uma categoria vazia, no interior do comentário. Quando o constituinte movido é um complemento oblíquo do predicador (verbal ou nominal), nomeamos a estrutura de *topicalização de oblíquo nuclear*. Em (1), o argumento interno do predicado verbal *depender*, que é um verbo de dois lugares, foi movido para a posição sintática de tópico, deixando, na posição sintática do oblíquo nuclear, a categoria vazia.

(1) FAUSTINO – Nada digas do que ouviste; é preciso que ninguém saiba que eu estou aqui incógnito. *[Do segredo]$_i$ depende ____$_i$ a nossa dita*[1]. (*Judas em Sábado de Aleluia*, Martins Pena)

Um sintagma na função de complemento oblíquo de um predicador nominal também pode se mover para a posição de tópico, como exemplificado em (2), em que *Nesse ponto* é complemento do predicador *razão*.[2]

(2) DOLORES – *[Nesse ponto]$_i$, o Nilson tem razão ____$_i$*. Juntar os panos quando a gente não tem nada é uma besteira. E essas histórias terminam tudo igual a Neiva. Deus que me perdoe, terminar igual a Neiva! Tendo que fugir pra cá, quando a menina dorme, tendo que fazer comida pra aquele cachorro num fogareiro de duas bocas. Eu tenho horror a fogareiro. Eu posso morar embaixo da ponte, mas levo meu fogão comigo. (*No coração do Brasil*, Miguel Falabella)

[1] Optamos por colocar em itálico a construção de topicalização de oblíquo nuclear a fim de destacá-la das demais frases da fala do personagem.

[2] A construção "ter razão em" é prevista por Luft (2002, p. 426), configurando-se em mais um dado de topicalização de oblíquo nuclear.

Nas construções de *topicalização de oblíquo nuclear* exemplificadas em (1) e (2), existe elevado grau de sintatização, pois o tópico exibe conformidade referencial, categorial, casual e temática com o constituinte no interior do comentário. Mateus *et al.* (2003) e Raposo *et. al.* (2013) afirmam ainda que essas construções caracterizam-se, em linhas gerais, por serem sensíveis a fenômeno de ilha sintática, não podendo o correferente situar-se no interior de uma oração subordinada relativa ou adverbial, enquanto o tópico encontra-se à esquerda da oração matriz. Embora as gramáticas portuguesas (assumimos aqui PE) pesquisadas não apresentem dados de topicalização de oblíquo nuclear com predicadores nominais, optamos por agrupar, neste estudo, as ocorrências de topicalização de oblíquo nuclear projetados por predicadores verbais e nominais.

Tendo definido o objeto sintático de estudo, passemos agora para uma discussão referente às restrições que limitam os contextos de produção das estruturas aqui focalizadas.

As gramáticas estudadas afirmam que o português (e aqui assumimos tratar-se do PE) admite construções em que o constituinte topicalizado não respeita todas as propriedades de subcategorização do predicador, sendo estas uma variante da topicalização, típica da modalidade oral. Nos exemplos em (3) a preposição *de* está ausente, tanto em contexto de frase raiz (3a), quanto em contexto de ilha fraca (3b). Empregamos o termo *frase raiz* para nos referirmos às construções em que tópico e correferente estão no interior de uma oração absoluta, como em (3a), uma oração matriz ou ainda uma oração coordenada. No contexto de ilha sintática fraca, o tópico está na oração matriz enquanto o correferente está no interior de uma oração subordinada completiva de verbo, como vemos em (3b).

(3) (a) Essa cerveja$_i$ eu não gosto ____$_i$. (Mateus *et al.*, 2003, p. 501)
(b) Esse relatório$_i$ acho que não precisamos ____$_i$ para a reunião de hoje. (Mateus *et al.*, 2003, p. 501)

Essa construção, exemplificada em (3), é denominada de *topicalização selvagem* ou *topicalização não canônica* e apresenta restrições sintática e semântica: (i) ocorre apenas em frase raiz e (ii) tolera apenas supressão de preposição com menos conteúdo semântico. Com base nestas restrições, analisemos os exemplos arrolados por Mateus *et al.* (2003) em (4) e (5).

(4) a) */? Todos sabem que [essa cerveja]$_i$ eu não gosto ____$_i$. (Mateus *et al.*, 2003, p. 501)
(b) */? Imprime o ficheiro inteiro, embora [esse relatório]$_i$ acho que não precisamos ____$_i$ para a reunião de hoje. (Mateus *et al.*, 2003, p. 501)

(5) (a) *[O João]$_i$ conversei ____$_i$ na festa. (Mateus *et al.,* 2003, p. 501)
(b) *[Paris]$_i$ aterrei ____$_i$ às 10h horas. (Mateus *et al.,* 2003, p. 501)

Em (4a), o tópico e sua categoria vazia estão no interior da oração subordinada, introduzida pelo complementizador *que*; em (4b), o tópico está fora da subordinada completiva de verbo (e dentro de uma oração adverbial concessiva) enquanto a categoria vazia está no interior da completiva verbal. Desta maneira, os dados em (4) aplicam a restrição sintática, configurando-se, para os portugueses, em estruturas agramaticais ou inaceitáveis[3].

Os dados em (5) atualizam a restrição de ordem semântica, já que a agramaticalidade (e não uma possível inaceitabilidade) de (5a) e (5b) decorre da supressão das preposições *com* e *em*, respectivamente, ambas preposições que carregam conteúdo semântico quando combinadas às formas verbais *conversar* e *aterrar*.

Dos dois critérios arrolados por Mateus *et al.* (2003) que atuam como restrições para a gramaticalidade de uma construção de topicalização de oblíquo nuclear, o segundo, isto é, o critério referente à natureza semântica, que diz respeito à possibilidade de supressão de uma preposição a depender do seu conteúdo semântico, revela-se mais forte entre os portugueses, se comparado ao critério sintático. Assim, para os portugueses, enquanto as orações em (4) podem ser agramaticais ou inaceitáveis, as construções em (5) são exclusivamente agramaticais.

Mateus *et al.* (2003), ao tratar especificamente do PB, afirmam que os falantes, apenas na modalidade oral, aceitam como gramaticais as frases ilustradas em (6), o que não é aceito ou produzido por falantes do PE. Neste dado, a preposição *com*, que veicula sentido comitativo, foi suprimida.

(6) (*PE, OK PB) Linguista$_i$ a gente não pode conversar ____$_i$ mais não. (Mateus *et al.*, 2003, p. 502)

[3] Vale registrar que o (*) representa a estrutura agramatical, isto é, a construção ausente na gramática de uma língua, enquanto o ponto de interrogação (?) aponta para a inaceitabilidade da estrutura, isto é, a construção não é interpretada como sendo boa, aceitável em uma dada língua, ainda que seja gramatical.

O fenômeno da topicalização de oblíquo nuclear também foi objeto de estudo de alguns trabalhos diacrônicos. A seguir, buscamos averiguar o que eles dizem acerca dessa construção.

Decat (1989) analisa dados extraídos de diários e correspondências oficiais e pessoais escritos entre o início do século XVIII e a primeira metade do século XIX. Embora seu objetivo não tenha sido estudar especificamente as construções de topicalização de oblíquo nuclear, não foram encontradas ocorrências em que tivesse havido supressão de preposição, diferentemente dos trabalhos sincrônicos de Braga (1986) e Pontes (1987), que registram, no PB oral do século XX, dados de topicalização de oblíquo nuclear com supressão de preposição, como exemplificado em (7) e (8).

(7) [Pão com presunto]$_i$, eu também gosto ____$_i$. (Decat, 1989, p. 131 *apud* Braga 1986, p. 57)

(8) [Meu cabelo]$_i$ eu não gostei nem um pouco (*dele*)$_i$. (Decat, 1989, p. 125 *apud* Pontes, 1987, p. 66)

Assim, enquanto Decat (1989) não encontra dados de topicalização de oblíquo com supressão de preposição nos séculos XVIII e XIX, Braga (1986) e Pontes (1987), em estudos sincrônicos, com dados do PB oral, encontram-nos. A pesquisa de Decat (1989) aponta para uma mudança no comportamento das construções de tópico em decorrência do seu encaixamento em um conjunto de mudanças ocorridas no Português Brasileiro, com destaque para o enfraquecimento do sistema de clíticos.

Araújo (2006) realiza um estudo comparativo das construções de tópico com dados coletados dos séculos XVIII e XIX do PE e do século XIX do PB[4], tratados pela autora como variedades do português, a fim de identificar diferenças e/ou semelhanças sintáticas e discursivas. Em sua extensa tese, porém, não foram encontrados exemplos com supressão de preposição em construções de topicalização de oblíquo, de objeto indireto ou de complemento nominal.

Ferreira (2014) estuda as construções de tópico marcado em cartas pessoais brasileiras dos séculos XVIII, XIX e XX, com o objetivo de verificar quais construções estão presentes na escrita de brasileiros nascidos nos séculos em questão. Os resultados mos-

[4] Os dados do PE foram coletados de cartas e peças de teatro escritas por portugueses no decorrer dos séculos XVIII e XIX. Estas fazem parte do *Corpus* do Projeto Tycho Brahe, organizado pela equipe coordenada pela professora Drª. Charlotte Galves, da UNICAMP. Em relação às cartas e às peças de teatro escritas por brasileiros, no século XIX, foram coletadas do acervo da Biblioteca Pública do Estado da Bahia.

tram que, embora as construções de tópico marcado apareçam na escrita de brasileiros desde a segunda metade do século XVIII, não há dados de topicalização de oblíquo com supressão de preposição, resultado que vai ao encontro dos achados de Araújo (2006).

Este trabalho, portanto, objetiva descrever e analisar, em perspectiva diacrônica, como se comportam as construções de topicalização de oblíquo nuclear, com predicadores verbais e nominais, em dados colhidos de peças teatrais escritas por brasileiros, naturais do Rio de Janeiro, nascidos no curso dos séculos XIX e XX. De forma particular, interessa-nos olhar, qualitativamente, para os dados de cada autor a fim de investigar (i) se há manutenção (ou não) da preposição, descrevendo sua natureza semântica, nos casos em que esta é suprimida, e (ii) se as construções violam (ou não) a restrição de ilha sintática. Pretendemos, assim, verificar qual(is) gramática(s) encontram-se instanciada(s) na escrita desses autores.

Para este fim, ancoramo-nos no modelo de Competição de Gramáticas proposto por Kroch (1989; 2001[2021]), já que os autores brasileiros podem apresentar dados que instanciam a gramática do PB, aqui considerada inovadora, ou do PE. Nossa hipótese é a de que a escrita dos autores nascidos nos séculos XIX e XX apresenta indícios de competição entre as duas gramáticas do português, PB e PE, mas a gramática do PB se instancia com mais força na escrita daqueles nascidos no curso dos novecentos. Nesse sentido, encontraremos, com mais expressividade, estruturas de topicalização de oblíquo nuclear com supressão de preposição com mais conteúdo semântico em textos de autores brasileiros nascidos no século XX, nos valendo da acertada afirmação de Tarallo (1996) de que é nesse período "que o Português Brasileiro deixa correr sua tinta com mais expressividade".

Este artigo encontra-se estruturado da seguinte forma: na seção 2, descrevemos a metodologia utilizada e listamos os autores das peças teatrais nascidos no curso dos séculos XIX e XX; na seção 3; discutimos os resultados encontrados e, por fim, tecemos nossas considerações finais, seguidas das referências e do quadro com as peças lidas para a coleta dos dados dessa pesquisa.

2. Procedimentos metodológicos

Ao optamos por fazer uma análise qualitativa, considerando cada autor, de forma particular, intentamos verificar se os dados por eles produzidos, no que tange ao fenômeno em análise, refletem o instanciamento da gramática do PE, do PB ou de ambas as gramáticas.

A amostra constitui-se de 20 peças teatrais de diferentes autores cariocas nascidos no curso dos séculos XIX e XX. A escolha pelo gênero peça teatral, todas de cunho popular, se dá por este configurar-se em um gênero híbrido. Trata-se de um texto escrito para que sua produção ocorra por meio da modalidade oral (cf. Marcuschi, 2008). Desta forma, acreditamos que os dados coletados de cada autor reflitam a sua gramática, isto é, a sua Língua_I, nos termos de Mendívil-Giró (2019), fundamentado em Chomsky (1985).

O quadro 1 reúne os autores e seu ano de nascimento[5].

Autor	Ano de nascimento
Martins Pena	1815
Joaquim Manuel de Macedo	1820
França Junior	1838
Gastão Tojeiro	1880
Armando Gonzaga	1884
Silveira Sampaio	1914
Millôr Fernandes	1923
Carlos Eduardo Novaes	1940
Miguel Falabella	1956

Quadro 1: Autores que compõem a amostra e seu respectivo ano de nascimento

Do conjunto dos 9 autores, encontramos dados em 8 deles, não havendo ocorrências apenas em Silveira Sampaio. Desta forma, serão analisados dados de 8 autores, todos do gênero masculino, naturais da cidade do Rio de Janeiro, com exceção de Joaquim Manuel de Macedo, que nasceu em Itaboraí, cidade da região metropolitana do Rio de Janeiro, a 45 km de distância da Capital.

3. Análise dos dados

Começamos nossa análise com Martins Pena, nascido em 1815, primeiro quartel do século XIX. É o autor em que ocorreu o maior número de dados. De suas peças, foram coletadas 19 ocorrências de topicalização de oblíquo nuclear, sendo 17 de complemento de verbo e 2, de nome.

Listamos, a seguir, todos os dados encontrados de complemento oblíquo projetado por um predicador verbal.

[5] Apresentamos a lista das peças lidas de cada autor ao final do artigo.

(9) CHIQUINHA – Fazes bem. (Aqui Maricota faz uma mesura para [a] rua, como a pessoa que a cumprimenta depois a fazer acenos com o lenço.) Lá está ela no seu fadário! Que viva esta minha irmã só para namorar! É forte mania! *[A todos]*[6] *$_i$ faz festa* ____$_i$, *[a todos] namora* ____$_i$. (*Judas em Sábado de Aleluia*)

(10) CHIQUINHA, interrompendo – Meu Deus, quantos?... *E [a todos esses]$_i$ namoras* ____$_i$? *(Judas em Sábado de Aleluia)*

(11) CHIQUINHA – Os homens têm mais juízo do que pensas; *[com as namoradeiras]$_i$ divertem-se eles* ____$_i$, mas não se casam. (*Judas em Sábado de Aleluia*)

(12) PIMENTA – Torno a dizer, feliz a hora em que deixei o ofício para ser cabo-de-esquadra da Guarda Nacional! *[Das guardas, das rondas e das ordens de prisão]$_i$ faço* ____$_i$ *o meu patrimônio.* Cá as arranjo de modo que rendem, e não rendem pouco... Assim é que é o viver; e no mais, saúde, e viva a Guarda Nacional e o dinheirinho das guardas que vou cobrar, e que muito sinto ter de repartir com ganhadores. *(Judas em Sábado de Aleluia)*

(13) FAUSTINO – Nada digas do que ouviste; é preciso que ninguém saiba que eu estou aqui incógnito. *[Do segredo]$_i$ depende* ____$_i$ *a nossa dita*. *(Judas em Sábado de Aleluia)*

(14) PIMENTA – Assim pensei eu, ou me fizeram pensar; mas já abriram-me os olhos, e... enfim, passarei ainda esta vez, e será a última. Tenho filhos. Meti-me nisto sem saber bem o que fazia. *[E do senhor]$_i$ queixo-me* ____$_i$, porque da primeira vez abusou da minha posição; eu estava sem vintém. É a última! *(Judas em Sábado de Aleluia)*

(15) CAPITÃO – *[Aos insultos]$_i$ respondem-se* ____$_i$ *com as armas na mão!* Tenho uma patente de capitão que deu-me o governo, hei de fazer honra a ela! O senhor é um covarde! Digo-lhe isto na cara; não me mete medo! Há de ir preso! Ninguém me insulta impunemente! (Os três, à proporção que falam, vão reforçando a voz e acabam bramando.) *(Judas em Sábado de Aleluia)*

(16) AMBRÓSIO, só de calça preta e chambre — No mundo a fortuna é para quem sabe adquiri-la. Pintam-na cega... Que simplicidade! Cego é aquele que não tem inteligência para vê-la e a alcançar. Todo homem pode ser rico, se atinar com o verdadeiro caminho da fortuna. Vontade forte, perseverança e pertinácia são poderosos auxiliares. Qual o homem que, resolvido a empregar

[6] Neste período, temos duas orações coordenadas assindéticas com topicalização de obliquo nuclear. São, portanto, dois dados em um único período.

todos os meios, não consegue enriquecer-se? *[Em mim]$_i$ se vê ____$_i$ o exemplo*. *(O Noviço)*

(17) FLORÊNCIA — E era para tratarmos do meu sobrinho Carlos, mas já não é preciso. Aqui esteve o padre-mestre; *[sobre isso]$_i$ falamos ____$_i$*; está tudo justo e sem dúvida Vossa Reverendíssima já está informado. *(O Noviço)*

(18) FABIANA - Está bem, *[em tua mulher]$_i$ basta que só dês quatro (canelões) ____$_i$*. *(Quem casa quer casa)*

(19) JEREMIAS, entrando cauteloso – *[Nesta]$_i$ não veio ____$_i$ ninguém que me inquiete. (As casadas solteiras)*

(20) JOHN – Mostrai-me vossa mão. (Examinando sua mão e falando-lhe mais baixo); Esta linha me diz que teu coração não está livre. Aquele que amas não é da tua nação, mas é um homem honrado e leal; *[dele]$_i$ te podes fiar ____$_i$*. *(As casadas solteiras)*

(21) VIRGÍNIA – As contrariedades do estado nada seriam; *[com elas]$_i$ contava eu ____$_i$, razoavelmente falando*. Porém o que mais me desespera é ter de aturar as manias inglesas de nossos caros maridos... Ontem, o meu quis que eu comesse, por força, rosbife quase cru. *(As casadas solteiras)*

(22) HENRIQUETA – Ah, ah, ah! *[Disso]$_i$ me rio eu ____$_i$*. *(As casadas solteiras)*

(23) JEREMIAS – Fujo sim, *porque [da peste, dos demônios, e das víboras]$_i$ se foge ____$_i$*... Não quero mais te ver! (Fecha os olhos.) *(As casadas solteiras)*

(24) JEREMIAS – *Meu caro senhor Narciso, [a isto]$_i$ não se pode o senhor se opor ____$_i$*; elas querem... *(As casadas solteiras)*

Os dados acima reunidos, todos coletados das peças de Martins Pena, apresentam preposição expressa, independentemente de ser esta portadora de mais ou menos conteúdo semântico. Além disso, nos dados com sujeito expresso - (11), (13), (16), (21), (22), (24) – estes estão sempre pospostos ao verbo e o constituinte tópico à sua esquerda. Segundo Torres Morais (1996), no português escrito no Brasil, por interferência da gramática do PE, a topicalização com inversão verbo-sujeito ocorre até o início do século XX.

Berlinck *et al.* (2016), ao investigar a ordem [DP]V / V[DP], em sentenças do tipo sujeito-predicado, apontam que, no século XIX, a frequência de V[DP] é maior (24%), caindo para 12,5%, no século XX. Considerando as mudanças sintáticas em curso no PB, no decorrer dos séculos XIX e XX, os autores afirmam que os fatores linguísticos que atu-

am na preferência pela ordem V[DP] no século XIX não continuam atuando na escrita do século XX.

Na esteira dos trabalhos sobre ordem do sujeito na história do PB, Berlinck e Coelho (2018), com base em dados coletados de cartas de leitores, anúncios e cartas pessoais dos séculos XIX e XX, disponíveis nos *corpora* do Projeto *Para a História do Português Brasileiro* (PHPB), reafirmam que a sintaxe da construção VS mudou do século XIX para o século XX, evidenciando uma perda dos padrões de inversão do DP sujeito.

No século XIX, ainda encontramos construções de inversão germânica - ([XP]VS), com verbo transitivo em segunda posição e sujeito posposto, como resquício de uma sintaxe V2 do Português Antigo (PA) e do Português Clássico (PC). Os dados encontrados em Martins Pena, com TOP OBL V S, reforçam a tese de que estamos diante da gramática do PE, que prefere construções de topicalização de oblíquo com presença de preposição.

Abaixo transcrevemos os dois dados de topicalização de oblíquo, sendo o SP o complemento de um predicador nominal.

(25) CARLOS - E que culpa tenho eu, se tenho a cabeça esquentada? Para que querem violentar minhas inclinações? Não nasci para frade, não tenho jeito nenhum para estar horas inteiras no coro a rezar com os braços encruzados. Não me vai o gosto para aí... Não posso jejuar; tenho, pelo menos três vezes ao dia, uma fome de todos os diabos. Militar é que eu quisera ser; *[para aí]$_i$ me chama a inclinação* ____$_i$. Bordoadas, espadeiradas, rusgas é que me regalam; esse é o meu gênio. Gosto de teatro, e de lá ninguém vai ao teatro, à exceção de Frei Maurício, que frequenta a plateia de casaca e cabelereira para esconder a coroa. (*O Noviço*)

(26) FLORÊNCIA - Oh, não falo desse direito, não o contesto. Direito de perseguí-lo quero eu dizer.
ROSA - *[Nisso]$_i$ estou de acordo* ____$_i$. (*O Noviço*)

Em (25), o SP *para aí*, que ocupa a posição de tópico, é complemento do substantivo *inclinação*, que constitui o núcleo do argumento externo do verbo *chamar*. O sujeito – *a inclinação para aí* – encontra-se posposto ao verbo, tendo o SP *para aí* se movido para a posição de tópico, uma sintaxe que reflete o resquício de um sistema V2, de períodos históricos anteriores.

Em (26), o SP *nisso* completa a estrutura *estar de acordo em*. O sujeito está nulo, uma preferência da gramática do PE. Duarte (2012), ao investigar o preenchimento do sujeito anafórico de 3ª pessoa em peças teatrais escritas por brasileiros no decorrer dos séculos XIX e XX, detecta o processo de mudança em direção ao sujeito expresso, influenciado fortemente pelo traço [+ humano] do antecedente. Sendo assim, a ausência de sujeito pronominal expresso, em (26), é mais uma evidência de que estamos diante de uma gramática do PE, que prefere manter a preposição em sintagmas preposicionados que estão na posição de tópico.

No que tange à configuração sintática em que ocorre o tópico e seu correferente, no interior do comentário, não há dados em contexto de ilha sintática, como preveem Mateus *et al.* (2003). Todos os dados estão em frase raiz, com exceção do dado transcrito em (18) e repetido em (27).

(27) FABIANA - Está bem, *[em tua mulher]$_i$ basta que só dês quatro (canelões)* ____$_i$. *(Quem casa quer casa)*

O predicador verbal *dar*, em (27), constitui com o SN *quatro canelões* uma unidade semântica, uma expressão idiomática, sendo o SP *em tua mulher* seu complemento oblíquo, movido para a posição de tópico. Nesta construção, o tópico está à esquerda da frase raiz, enquanto o correferente encontra-se no interior da subordinada completiva de nome.

A presença de um dado em que o constituinte tópico – SP *em tua mulher* - esteja na oração matriz e a categoria vazia no interior da oração subordinada na escrita de Martins Pena aponta para o fato de ou (1) de a restrição quanto à configuração sintática da estrutura em que ocorrem tópico e correferente no PE seja menos forte que a restrição referente à natureza semântica da preposição; ou (2) a construção de topicalização de oblíquo nuclear em contexto de ilha sintática é já um indício da gramática do PB na escrita desse autor, uma vez que tal estrutura ou é agramatical ou não aceitável em PE.

Em Joaquim Manuel de Macedo, nascido em 1820, encontramos apenas três dados, todos complementos oblíquos de verbo, movidos para a periferia esquerda da sentença. Abaixo, transcrevemos esses dados. Em todos eles, as preposições são [+ lexicalizadas], já que carregam mais conteúdo semântico, e estão presentes.

(28) ANASTÁCIO - Pois não deve ufanar-se disso. O que mais convém a uma senhora honesta é que não se fale muito em seu nome, nem em bem e ainda menos em mal; *e [a uma menina solteira]$_i$ o que melhor assenta ____$_i$ é, recolhida no seio da modéstia, fazer-se notar pela virtude que não se ostenta*, e que, no entanto, excita a admiração, por isso mesmo que não procura louvores. (*Luxo e vaidade*)

(29) ANASTÁCIO - Escuta: há vinte cinco anos aquela mulher supunha-se amada por ti, e viu em Hortênsia uma rival preferida, *quando [com esta]$_i$ te ligaste ____$_i$ em casamento*. O desprezo de um homem abre no seio da mulher uma ferida envenenada que nunca cicatriza. A ofensa, foste tu que a fizeste, mas a mulher desprezada detesta ainda mais que ao ofensor a rival que triunfou. Assim, pois, diz a lógica, que Fabiana aborrece profundamente a tua esposa. (*Luxo e vaidade*)

(30) ANASTÁCIO - Acima dos meus parentes está a nação que pode colher benéficos resultados da lição que oferece a sua desgraça. A sociedade acha-se corrompida pelo luxo e pela vaidade, e um quadro vivo das consequências desastrosas dessas duas paixões talvez lhe seja de prudente aviso. *[Em Maurício]$_i$ verá ____$_i$ o homem de medíocre fortuna e especialmente o empregado público, que a ostentação e o fausto de alguns anos determinam a miséria de todo o resto da vida*; nas suas lágrimas de esposa e de mãe, as mães e as esposas verão os horrores a que as pode levar o abuso do amor de um marido extremoso e cego e a falsa educação dada às filhas. (*Luxo e vaidade*)

Em (28) o verbo *assentar* solicita preposição *a*, recebendo o complemento papel temático de tema. Neste caso, a preposição carrega menos conteúdo semântico. O constituinte tópico, o SP *a uma menina solteira*, encontra-se na periferia esquerda do sintagma complexo *o que melhor assenta a uma menina solteira* que, por sua vez, é sujeito da oração subordinada.

Em (29), o verbo *ligar* é regido pela preposição *com*, que exprime a ideia de conexão, combinação, como descreve Luft (2010). Tópico e correferente encontram-se no interior de uma oração subordinada adverbial de tempo, o que não compromete a restrição de ilha prevista para esta estrutura.

Em (30), deparamo-nos com a seguinte estrutura argumental: *ver em algo ou alguém + predicativo,* o que, também, não compromete a restrição de ilha imposta às cons-

truções de topicalização no PE. A preposição *em* carrega mais conteúdo semântico ao introduzir o tema.

França Junior nasceu em 1838 e é o último escritor aqui analisado nascido na primeira metade do século XIX. De suas peças, foram coletados 12 dados: 11 de topicalização de oblíquo vinculado a um verbo; 1 de topicalização de oblíquo vinculado a um predicador nominal. Abaixo, transcrevemos o conjunto de dados de topicalização de oblíquo projetado por um verbo.

(31) MATIAS - (Canta) *[De ter alamões em casa]$_i$ / Ninguém deve se queixar ____$_i$;* / Pois é gente papafina. / Para uma casa guardar. (*Defeito de família*)

(32) MATIAS - (...) Quem quiser ter o sossego / E a paz no coração, / *[Lá da terra das bengalas]$_i$ / Mande vir ____$_i$ um alamão*. (*Defeito de família*)

(33) ERNESTO - Com os diabos! *[Por esta]$_i$ não esperava eu ____$_i$*. Estou aqui, estou demitido. (*Caiu o Ministério*)

(34) BEATRIZ – *[Disto]$_i$ lá nunca vi ____$_i$;* pelo menos em Paris. (*Caiu o Ministério*)

(35) RAUL - *[Nessa]$_i$ não caio eu ____$_i$*! É pobre como Jó, e mulher sem isto (Sinal de dinheiro) está se ninando. Vamos embora. (Saem.) (*Caiu o Ministério*)

(36) FILOMENA - *Creio que [entre nós]$_i$ nada se tem passado ____$_i$ que possa porventura interromper, sequer de leve, as nossas relações amistosas*. (*Caiu o Ministério*)

(37) CORO DE CAPANGAS - Que o voto é livre / Ninguém duvida! / *[Por nossos amos]$_i$ / Demos a vida ____$_i$*. (*Com se fazia um deputado*)

(38) MARIA (à parte.) - Parece incrível! *[Isto contado]$_i$ ninguém acredita ____$_i$*. (*As Doutoras*)

(39) MARIA - Nas rixas que se dão constantemente nesta casa já viste envolvido o meu nome? Sou para o Doutor Pereira uma criatura completamente indiferente. *[Dos seus lábios]$_i$ ainda não partiu contra mim ____i a mais pequena censura, ou uma palavra sequer que deixasse transparecer, embora sutilmente, o veneno do epigrama*. (*As Doutoras*)

(40) EULÁLIA — O quê? Meter essa seringa no meu corpo? Com caldo de febre amarela? *[Em mim]$_i$ o senhor não mete isto ____$_i$, não, mas é o mesmo*. Chegue-se para lá, patrão. (*As Doutoras*)

(41) EULÁLIA — Avisá-la? *[Nessa]$_i$ não caio eu ____$_i$!* (*As Doutoras*)

O dado (38) deve ser tratado de forma separada por apresentar a supressão da preposição. Nele, o verbo *acreditar*, transitivo indireto, segundo Luft (2010), encontra-se desacompanhado da preposição *em*. Podemos interpretar que houve supressão de uma preposição com menos conteúdo semântico, o que é previsto por Mateus *et al.* (2003) e Raposo *et al.* (2013), para gramática do PE. Os demais têm um constituinte tópico encabeçado por preposição, com a função de complemento oblíquo, independentemente de ser a preposição mais ou menos lexicalizada. Não há dados com supressão de preposição com mais conteúdo semântico, confirmando nossa hipótese de que estas estruturas por serem próprias do PB somente aparecerão apenas na escrita dos autores nascidos no século XX.

Nos dados (33), (35) e (41), o sujeito expresso encontra-se posposto, comportamento que reforça a tese de se tratar da gramática do PE, como já discutido, com base nos trabalhos de Torres Morais (1996), Berlinck *et al.* (2016) e Berlinck e Coelho (2018).

O dado com predicador nominal, encontra-se transcrito em (42).

(42) LUÍSA - *[Neste ponto]$_i$, minha mãe, o Pereira tem razão* ____$_i$! (*As Doutoras*)

Em (42), a forma ter razão é regida pela preposição *em*, que se mantém presente no Sintagma Tópico.

Em relação à configuração sintática da estrutura em que ocorrem tópico e correferente, constatamos que os dados respeitam a restrição prevista em Mateus *et al.* (2003), segundo a qual a topicalização é sensível à ilha sintática, estando todos em frase raiz ou no interior de oração subordinada.

Passemos, agora, aos autores da segunda metade do século XIX: Gastão Tojeiro e Armando Gonzaga. Em Gastão Tojeiro, nascido em 1880, encontramos 4 dados de topicalização de oblíquo projetado por predicador verbal. Não há dados de complemento de nome. Transcrevemos, abaixo, os 4 dados encontrados.

(43) INACIA: *[De passear]$_i$ você nunca se esquece* ____$_i$. (*Onde canta o Sabiá*)

(44) IRENE, intencionalmente, - Não é só com os heróis que sucede isso; *[com certas heroínas]$_i$ dá-se o mesmo* ____$_i$. (*A inquilina de Botafogo*)

(45) DESIDERIO – Por que não hei de repetir? Ou a senhora muda-se ou mando despejá-la. *E [disso]$_i$ nem o papa a livra* ____$_i$! (*A inquilina de Botafogo*)

(46) DESIDERIO – Sim, sim ... Estou percebendo. *[Para um bom entendedor]$_i$ poucas reticências bastam* ____$_i$. (*A inquilina de Botafogo*)

Em (43a), abaixo, colocamos o tópico *De passear* na sua posição *in situ*, constituindo a perífrase verbal *esquecer de passear*. Nesta posição, a supressão da preposição gera, tanto para a gramática do PE quanto para a gramática do PB, uma construção agramatical, transcrita em (43b). Porém a construção (43c), no entanto, na qual o constituinte *De passear* está na posição de tópico apresenta a supressão da preposição *de*, é licenciada na gramática do PB falado[7], não o sendo na do PE falado, considerando as restrições descritas por Mateus *et al.* (2003). Essa diferença de comportamento aponta para o distanciamento entre as gramáticas do PB e do PE, no decorrer do século XX, confirmando a hipótese de que é neste século que a gramática do PB começa a se distanciar da gramática do PE. (cf. Tarallo, 1996)

(43) a. Você nunca se esquece de passear. (PE / PB)
b. *Você nunca se esquece passear. (*PE / *PB)
c. [Passear]$_i$ você nunca se esquece ____$_i$. (*PE / PB)

Construções como (43)d não foram encontradas na escrita de Gastão Tojeiro. Todas as construções de topicalização de oblíquo nesse autor têm manutenção de preposição, independentemente da sua natureza semântica, e estão em contexto sintático de frase raiz, o que aponta para o instanciamento da gramática do PE.

Em (44), temos a forma verbal *dar-se*, que, no contexto em foco, significa *ocorrer*, com a manutenção da preposição *com*, [+ lexicalizada]; em (45), há a manutenção da preposição *de*, [- lexicalizada].

Em (46), temos uma estrutura análoga a um dito popular *Para bom entendedor, meia palavra basta*. Sendo assim, não acreditamos que seja possível o licenciamento da estrutura sem a preposição, tanto no PE quanto no PB do século XIX, pois em ambas as línguas a construção gerada seria agramatical.

(46) a. *[Um bom entendedor]$_i$ poucas reticências bastam ____$_i$. (*PE / *PB)

Na peça de Armando Gonzaga, coletamos apenas uma ocorrência, sendo esta de complemento de nome. Em (47), a preposição *em* encontra-se contraída ao pronome demonstrativo *este*.

[7] O teste de gramaticalidade foi realizado pelos autores desse artigo, ambos brasileiros, nascidos no século XX.

(47) ZELIA (desviando o golpe) – [Neste ponto]$_i$ o senhor tem razão ____$_i$ (vagamente). Nem tudo que a gente deseja pode ser alcançado. Eu que o diga. (*O hóspede do quarto nº 2*)

Todos os autores do século XIX parecem instanciar a gramática do PE, segundo a qual não há supressão de preposição semanticamente plena nas topicalizações. O comportamento do sujeito posposto sem restrições reforça esta tese. Em apenas um dado houve supressão deste item, exemplo (38), pertencente ao *corpus* produzido por França Junior, licenciado pelo PE por tratar-se de uma preposição com menos conteúdo semântico. Passamos, agora, para a análise dos dados dos autores nascidos no século XX: Millôr Fernandes; Carlos Eduardo Novaes e Miguel Falabella.

Millôr Fernandes, primeiro autor nascido no século XX, no ano de 1923, produziu 11 dados, sendo 9 de topicalização de oblíquo com predicador verbal e 2, com predicador nominal. Os 9 dados de complemento verbal foram produzidos com manutenção de preposição, tenha ela mais ou menos conteúdo semântico. Abaixo, estão transcritos os dados.

(48) MARIA – *[Até nisso]$_i$ você me enganou* ____$_i$: a barba era porque ia entrar num clube de mágicos! (*Um elefante no caos*)

(49) ROSA – Crescidinho (pausa). *[Com os filhos]$_i$ sucede* ____$_i$ *o contrário* – as empregadas é que tem que tomar cuidado. (Suspira) A senhora não acha que as relações entre as empregadas domésticas são altamente explosivas? (*Um elefante no caos*)

(50) ROSA – É uma história comprida, mas vou contar bem curta. Naquela noite, quando nos separamos, percebi *que [de ti]$_i$ sabia* ____$_i$ *apenas teu primeiro nome: Paulo.*

(51) ROSA – *E que [de meu]$_i$, nem o nome você sabia* ____$_i$. (*Um elefante no caos*)

(52) PAULO – Não temos salvação, *então, [nem para hineiro]$_i$ eu dou* ____$_i$. De que vamos viver? (*Um elefante no caos*)

(53) CARLOS – Imprensa é oposição: o resto é armazém de secos e molhados. *Mas [em armas]$_i$ eu não peguei* ____$_i$ *não*. Não sei nem onde é que fica o gatilho de um revólver. (*Os órfãos de Jânio*)

(54) BETO – E o Bob Hope? Gastava mais de um milhão de dólares por mês para ser esculhambado pela esquerda. *[Com isso]$_i$, ele ganhava* ____$_i$ *vários milhões, com o patrocínio da direita.* (*Os órfãos de Jânio*)

(55) BETO - (...) provocou uma rearrumação apressada e caótica de todos os músculos e órgãos da cabeça, sendo que o preto do olho se voltou para dentro, me dando pela primeira vez a oportunidade de constatar minha alma branca, os miolos, comprimidos e descomprimidos pela força da inércia e da procura, pensaram coisas até então jamais pensadas, *[da bochecha em fogo]*$_i$ *saiu uma fumaça azul celeste* ____$_i$ e, no interior da boca, a língua enlouquecida apanhou com avidez dois dentes brancos que nadavam alegremente num litro de sangue (...). (*Os órfãos de Jânio*)

(56) BETO – Olha, eu sei, sou bem barra pesada, já vi muita gente levar porrada em minha vida, mas nunca, nem antes nem depois, vi ninguém levar porrada igual àquela. *[Disso]*$_i$ *eu me orgulho* ____$_i$. Foi assim que eu fui descoberto. (*Os órfãos de Jânio*

Em todos os dados de topicalização de oblíquo nuclear, produzidos por Millôr Fernandes, temos presença de preposição, com menos conteúdo semântico, em (48), (50), (51), (53) e (56), e com mais conteúdo semântico, em (49), (52), (54), (55). Tal comportamento leva-nos a interpretar os dados como uma característica do PE, como prevê Mateus *et. al.* (2003) e Raposo *et al.* (2013) para o PE, e esse padrão é o encontrado na escrita de Millôr, que nasceu na primeira metade do século XX.

A presença da forma *você*, em (48), sugere que estamos diante da gramática do PB, em cuja sintaxe o constituinte passa de pronome de tratamento para pronome nominativo. Lopes (2007, p. 104) afirma que "Em termos evolutivos, o uso majoritário de *tu* – forma recorrente no século XIX – só será suplantado por *você* por volta dos anos 1920-30 do século passado." Duarte (1996), em trabalho anterior ao de Lopes (2007), afirma que, no início do século XX, já temos o pronome *você* incorporado ao paradigma pronominal do PB, em variação com a forma pronominal *tu*. Portanto, Millôr Fernandes, nascido em 1923, já é exposto a uma gramática que utiliza a 2ª pessoa indireta, em situações informais de comunicação, entre pessoas que possuem bastante intimidade, como temos no diálogo entre Maria e Paulo, mãe e filho.

No caso do fenômeno aqui investigado, fica, portanto, evidente que PE e PB compartilham a mesma propriedade de mover para a posição de tópico constituintes preposicionados, com preferência para a sua manutenção. E quanto à configuração sintática das sentenças em que ocorrem tópico e correferente, não há dados que violem a restrição de ilha imposta às topicalizações, prevista por Mateus *et al.* (2003).

Já os dois dados de topicalização de oblíquo de um termo nominal têm a preposição suprimida. Eles estão transcritos abaixo.

(57) NELITA – Eu, desde menina, não aguentava Cuiabá. Vivia de olho no Rio e em São Paulo. Lia tudo quanto era jornal, sobretudo notícia social e as fofocas sobre intelectuais prafrentex, grandes artistas de vanguarda. E fingia que lia os livros sérios, *mas [esses]$_i$, francamente, eu não tinha saco* ____$_i$. Acho que ninguém tem. (*Os órfãos de Jânio*)

(58) CONCEIÇÃO – O presidente tinha pegado a ideia na Índia, não, no Egito – ele era tarado pelo Nasser: eu era tarada por ele. *[Tudo o que ele fazia]$_i$ eu estava de acordo* ____$_i$. (*Os órfãos de Jânio*)

Em ambos os dados, vemos o instanciamento da gramática do PB, que licencia a supressão de preposição com mais conteúdo semântico. Em (57), temos a supressão da preposição *para*, com ideia de finalidade. O pronome demonstrativo *esses* retoma o referente os *livros sérios*. Temos, então, a estrutura "*não ter saco para os livros sérios*". O tópico e seu correferente encontram-se dentro de uma oração de caráter explicativo, já que a conjunção *mas* parece remeter ao valor semântico da conjunção *porque*. Assim, Nelita "apenas *fingia* ler os livros sérios, porque para esse tipo de livro não tinha saco". Temos, portanto, na oração introduzida pelo Sintagma Tópico *esses*, a explicação para a ação de *fingir ler*.

Em (58), a preposição *com*, dotada de mais conteúdo semântico, foi suprimida, visto tratar-se da construção *estar de acordo com*. Neste dado, o Sintagma Tópico *tudo que ele fazia* apresenta-se como um SN complexo, já que seu núcleo é modificado pela oração relativa restritiva. Dados com SN complexo não estão presentes nas descrições de Mateus *et al.* (2003) e Raposo *et al.* (2013).

A análise dos dados de Millôr Fernandes aponta para uma escrita em que verificamos a competição entre as gramáticas do PE e do PB, nos termos de Kroch (1989; 2021[2001]), numa diglossia internalizada. A inter-relação entre história social e gramatical da escrita no Brasil, nos termos de Duarte (2025, neste volume), explica esse conhecimento compartilhados de diferentes de gramáticas do português. A ausência das preposições *para* e *com*, nos dados (57) e (58), respectivamente, ambas [+ lexicalizadas], refletem o instanciamento da gramática do PB, que licencia a supressão de preposição semanticamente plena, em variação com a sua realização, desde a primeira metade do século XX. Em contraposição, a gramática do PE, tanto no passado – século XIX e XX

- quanto no presente, só permite supressão de preposição com menos conteúdo semântico, conforme descrição das gramáticas que tomamos como referência para esta investigação. Assumimos, neste estudo, que a presença da preposição e a supressão de preposição com menos conteúdo semântico são comuns a ambas as gramáticas.

Em Carlos Eduardo Novaes, nascido em 1940, encontramos 8 dados de topicalização de oblíquo com predicador verbal e 2, com predicador nominal. Abaixo, seguem os 8 dados de topicalização de oblíquo com predicador verbal.

(59) CRISTINA – *[Teatro]$_i$ vamos ____$_i$ de vez em quando*, só quando tem alguma comédia porque ele diz que só vai ao teatro para se divertir. Outro dia fomos ver A mulher integral. Você já viu? (*A mulher integral*)

(60) JULIA – *Sim, [no assunto]$_i$ não toca ____$_i$*, mas toca na sua bunda? (Pausa) Bem, isso não interessa agora. Já procurou alguma vez se autosatisfazer? (*A mulher integral*)

(61) ARNALDO – É uma mulher maravilhosa, compreensiva, equilibrada, inteligente. Foi dessas paixões fulminantes. Quando bati os olhos nela, senti alguma coisa me dizendo: Arnaldo, essa é a mulher da sua vida. Fui conferir e era mesmo. *[Por ela]$_i$ estou disposto a largar tudo ____$_i$*. (*A mulher integral*)

(62) CRISTINA – *[O (dinheiro) que você me deu a mais]$_i$, amorzinho, eu comprei manteiga ____$_i$*. Nós vamos gastar muita manteiga, não vamos, meu amor? (*A mulher integral*)

(63) SERGIO - Já perdi. *[Por você]$_i$ eu perco tudo ____$_i$*. Deixe-me beijá-la; se você não deixar, eu não vou embora. (*A mulher integral*)

(64) ALICE – Bem, se você me aconselha, eu vou. Você sabe que *[de homem]$_i$ eu só entendo ____$_i$ do plano ideológico*. (*A mulher integral*)

Em (59) e (62), as preposições *a* e *com*, respectivamente, são suprimidas, ambas com mais conteúdo semântico: *a*, em (59), combinada ao verbo de movimento, introduz a ideia de destino; *com*, em (62), carrega o sentido de custo, valor. Esses dados revelam que o Carlos Novaes usa estruturas que instanciam a gramática do PB, e que se encontram em competição com a do PE. Os demais dados refletem a gramática tanto do PE quanto do PB, já que se assemelham no que tange ao comportamento das construções de topicalização de oblíquo nuclear com manutenção de preposição.

Os dois dados de topicalização de oblíquo com predicador nominal são produzidos com manutenção de preposição, ambas com mais conteúdo semântico.

(65) Alice – Não me admira que você tenha que empurrar o caminhão sozinha. *[Pra ele]$_i$ a situação ____$_i$ é muito cômoda.* Tem uma mulherzinha que defende a retaguarda, cuida da casa, das roupas, da comida... (*A mulher integral*)

(66) SERGIO - Sim, *mas [por nenhuma]$_i$ tão apaixonado ____$_i$ como por você.* Sabe? Nós nos identificamos muito. Nas menores coisas... veja, você está com as cores do Império Serrano e eu sou Império. (*A mulher integral*)

Quanto à configuração sintática das construções coletadas em Carlos Eduardo Novaes, não identificamos dados em que seja violada a restrição de ilha sintática, comportamento semelhante ao observado nos dados de Miguel Falabella, nascido em 1956, segunda metade do século XX.

Em sua peça, *No coração do Brasil*, encontramos 10 ocorrências de topicalização de oblíquo, sendo 4 dados de complemento de verbo e 6, de nome na escrita de Miguel Falabella. Abaixo, transcrevemos, primeiramente, os 4 dados de topicalização de oblíquo, projetado por um verbo.

(67) DONA IRENE – Uma coisa terrível, meu filho. Os comunistas construíram o muro de surpresas. *[Essa foto]$_i$, nunca hei de esquecer ____$_i$: a moça muito jovem, loura, bonita, vestida de noiva, chorava, estendendo a mão pelo meio do arame farpado, tentando tocar uma outra mão que se aproximava*. E sabe de quem era a outra mão? (*No coração do Brasil*)

(68) MARGARETH - A da mãe dela. *[Dessa parte]$_i$ eu não vou esquecer nunca ____$_i$.* Que horror, né, Dona Irene? (*No coração do Brasil*)

(69) DOLORES – Eu sou muito fértil. *[Comigo]$_i$ não tem mistério ____$_i$*. No que o fuzileiro veio encostando, eu já peguei barriga. Diz que tem uma tal de tabela, mas aí já viu o rolo que ia dar, né? Eu ia ter que ter uma tabela pras regras e outra pros fuzileiros. Ah, deixa pra lá! (*No coração do Brasil*)

(70) Gabriel – Foi mais ou menos nessa época que eu comecei a partir. Eu tinha vontade de voltar ao cinema, de rever aquela gente toda, mas a ideia de voltar sempre me assustou. Era como se houvesse uma armadilha me esperando. Uma armadilha que me manteria aprisionado lá para sempre. Um dia desses eu li alguma coisa sobre Anais de Akasha. *[Neles]$_i$ está guardada ____$_i$ a memória prodigiosa do universo*. Lá, como numa gigantesca tela, pode-se ver tudo o que aconteceu na terra. (*No coração do Brasil*)

A presença de dois dados com o mesmo verbo, *esquecer,* ora com a preposição *de ausente,* dado (67), ora presente, dado (68), é bastante elucidativa. Essa alternância entre presença x ausência de preposição, mesmo sendo esta [- lexicalizada], vai ao encontro da tese de que a gramática do PB, instanciada na escrita dos autores nascidos no século XX, pois licencia a variação entre presença e ausência de preposição no Sintagma Tópico, independentemente da sua natureza semântica, comportamento diverso da gramática do PE, que parece preferir a presença da preposição, com poucos dados de supressão, somente quando a preposição apresenta menos conteúdo semântico. Por isso, temos predominância quase absoluta de dados com preposição presente na escrita dos autores brasileiros nascidos no século XIX, momento em que predomina a gramática do PE.

Em (69), o tópico é constituído pelo pronome oblíquo tônico, *comigo*, o que inviabiliza a ausência de preposição. Em outras palavras, não existe a forma pronominal tônica de 1ª pessoa do singular sem preposição no sistema linguístico. Em (70), o Sintagma Tópico *Neles* ocupa a primeira posição na sentença, estando o sujeito posposto. É interessante observar que, embora a gramática do PB tenha sofrido mudanças em relação à ordem do sujeito, sendo SVO a forma canônica, a posposição do sujeito, neste dado, parece levar à topicalização para evitar que o verbo fique em posição inicial.

Miguel Falabella produziu 6 dados de topicalização de oblíquo, complementos de nome. Desse total, 4 deles reportam um diálogo por telefone entre os personagens Dolores e "o cara da distribuidora de filmes". Durante a conversa, os personagens Margareth e Gabriel insistem que Dolores faça perguntas sobre os filmes que serão entregues. Temos, mais uma vez, a alternância entre presença *x* ausência da preposição *de*, encontrada para o verbo *esquecer*, em Miguel Falabella. Neste caso, essa alternância pode se justificar pela produção em sequência dos dados pela mesma personagem.

(71) MARGARETH – Pergunta aí se não vem filme da Kim Novak, Dolores?
DOLORES – *[Da Kim Novak]$_i$ não vem nada* ____$_i$*?* Não, não, estou ouvindo sim. Pode continuar. No calor da noite. Nossa, esse dever ser bárbaro. (*No coração do Brasil*)

(72) GABRIEL– *[Da Doris Day]$_i$, não vem nada* ____$_i$*?* (*No coração do Brasil*)

(73) MARGARETH – *[E da Kim Novak]$_i$ não vem nada* ____$_i$*?* (*No coração do Brasil*)

(74) DOLORES– Ele disse que a Kim Novak tá de férias. *[E filme de marinheiro]$_i$ não vem nada* ____$_i$*.* Ou então de guerra! (Ela tapa o bocal e fala para os

outros). Eu sou doida numa farda. O quê? O senhor me desculpe, mas isso eu não sei repetir não. Vem cá, Margareth, vê se tu entende. (*No coração do Brasil*)

(75) DOLORES – *[Nesse ponto]$_i$, o Nilson tem razão* ____$_i$. Juntar os panos quando a gente não tem nada é uma besteira. E essas histórias terminam tudo igual a Neiva. Deus que me perdoe, terminar igual a Neiva! Tendo que fugir pra cá, quando a menina dorme, tendo que fazer comida pra aquele cachorro num fogareiro de duas bocas. Eu tenho horror a fogareiro. Eu posso morar embaixo da ponte, mas levo meu fogão comigo. (*No coração do Brasil*)

(76) DOLORES - (...) E depois lá em casa não tem espaço. Não vou jogar a criança na sala. Eu tenho um fogão de seis bocas na sala. Não cabe na cozinha. *E [do meu fogão]$_i$ eu não abro mão* ____$_i$. De jeito nenhum. (*No coração do Brasil*)

Nesta sequência de 4 dados de topicalização de oblíquo, todos projetados pelo pronome indefinido *nada,* apenas o (74) tem a supressão da preposição *de*. Os outros três apresentam a manutenção da preposição. Em (75), o Sintagma Tópico *Nesse ponto* é complemento da estrutura *ter razão* que solicita preposição *em*. Em (76), temos a construção *abrir mão de,* estando o Sintagma Tópico introduzido pela preposição *de*. Parece-nos que a gramática de Miguel Falabella, o mais jovem dos escritores que compõem a amostra dessa pesquisa, evidencia bem o comportamento do PB, alternando entre a presença e a ausência de preposições, sejam eles com mais ou menos conteúdo semântico.

4. Considerações finais

Esta pesquisa teve como objetivo verificar qual ou quais gramáticas pode(m) ser instanciada(s) na escrita de autores brasileiros nascidos entre os séculos XIX e XX, ao focalizar os dados de topicalização de oblíquo nuclear produzidos por 9 autores brasileiros nascidos no decorrer dos séculos XIX e XX, em um conjunto de 20 peças teatrais. Em 8 dos 9 autores, encontramos ocorrências do fenômeno em estudo. Para diferenciar PE de PB, tomamos como referencial descritivo as gramáticas portuguesas de Mateus *et al.* (2003) e Raposo *et. al.* (2013).

Partimos da hipótese de que a escrita de autores brasileiros nascidos nos séculos XIX e XX apresentam indícios de competição entre as duas gramáticas do português, PB e PE, nos termos de Kroch (1989, 2021[2001]). Os resultados confirmaram parcialmente nossa hipótese, já que a escrita dos autores nascidos no século XIX parece instanciar

exclusivamente a gramática do PE. É na escrita dos autores nascidos no século XX que vislumbramos a competição de gramáticas

Esta conclusão resulta do comportamento dos dados de topicalização de oblíquo nuclear em relação, principalmente, à possibilidade de supressão de preposição com mais conteúdo semântico. Assim, no PE a supressão de preposição está restrita àquelas que carregam menos conteúdo semântico, enquanto no PB a supressão de preposição com mais conteúdo semântico é licenciada. No quadro abaixo, apresentamos, por autor, a frequência de dados, reunindo tanto os casos de complemento oblíquo de verbo quanto de nome, e o comportamento em relação à preposição, no que tange ao seu apagamento e a sua natureza semântica.

Autor	Ano de nascimento	Topicalização de oblíquo nuclear	
		Com preposição	Sem preposição
Martins Pena	1815	19	-
Joaquim Manuel de Macedo	1820	3	-
França Junior	1838	10	1 [preposição com – menos conteúdo semântico (*em*)]
Gastão Tojeiro	1880	4	-
Armando Gonzaga	1884	1	-
Millôr Fernandes	1923	9	2 [preposições com mais conteúdo semântico (*para* e *com*)]
Carlos Eduardo Novaes	1940	8	2 [preposições com mais conteúdo semântico (*a* e *com*)]
Miguel Falabella	1956	8	2 [preposições com menos conteúdo semântico (*de* em ambos os dados)]

Quadro 2: Distribuição dos dados por autor, segundo o comportamento da preposição

O quadro 2 mostra que, em todos os autores do século XIX, não há supressão de preposição [+ lexicalizada] nas topicalizações de oblíquo nuclear, seja este constituinte um complemento de verbo ou de nome, movido para a periferia esquerda da sentença. Em apenas um dado, produzido por França Junior, houve supressão de preposição com menos conteúdo semântico: exemplo (38).

Isto é um indicador de que os autores brasileiros nascidos no século XIX instanciam a gramática do PE, que não parece licenciar a supressão de preposições com mais conteúdo semântico, nas topicalizações de oblíquo. Esta afirmação encontra respaldo no fato de boa parte dos dados desses autores apresentarem uma sintaxe VS, característica do PE oitocentista, como um resquício do Português Antigo (PA) e do Português Clássico (PC).

Os dados dos autores do século XX esboçam, por sua vez, uma competição entre as gramáticas do PE e do PB, já que nesta última parece haver o licenciamento da supressão de preposição com mais conteúdo semântico, nas construções de topicalização de oblíquo nuclear, ausente naquela. Millôr Fernandes é o primeiro autor que apresenta dados com supressão de preposição com mais conteúdo semântico. Comportamento semelhante encontramos em Carlos Eduardo Novaes, como evidencia o quadro 2.

Os resultados apresentados neste artigo apontam para o fato de a natureza da preposição ser o critério que efetivamente diferencia as gramáticas do PE e do PB, sendo naquela licenciada apenas a supressão de preposição com menos conteúdo enquanto, enquanto nesta existe a possibilidade da supressão de preposição com menos ou mais conteúdo semântico. Os dados com supressão de preposição com mais conteúdo semântico somente surgem no século XX, tendo os autores nascido a partir de 1920, fato que converge para a interpretação de que é neste período que a gramática do PB se consolida, afastando-se da do PE. Ao fenômeno do apagamento da preposição do Sintagma Tópico em topicalizações de oblíquo nuclear somam-se outros, como a inserção do *você* no paradigma pronominal, a fixação da ordem SVO e a preferência pelo preenchimento de sujeitos referenciais definidos, conjunto de mudanças que apontam para o surgimento de uma gramática do PB (cf. Duarte, 1995; Duarte, Freire, Vasco, 2003; Cyrino, Duarte, Kato, 2000; Kato, Duarte, Cyrino, Berlinck, 2006, entre outros).

Esperamos, com este trabalho, contribuir para as discussões acerca da constituição de uma gramática do PB, que se afasta do PE, tendo em vista as mudanças sintáticas por que passou (ou ainda passa).

Referências

ARAUJO, Edivalda A. *As construções de tópico do Português nos séculos XVIII e XIX: uma abordagem sintático-discursiva*. Tese de doutoramento. Salvador: UFBA: Instituto de Letras, 2006.

BERLINCK, R. A.; COELHO, I. L. A ordem do sujeito em construções declarativas na história do Português Brasileiro. In: CYRINO, S. M. L.; TORRES MORAIS, M. A. (Org.). *História do Português Brasileiro. Mudança sintática do Português Brasileiro: perspectiva gerativa.* V. 6. São Paulo: Contexto, 2018, pp. 308-381.

BERLINK. R. A.; COELHO, I. L.; CYRINO, S. M. L.; DUARTE, M. E. L.; MARTINS, M. A. Mudança sintática e a história do PB nos séculos XIX e XX. In: ARAÚJO SÁ JR., L.; MARTINS, M. A. (Org.). *Rumos da Linguística Brasileira no Século XXI: historiografia, gramática e ensino*. São Paulo: Blucher Editora, 2016, v. 1, pp. 155-187. Disponível no link: https://www.blucher.com.br/livro/detalhes/rumos-da-linguistica-brasileira-no-seculo-xxi-1219

CHOMSKY, N. *Knowledge of language: its nature, origins and use*. New York: Praeger, 1985.

CYRINO, S.; DUARTE, M. E. L.; KATO, M. Visible subjects and invisible clitics in brazilian portuguese. In: KATO, M; NEGRÃO, E. V. (Org.). *Brazilian portuguese and the null subject*. Frankfurt am Main: Vervuert Verlag, 2000. p. 55-73.

DECAT, M. B. Nascimento. Construções de tópico em português: uma abordagem diacrônica à luz do encaixamento no sistema pronominal. In: TARALLO, F. (org.) *Fotografias sociolinguísticas*. Campinas, SP: Pontes: Editora da UNICAMP, 1989, pp. 113-137.

DUARTE, Maria Eugenia L. *A perda do princípio Evite Pronome no português brasileiro.* Tese de Doutorado. Campinas: Instituto de Estudos da Linguagem / UNICAMP, 1995.

DUARTE, Maria Eugenia L. Do pronome nulo ou pronome pleno: a trajetória do sujeito no português do Brasil. In: KATO, M. A; ROBERTS, Ian. (org.). *Português brasileiro: uma viagem diacrônica*. 2ª ed. Campinas, SP: Editora da UNICAMP, 1996. pp. 107-128.

DUARTE, Maria Eugenia L. Os sujeitos de 3ª pessoa: revisitando Duarte 1993. In: DUARTE. M. E. L. (org.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Editora Parábola, 2012.

DUARTE, Maria Eugênia L.; FREIRE, Gilson Costa, VASCO, Sérgio Leitão. Português europeu e português brasileiro: alguns aspectos morfossintáticos. In: HENRIQUES, C. (org.). *Linguagem, Conhecimento e Aplicação*. Rio de Janeiro: Europa, 2003, pp. 253-266.

FERREIRA, Renato Kledson. *Construções de tópico marcado em cartas pessoais brasileiras dos séculos XVIII, XIX e XX*. Dissertação de Mestrado. Natal, RN: Universidade Federal do Rio Grande do Norte, 2014.

KATO, M.; DUARTE, M. Eugenia L.; CYRINO, Sonia; BERLINCK, Rosane. Português brasileiro no fim do século XIX e na virada do milênio. In: CARDOSO, S; MOTA, J.; MATTOS E SILVA, R. V. (Org.). *Quinhentos anos de história linguística do Brasil*. Salvador: Empresa Gráfica da Bahia / Funcultura / Governo do Bahia. 2006, pp. 413-438.

KROCH, Anthony. *Reflexes of grammar in patterns of language change*: *Language Variantion and change*. v. 1, 1989.

KROCH, Anthony. Mudança Sintática. Tradução de Silvia Regina de Oliveira Cavalcante e Anna Lyssa do Nascimento Donato Machado. In: *Work. Pap. Linguíst.,* 22(2), Florianópolis, 2021.

LOPES, Celia. Pronomes pessoais. In: VIEIRA, S.; BRANDÃO, S. (Org.). *Ensino de gramática: descrição e uso*. SP: Contexto, 2007. pp. 103-119.

LUFT, Celso Pedro. *Dicionário prático de regência nominal*. 4ª edição, 2ª impressão, São Paulo: Ática, 2002.

LUFT, Celso Pedro. *Dicionário prático de regência verbal*. 9ª edição, 8ª impressão. São Paulo: Ática, 2010.

MATEUS, Maria Helena Mira *et alii*. *Gramática da língua portuguesa*. Lisboa: Editorial Caminho, 2003, pp. 433-506.

MENDÍVIL-GIRÓ, José Luis. Por que as línguas não se adaptam ao seu ambiente? *ReVEL*. V. 17, n. 32, 2019. pp. 18-42. Tradução de Gabriel de Ávila Othero e Mariana Terra Teixeira.

RAPOSO, Eduardo *et alii*. *Gramática do Português*. v. 1. Lisboa: Fundação Calouste Gulbenkain, 2013.

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aquém-mar e d'além-mar ao final do século XIX. In: KATO, M. A; ROBERTS, Ian. (org.). *Português brasileiro: uma viagem diacrônica*. 2ª ed. Campinas, SP: Editora da UNICAMP, 1996. pp. 69-105.

TORRES MORAIS, Maria Aparecida C. R. de. Aspectos diacrônicos do movimento do verbo, estrutura da frase e caso nominativo no português do Brasil. In: KATO, M. A; ROBERTS, Ian. (org.). *Português brasileiro: uma viagem diacrônica*. 2ª ed. Campinas, SP: Editora da UNICAMP, 1996. pp. 263-306.

Anexo 1 – Referências das peças teatrais que compõem a amostra

FALABELLA, Miguel. *No coração do Brasil*. SBAT (Sociedade Brasileira de Autores Teatrais), 1992.

FERNANDES, Millôr. *Um elefante no caos*. Porto Alegre: L&PM Editores, 1979.

FERNANDES, Millôr. *Os órfãos de Jânio*. Porto Alegre: L&PM Editores, 1980.

FRANÇA JUNIOR, Joaquim J. de. *Defeito de família*. SBAT (Sociedade Brasileira de Autores Teatrais), 1870.

FRANÇA JUNIOR, Joaquim J. de. Como se fazia um deputado. In: CAFEZEIRO, E. *et al.* (Org.). *Teatro de França Junior.* Tomo II. Rio de Janeiro. Serviço Nacional de Teatro. Fundação da Arte, 1980, pp. 123-168.

FRANÇA JUNIOR, Joaquim J. de. Caiu o ministério. In: CAFEZEIRO, E. *et al.* (Org.). *Teatro de França Junior.* Tomo II. Rio de Janeiro. Serviço Nacional de Teatro. Fundação da Arte, 1980, pp. 169-221.

FRANÇA JUNIOR, Joaquim J. de. As doutoras. In: CAFEZEIRO, E. *et al.* (Org.). *Teatro de França Junior.* Tomo II. Rio de Janeiro. Serviço Nacional de Teatro. Fundação da Arte, 1980, pp. 223-291.

GONZAGA, Armando. *O hóspede do quarto nº 2*. SBAT (Sociedade Brasileira de Autores Teatrais), 1937.

MACEDO, Joaquim Manuel de. *Luxo e Vaidade*. SBAT (Sociedade Brasileira de Autores Teatrais). Escrita em 1860.

MARTINS PENA, Luiz C. O juiz de paz na roça. In: *As melhores comédias de Martins Pena*. Apresentação de Guilhermino César. Porto Alegre: Mercado Aberto. 1987. pp. 9-28.

MARTINS PENA, Luiz C. Judas em Sábado de Aleluia. In: *As melhores comédias de Martins Pena.* Apresentação de Guilhermino César. Porto Alegre: Mercado Aberto. 1987. pp. 151-176.

MARTINS PENA, Luiz C. Quem casa quer casa. In: *As melhores comédias de Martins Pena.* Apresentação de Guilhermino César. Porto Alegre: Mercado Aberto. 1987. pp. 177-196.

MARTINS PENA, Luiz C. O noviço. In: *As melhores comédias de Martins Pena.* Apresentação de Guilhermino César. Porto Alegre: Mercado Aberto. 1987. pp. 73-118.

MARTINS PENA, Luiz C. As casadas solteiras. In: *As melhores comédias de Martins Pena.* Apresentação de Guilhermino César. Porto Alegre: Mercado Aberto. 1987. pp. 29-72.

NOVAES, Carlos Eduardo. *A mulher integral*. Biblioteca da UNIRIO, 1975.

SAMPAIO, Silveira. *A garçonière de meu marido*. SBAT (Sociedade Brasileira de Autores Teatrais), 1949.

SAMPAIO, Silveira. *Flagrantes do Rio*. SBAT (Sociedade Brasileira de Autores Teatrais), 1945.

TOJEIRO, Gastão. *O simpático Jeremias*. SBAT (Sociedade Brasileira de Autores Teatrais), 1918.

TOJEIRO, Gastão. *Onde canta o sabiá*. SBAT (Sociedade Brasileira de Autores Teatrais), 1920.

TOJEIRO, Gastão. *A inquilina de Botafogo*. SBAT (Sociedade Brasileira de Autores Teatrais), 1920.

Anexo 2 – Quadro com os títulos das peças que compõem a amostra

Autor	Peça
Martins Pena	O juiz de paz na roça
	Judas em Sábado de Aleluia
	Quem casa quer casa
	O noviço
	As casadas solteiras
Joaquim Manuel de Macedo	Luxo e vaidade
França Junior	Defeito de família
	Como se fazia um deputado
	Caiu o ministério
	As doutoras
Gastão Tojeiro	O simpático Jeremias
	Onde canta o sabiá
	A inquilina de Botafogo
Armando Gonzaga	O hóspede do quarto nº 2
Silveira Sampaio	A garçoniére de meu marido
	Flagrantes do Rio
Millôr Fernandes	Um elefante no caos
	Os órfãos de Jânio
Carlos Eduardo Novaes	A mulher integral
Miguel Falabella	No coração do Brasil

CC BY NC

http://doi.org/10.5007/1984-8420.2025e101319

"REDES SOCIAIS E SOCIOLINGUÍSTICA HISTÓRICA": VARIAÇÃO DOS PRONOMES DE SEGUNDA PESSOA EM CARTAS (1956-1994)

"SOCIAL NETWORKS AND HISTORICAL SOCIOLINGUISTICS": VARIATION IN SECOND-PERSON PRONOUNS IN LETTERS (1956-1994)

Célia Lopes | Lattes | celiar.s.lopes@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Janine da Silva Barbosa | Lattes | janinebarbosa@letras.ufrj.br
Universidade Federal do Rio de Janeiro

João Pedro de Carvalho Gonçalves da Silva | Lattes | joaopedrocarvalho@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Mariana da Silva Soares | Lattes | marianasoares@letras.ufrj.br
Universidade Federal do Rio de Janeiro

Resumo: O objetivo do estudo é analisar se a difusão das formas de *você* já estava consolidada na segunda metade do século XX, como mostrou Souza (2012), em cartas do Rio de Janeiro. A amostra é composta por 129 cartas escritas no período de 1956-1994 e trocadas entre membros de uma família, cuja filha passou a viver em Paris. O estudo se baseia nos pressupostos teóricos da *Sociolinguística Histórica* (Conde Silvestre, 2007), da teoria do *Poder e Solidariedade* (Brown e Gilman, 1960) e incorpora o conceito de *Redes Sociais* na explicação da mudança em sincronias passadas (Bergs, 2005). Os resultados mostraram que o padrão canônico *tu-te-ti-contigo,* mais conservador, ficou patente nas cartas da matriarca da família (E) que participava de redes sociais mais densas. As formas também uniformes *você-o-lhe-com você* prevaleceram nas cartas do pai (W), atuante em redes um pouco menos densas do que a esposa. Diferentemente dos pais, a filha (A), uma "cidadã do mundo" e pertencente a redes sociais bastantes difusas e abertas, adota o *padrão híbrido* (*você-te-com você)* vigente na fala do Rio de Janeiro atualmente.

Palavras-chave: Redes sociais; Formas de tratamento; Cartas cariocas.

http://doi.org/10.5007/1984-8420.2025.e101319

Abstract: The aim of the study is to analyze whether the diffusion of the forms of "você" (you) was already consolidated in the second half of the 20th century, as shown by Souza (2012), in letters from Rio de Janeiro. The sample consists of 129 letters from 1956-1994 exchanged between members of a family, whose daughter moved to live in Paris. The study is based on the theoretical assumptions of Historical Sociolinguistics (Conde Silvestre, 2007), the theory of Power and Solidarity (Brown and Gilman, 1960), and incorporates the concept of Social Networks in explaining change in past synchronies (Bergs 2005). The results showed that the *canonical pattern tu-te-ti-contigo*, more conservative, was evident in the letters of the family's matriarch (E) who participated in denser social networks. The also uniform forms *você-o-lhe-com você* prevailed in the father's (W) letters, who was involved in networks somewhat less dense than his wife. Unlike the parents, the daughter (A), a "citizen of the world" and belonging to quite diffuse and open networks, adopts the *hybrid pattern* (*você-te-com você*) currently prevailing in Rio de Janeiro's speech.

Keywords: Social networks. Address forms. Rio de Janeiro letters.

1 Introdução

O objetivo inicial do estudo é analisar se a difusão das formas do paradigma de *você* no lugar das formas de *tu*, como apontou Souza (2012), já estava consolidada na segunda metade do século XX. Para tanto, foram analisadas 129 cartas de uma família brasileira, denominada apenas de FB para a garantia da anonimidade dos remetentes. As cartas foram produzidas entre 1956-1994 e o eixo Rio de Janeiro-Paris-Rio de Janeiro é o principal ponto de referência dessas cartas trocadas entre familiares, em particular, as enviadas dos pais para a filha que residia fora do país.

A escolha de materiais produzidos especificamente na segunda metade do século XX se justifica pelo fato de buscarmos confirmar os resultados de Souza (2012). A autora assinala que a implementação, no Rio de Janeiro, de *você-sujeito* como um efetivo pronome de segunda pessoa (2SG) se deu a partir de 1940. Antes disso, Souza (2012) detecta duas fases de transição com distintos comportamentos. A primeira fase corresponderia ao período de 1870-1900 em que *tu* se sobrepunha a *você* nas cartas estudadas pela autora. Em um segundo momento, entre 1901-1939, a autora observa uma coexistência acirrada entre as duas formas variantes. Só a partir de 1940 é que as frequências de *você* suplantam significativamente o emprego de *tu* e a forma inovadora de segunda pessoa

(*você*) se estabelece nas cartas pessoais estudadas em quase todas as relações estabelecidas entre remetente-destinatário (simétricas e assimétricas).

A análise de Souza (2012) é nosso ponto de partida, mas é preciso considerar que a autora analisou um conjunto reduzido de cartas da segunda metade do século XX, o que pode fragilizar, de certa forma, suas conclusões. No estudo proposto aqui, incorporamos materiais novos editados pela equipe do projeto *HistLing – Corpus Histórico da Língua Portuguesa*: as cartas da família FB, cujos remetentes não são pessoas conhecidas. Por essa razão, outro intuito do trabalho é recuperar a história social dessa família a partir de pistas textuais das próprias fontes documentais sob análise e correlacionar esses aspectos histórico-sociais com o comportamento linguístico adotado.

Para dar conta desse outro objetivo, estamos levando em conta, em termos teóricos, os pressupostos da *Sociolinguística Histórica* (Conde Silvestre, 2007) e da teoria do *Poder e Solidariedade* (Brown; Gilman, 1960). Dá-se destaque ainda à discussão da pertinência do conceito de Redes Sociais para explicação da mudança em sincronias passadas (Bergs, 2005). Em termos metodológicos, será utilizada a ferramenta estatística GOLDVARB-X para a quantificação dos dados variantes de segunda pessoa do singular em suas diferentes funções: nominativa (ou sujeito) (*tu/você);* acusativa (*te~você~lhe~o/a*), dativa *(te~-lhe~para/a você/ti*), oblíqua (prep. *você~ prep. ti/contigo*) e genitivo (*teu~seu*).

Além de confirmar a hipótese de Souza (2012) sobre a implementação de *você* nas relações mais pessoais a partir da segunda metade do século XX, buscamos nessa investigação trazer alguns problemas gerais e específicos sobre o tema.

Em termos mais amplos, as perguntas a serem respondidas são: (i) De que maneira a reconstrução da história social de uma família desconhecida, a partir de sua produção escrita, ajuda a compreender os seus usos linguísticos? (ii) O comportamento linguístico mais ou menos inovador dos membros da família FB está relacionado às redes sociais (mais ou menos) densas e abertas das quais cada um participava?

Em termos especificamente linguísticos, nossas questões são: (iii) As formas do paradigma pronominal "híbrido/misto" de 2SG (*você~tu*), descritas em Lopes *et al.* (2018) para o Rio de Janeiro, se fazem notar na amostra e no período em análise? (iv) Que paradigmas de segunda pessoa são mais produtivos entre os remetentes de uma família abastada da sociedade carioca de fins do século XX: *paradigmas canônicos*, tais como *tu-te-ti-contigo-teu*; *você-o/a-lhe-com você-seu;* ou *paradigmas híbridos* do tipo *você(~tu)-te-com você-seu/teu*?

O artigo está assim estruturado. Além dessa introdução, na seção 2 situamos brevemente o estado da questão a partir de estudos sobre o comportamento de *você* e *tu* em

cartas do Rio de Janeiro produzidas no decorrer do século XX. A seção 3 traz os postulados teóricos adotados e a descrição da amostra. Na seção 4, são apresentados os resultados obtidos que foram organizados da seguinte maneira: primeiramente, descrevemos a distribuição das formas de segunda pessoa em suas diferentes realizações sintáticas já listadas; na sequência, discutimos os resultados das formas de 2SG na posição de sujeito, observando as relações de *Poder* e *Solidariedade*. Por fim, tentaremos interpretar os resultados obtidos a partir das Redes Sociais do pai-mãe-filha (re)construídas a partir de pistas textuais detectadas no material em análise. As conclusões constam da seção 5 que é seguida pelas referências utilizadas.

2 O estado da arte: a difusão de *você* em cartas escritas no Rio de Janeiro (1870-1979)

Os trabalhos de Souza (2012) e de Lopes e Souza (2018) descreveram a distribuição de *você* e *tu* na posição de sujeito em 366 cartas produzidas por diferentes grupos de famílias que viviam no Rio de Janeiro entre 1870 e 1979. Esses estudos analisaram 1525 dados com uma distribuição geral equilibrada (763 dados de *você* e 762 de *tu*) como sujeito pleno ou nulo. A diferença de comportamento das formas variantes de 2SG identificada ao longo de 100 anos foi interpretada pelas autoras como etapas de um percurso de mudança. Em um primeiro momento, nos idos de 1870 até a virada do século XIX, o pronome *tu* era mais produtivo que a forma *você*. Em uma segunda fase, que vai do início do século XX até 1939, as duas formas coexistem com frequências aproximadas (em torno dos 50%), o que se altera a partir dos anos 1940, quando *você* se torna majoritário, atingindo quase 100% de frequência em 1979.

Além do panorama geral das formas variantes, os estudos mencionados analisaram como as relações de Poder e Solidariedade (cf. Brown e Gilman, 1960)[1] e o grau de parentesco do remetente-destinatário podem ter influenciado nessa distribuição diferenciada ao longo do tempo.

Segundo a interpretação de Lopes e Souza (2018), o pronome *você* ainda funcionava, nas cartas oitocentistas, como estratégia de poder nas relações *assimétricas ascendentes* (inferior-superior), por não ter perdido completamente os traços de respeito e distân-

[1] Com variação de estratégias linguísticas de língua para língua, mas usando o francês como referência, os autores mostram que, em relações de maior intimidade, solidariedade e proximidade, haveria o favorecimento de *tu* (T), enquanto as relações de menor intimidade e de maior poder entre os participantes favoreceriam o emprego do tratamento menos solidário *vous* (V). Em princípio, no português do século XIX e início do XX, a oposição seria entre *tu* e *você* (ou formas de base nominal como *o senhor*), mas nas sociedades contemporâneas essa dicotomia tem se tornado mais fluida e menos marcada.

cia social do tratamento abstrato *Vossa Mercê* (Koch, 2008). No decorrer do século XX, contudo, a variante *você* foi deixando de ser uma estratégia de tratamento que marcava *poder* e passou a funcionar no eixo da *solidariedade*. Tal comportamento acompanhou a dinâmica das sociedades contemporâneas que seriam mais abertas e igualitárias do que se observara no Brasil Colônia.

Os resultados obtidos mostraram ainda que o gênero e a idade do missivista também foram fatores atuantes quando correlacionados ao grau de parentesco nas *relações ascendentes* (inferior-superior). Os estudos identificaram menor frequência de *você* nas cartas trocadas entre filho-pai (8%) e seu emprego categórico quando o destinatário era uma mulher mais velha, como uma tia ou mãe. Nos dados do início do século XX, o uso de *tu* não era comum nesse contexto assimétrico para mulheres. As autoras defendem que era mais comum o emprego de *você* ou uma forma mais distante, como *a senhora*, na escrita feminina.

Nas relações *assimétricas descendentes* (superior-inferior), por sua vez, a presença de *tu* seria a forma esperada, principalmente nas cartas de remetentes com maior *poder* como é caso de homens mais velhos em sociedades patriarcais. No estudo com cartas do Rio de Janeiro, os índices de frequência para *você* aumentaram (de 22%, em fins do XIX, para 38%, no início do XX), embora o emprego de *tu* ainda fosse mais produtivo. Nesse tipo de relação (descendente), o gênero do remetente também se mostrou relevante. Nas relações entre homens (pai-filho e avô-neto), prevaleceu o *tu*, ao passo que houve o predomínio de *você* quando o remetente era uma mulher (mãe-filho, tia-sobrinho).

Nas *relações simétricas*, as autoras observaram uma difusão gradativa da forma inovadora *você* no lugar de *tu* nas relações igualitárias, principalmente a partir das cartas escritas por amigos depois de 1930. A frequência de *você* era de 16% em 1870-1899, 75% em 1900-1939 e 99% em 1940-1979. Também aqui se identificou um comportamento distinto entre homens e mulheres. Nas cartas entre casais, os homens utilizavam *tu* e as mulheres, *você*. Nas cartas trocadas entre irmãs, os índices de *você* atingem 63%, entre 1900-1939, enquanto os irmãos apresentam 40% de frequência na mesma fase. Aqui também os remetentes masculinos usaram, ao longo do tempo, mais *tu* do que as remetentes femininas.

Por fim, falta-nos abordar brevemente o comportamento das formas variantes de 2SG nas demais funções sintáticas, além da função nominativa (sujeito). Os estudos de Oliveira (2014), Souza (2014) e Lopes, Marcotulio e Oliveira (2020) mostraram que a difusão de *você* não se deu na mesma velocidade nas funções de complemento acusativo

e dativo. Nesses casos, houve uma retenção de formas do paradigma de *tu* e o clítico *te* se manteve produtivo independentemente da forma que ocupa a posição de sujeito (*você* ou *tu*). Em cartas exclusivas de sujeito *tu* e em cartas em que coexistem os pronomes *tu* e *você* na posição de sujeito, por exemplo, o clítico *te* apresentou produtividade de 70% das ocorrências ou mais. Mesmo nas cartas em que o pronome *você* é exclusivo na posição de sujeito, a taxa de uso do dativo *te* é superior às demais formas variantes seja na função de acusativo (*te, o/a, lhe, você*), seja na de dativo (*te, lhe, para/a você, objeto nulo*). Na função oblíqua (*preposição* + *ti* ou + *você*), predominam as formas de *você* e na função genitiva (*teu* e *seu*), o possessivo *teu* prevaleceu nas cartas cariocas estudadas (cf. Lopes *et al.*, 2018).

No estudo aqui proposto, pretende-se confirmar, principalmente, os resultados relativos à posição de sujeito que sinalizam para uma perda gradativa do uso de *tu* e o espraiamento do inovador *você* nos diferentes tipos de relação (simétricas e assimétricas), principalmente a partir de 1940. Tal marco histórico nos motivou a incorporar uma amostra mais representativa de cartas da segunda metade do século XX, como é o caso das cartas da família FB. O objetivo é confirmar se houve realmente a generalização e consequente *polifuncionalidade*[2] de *você* como uma estratégia neutra para qualquer situação/relação. O controle das outras funções sintáticas não é o foco central do estudo, mas objetiva verificar se as cartas estudadas apresentam *paradigmas canônicos* do tipo *tu-te-ti-contigo-teu*; *você-o-lhe-com você-seu* ou os novos *paradigmas híbridos* de 2SG vigentes no Rio de Janeiro atual: *você~tu-te-com você-teu~seu.*

3 Base teórica e descrição geral da amostra

3.1 A *Sociolinguística Histórica* e o conceito de *redes sociais*

Os estudos no âmbito da Linguística Histórica (LH) procuram explicar as condições gerais da evolução linguística ao longo do tempo (Conde Silvestre, 2007). Para compreender como as línguas mudam, a LH costuma analisar e comparar manifestações textuais de épocas diferentes do passado com vistas a compreender os fatores que determinam a mudança linguística de maneira universal. Nem sempre, entretanto, as análises calcadas na LH dão conta da sociedade em que as mudanças linguísticas ocorreram, foco central da Sociolinguística Histórica (SH) (cf. Nevalainen, 2015). Esta última, além de se

[2] A depender do momento histórico, *você* pode funcionar para marcar algum distanciamento, herança da forma original *Vossa Mercê* e, mais recentemente, como variante de *o senhor/a senhora*, mas é possível identificar tal forma em contextos mais íntimos como uma variante de *tu*. Há contextos em que *você* funciona inclusive como uma estratégia de esquiva ou neutra, quando não se tem certeza da proximidade entre os interlocutores.

ocupar dos princípios gerais da teoria da mudança (objeto da LH), aplica as premissas da Sociolinguística Laboviana (SL) para o estudo e interpretação dos materiais históricos, correlacionando fatores linguísticos e sociais (Gimeno, 1983).

O maior dilema para a reconstrução histórica de uma língua é a natureza da fonte documental que serve de base para os estudos de mudança. Os materiais escritos do passado são esparsos, fragmentários e pouco representativos das comunidades pretéritas (Conde Silvestre, 2007). A documentação escrita fornece dados parciais pertencentes muitas vezes a uma parcela restrita da comunidade, constituída basicamente por alfabetizados que, a depender da época, eram prioritariamente homens de *status* mais alto, restringindo-se a certos estilos e registros.

No Brasil, capitaneados por Ataliba de Castilho (nosso homenageado), investigadores de diversas universidades brasileiras iniciaram, a partir da década de 1990, as pesquisas linguístico-históricas sobre o português *no/do* Brasil. O foco principal do mentor do projeto nacional *Para uma História do Português Brasileiro – PHPB* é reconstruir uma história social linguística do Brasil a partir da descoberta e da análise linguística de documentos não-literários do nosso passado.

De lá para cá, muito se avançou na pesquisa nessa área tanto em termos descritivos quanto em termos teóricos. A Sociolinguística Histórica é uma vertente de investigação que ganhou força, refinando os métodos e ferramentas da Sociolinguística de base laboviana para apreender a variação a partir dos textos escritos do passado (Hernández-Campoy; Conde Silvestre, 2012). Defende-se, como Bergs (2005, p. 21), que a SH é uma "subdisciplina com potencial próprio, que deve desenvolver seus próprios objetivos, metodologias e teorias, divorciadas da sociolinguística atual, por um lado, e da linguística histórica, por outro."

Obviamente, há pontos coincidentes entre a SL e a SH, pois ambas partem de três premissas fundamentais: (i) a observação da *heterogeneidade ordenada*, uma vez que consideram que as línguas não se constituem como uma realidade homogênea e estão em constante processo de variação e mudança sob restrições estruturais e sociais; (ii) o reconhecimento de que toda mudança emerge da variação linguística e (iii) a atuação das variáveis linguísticas é depreendida a partir da análise de *corpora* reais de uso e representativos de uma comunidade.

Entretanto, por lidar apenas com materiais escritos parciais e limitados que sobre-

viveram por sorte, a SH nem sempre tem a seu dispor um *corpus* histórico representativo de uma dada comunidade do passado. É preciso que o investigador esteja consciente dos entraves que enfrentará e para os quais deverá propor soluções e uma metodologia adequada, pois nem sempre se dispõe, por exemplo, de uma variedade de estilos e registros. As amostras são naturalmente enviesadas e os dados do passado são "meros restos de *corpus* textuais muitíssimo mais amplos" e "dificilmente vinculáveis à produção real de seus falantes" (Conde Silvestre, 2007, p. 35-36). O controle de fatores externos, como faixa etária, sexo, nível de escolaridade, classe social, entre outros, será lacunar e o investigador poderá recuperar outros dados sociais relevantes a partir de pistas textuais, como pretendemos mostrar nesse estudo.

Para minimizar esses problemas, tem sido bastante eficaz incorporar aos preceitos da Sociolinguística Histórica o conceito de *redes sociais*, discutido desde J. Milroy e L. Milroy 1992, na perspectiva adotada por Bergs (2005) para dados do passado. O conceito de rede prevê diferentes padrões estruturais organizados a partir da sua *densidade* e dos *laços* que podem influenciar no comportamento de seus membros individuais nos processos de difusão de mudanças linguísticas. Se essa proposta é pertinente para o presente, certamente falantes/escreventes do passado também participariam de diferentes estruturas de rede que poderiam influenciar no desenvolvimento de algumas mudanças historicamente comprovadas. A identificação do pertencimento a redes sociais mais ou menos abertas pode ser mais eficaz do que a recuperação de determinados fatores estáveis relacionados ao falante, como idade, classe e educação. Os falantes/escreventes estabelecem contatos (*rede social pessoal)* com outras pessoas na comunidade em que vivem (cf. Milroy 1992). As redes não são entidades fixas e estáveis, mas agregados de relacionamentos relativamente flexíveis pelo número de contatos entre os seus membros e pela qualidade desses contatos na comunidade.

Como não dispomos de informações precisas da biografia dos autores das cartas da família FB a serem analisadas, propomos reconstruir a história social da família à luz da Sociolinguística Histórica, incorporando premissas do conceito de redes sociais que serão discutidas em 4.3.

3.2 A história social da família em análise: o perfil dos autores

O *corpus* é constituído por 128 cartas de uma família brasileira, identificada aqui apenas como FB para garantir a anonimidade de seus membros. Os atores centrais que

escreveram a maior parte das cartas são o pai (W), a mãe (E) e a filha (A). Eles viviam em Copacabana, no Rio de Janeiro, até que a filha vai estudar em Paris e lá se estabelece. As cartas, escritas entre 1956 e 1994, retrataram prioritariamente esse período do afastamento da filha (A), mas há outros remetentes que compõem a amostra: a mãe de (W), os seus irmãos e os sobrinhos que viviam em Belém do Pará.

Embora não tenhamos certeza do local de nascimento dos remetentes, é provável que a filha (A) seja mesmo natural do Rio de Janeiro, pois ela diz, em dado momento, "sou realmente carioca demais para poder passar mais de 3 meses sem escutar o barulho do mar e sentir o sol no meu corpo" (carta 55 FB 25-04-1975). Os pais nasceram no primeiro quartel do século XX e a filha possivelmente nasceu depois de 1940. O pai (W) tinha uma alta patente militar[3], mas foi preso e julgado pelo Supremo Tribunal Militar[4], em 1965.

A mãe de A, identificada como (E), era católica praticante. Além de expressar textualmente sua fé, menciona sempre as suas idas à missa e a saudade que tem de suas amigas de paróquia quando passa uma temporada com a filha (A) em Paris por ocasião do nascimento de suas duas netas. De lá, escreve a seu marido se queixando do frio, do comportamento um tanto arrogante da filha, das idas à missa, do fato de não saber francês e da vontade de retornar ao Brasil.

As cartas retratam bem o período em que a filha (A) viveu na França. Ela se muda para a capital francesa nos anos de 1970 com o intuito de estudar. Começa a cursar um "Doctorat d'Etat" (Carta 23-FB-25/10/1971), mas assume alguns empregos informais, ora fazendo figuração em alguns filmes, ora como auxiliar de produção. Nesse contexto, a filha (A) se casa com um cineasta francês (D) e tem duas filhas (G e M), na cidade onde passou a viver. Sua primeira carta aos pais foi escrita em 1970 e a última em 1994, assim a amostra FB permite que se realize um estudo do comportamento de uma mulher nascida no Rio de Janeiro (*Panel Study*) ao longo de sua juventude e da vida adulta, pois são 20 anos de escrita da mesma pessoa (cf. Rumeu, 2008; Rumeu; Callou, 2021).

Na sequência, serão apresentados os resultados quantitativos gerais.

[3] Carta 05 de 13-09-1962, a mãe de W diz "ainda devia estár vivo o teu pai para faser parte da minha alegria, se quando ele recebeu noticia tinhas recebido a divisa de Cabo ele ficou convencido, já mais agora vendo o nosso filho ocupar uma alta patente como estás ocupando mas Deus não quis"

[4] O irmão faz menção ao acontecido na carta 10 de 20/11/1965. Na carta 12 de 16-01-1966, uma prima fala da sua prisão.

4 Análise geral dos resultados das cartas da família FB: 2a metade do século XX

4.1 Distribuição geral dos dados quantificados

No total foram levantados 532 dados nas 129 cartas da família FB, como se observa na tabela a seguir. As formas de referência ao interlocutor levantadas foram organizadas a partir do paradigma ao qual pertenciam. Assim temos, por exemplo, formas do paradigma de *você* em suas diferentes funções, ou seja, como sujeito (*você*), como acusativo (*você, lhe, o/a*), como dativo (*a/para você, lhe*), como oblíquo (*preposição + você*), como possessivo (*seu/sua*). No paradigma de *tu*, as variantes seriam *tu* ou verbo com desinência 2ª pessoa, *te* como acusativo, *para/a ti* como dativo, *preposição + ti* e *contigo* como oblíquo e *teu/tua* como possessivo. As formas dos outros paradigmas seguem a mesma lógica dos que foram mencionados. Os resultados estão na tabela 1:

Tabela 1: Distribuição geral dos dados de 2SG nas cartas da família FB por função

	FORMAS LEVANTADAS NA AMOSTRA A PARTIR DOS PARADIGMAS DE *VOCÊ, TU* E *O SENHOR*			
FUNÇÕES	**Você**	**Tu**	**O senhor**	**Terceira pessoa (sem referência)**
Sujeito	111/243 46%	**120/243** **49%**	6/243 2,5%	6/243 2,5%
Acusativo	16/43 37,2%	**27/43** **62,8%**	-	-
Dativo	24/69 34,8%	**39/69** **56,5%**	1/69 1,4%	5/69 7,2%
Oblíquo	**27/44** **61,4%**	15/44 34,1%	2/44 4,5%	-
Possessivo	55/133 41,4%	**73/133** **54,9%**	-	5/133 3,8%
Total	**233/532** **43,8%**	**274/532** **51,5%**	**9/532** **1,7%**	**16/532** **3%**

Fonte: Levantamento feito pelos próprios autores.

Como era de se esperar em um conjunto de cartas familiares mais pessoais, as formas do paradigma de *você* e *tu* apresentaram frequências bastante próximas com leve predomínio de *tu* (51,5%) sobre *você* (43,8%) no cômputo geral. As outras estratégias foram bastante marginais na amostra e não atingiram 5% de uso: 3% para o verbo na terceira pessoa sem referência explícita no contexto da carta e 1,7% para o tratamento mais distante *o/a senhor/a.*

A predominância de formas de *tu* na amostra era, de certa forma, esperada para as funções acusativas (62,8%) e dativas (56,5%) nesse *corpus* de cartas do Rio de Janeiro, como bem mostraram os trabalhos de Souza (2014) e Oliveira (2014). Esses estudos já demonstraram que houve retenção do clítico *te* mesmo quando o falante ou escrevente emprega *você* na posição de sujeito, em construções do tipo ***você*** *quer que* ***te*** (DAT) *dê o livro* ou ***você*** *disse que* ***te*** (ACUS) *vi*. Entretanto, chama atenção, nos resultados da amostra de cartas do Rio de Janeiro em análise, a maior produtividade do sujeito *tu* (49%) em relação ao sujeito *você* (46%) em cartas da segunda metade do século XX. Nesse período, como já mostrou Souza (2012), a forma pronominal *você* já teria suplantado o emprego de *tu* em cartas produzidas na maior parte do Brasil, principalmente no Rio de Janeiro e, desse modo, esperava-se que a difusão de *você* na posição de sujeito estivesse consolidada como previsto em nossas hipóteses descritas na seção 2.

4.2 As formas de 2SG na posição de sujeito e as relações sociais de Poder e Solidariedade

Para compreender por que os resultados foram um tanto destoantes do que fora apontado por Souza (2012), buscou-se identificar possíveis motivações sociais que ajudariam a explicar o comportamento encontrado nessas cartas da família FB. Para tanto, analisamos os fatores tradicionalmente utilizados nos estudos sobre formas de tratamento, controlando o tipo de relação social estabelecida entre remetente e destinatário a partir do modelo clássico de Poder e Solidariedade (Brown e Gilman, 1960). Atrelado a isso, observamos o grau de parentesco estabelecido entre remetente e destinatário e, em um segundo momento, recuperamos as redes sociais das quais os principais agentes participavam. O objetivo é tentar responder às novas perguntas:

i. Por que essas cartas da segunda metade do século XX ainda apresentam tantos dados de *tu* como sujeito?
ii. Essa alta frequência de *tu* permeia todas as cartas da família ou está circunscrita a alguns remetentes mais velhos?
iii. Afinal, quem são os remetentes que ainda empregam *tu* nessas cartas mais recentes (da segunda metade do século XX) e em que relações?

Desse modo, na tabela 2 são apresentados os resultados das formas de tratamento empregadas somente na posição de sujeito para responder essas questões. O grau de parentesco entre os escreventes foi agrupado por relações simétricas, assimétricas descendentes (superior-inferior) e assimétricas ascendentes (inferior-superior) para verificarmos se as relações de poder atuaram nesses resultados:

Tabela 2: Distribuição das formas de 2SG nas cartas da família FB por tipo de relação e parentesco

RELAÇÕES	PARENTESCO	FORMA CONCRETA REALIZADA NA POSIÇÃO DE SUJEITO			
		Você	**Tu**	**Senhor**	**Terceira pessoa**
Simétricas	Marido-Mulher	**28/30 – 93.3%**	2/30 – 6.7%	-	-
	Mulher-Marido	9/81 – 11.1%	**72/81 – 88.9%**	-	-
	Irmãos	-	**28/28 – 100%**	-	-
	Primos	-	-	-	2/2 – 100%
	Amigos	**6/7 – 85.7%**	1/7 – 14.3%	-	-
Assimétricas descendentes (superior-inferior)	Mãe-Filho	-	**14/14 – 100%**	-	-
	Pais-Filho(a)	**39/40 – 97.5%**	1/40 – 2.5%	-	-
Assimétricas ascendentes (inferior-superior)	Filha-Pai	**12/15 – 80%**	-	1/15 – 6.7%	2/15 – 13.3%
	Filha-Mãe	**16/18 – 88.9%**	2/18 – 11.1%	-	-
	Sobrinho-Tio	-	-	-	2/2 – 100%
	Sobrinho-Tia	-	-	3/3 – 100%	-
	Neto-Avô	1/3 – 33.3%	-	**2/3 – 66.7%**	-

Fonte: Levantamento feito pelos próprios autores.

Nas relações simétricas, estudos anteriores defendem que o *tu* predominaria nas relações mais íntimas entre casais, principalmente na escrita masculina, como mostrou o estudo de Souza (2012) para o fim do século XIX e início do XX. As esposas empregavam mais o *você* no trato com os seus esposos por um vestígio de respeito que essa forma trouxe do *Vossa Mercê* do século XIX. No *corpus* da família FB em análise, o principal casal da família (E-esposa e W-marido) apresentou, entretanto, um comportamento completamente inesperado quanto ao uso das variantes de 2SG. Enquanto o marido empregava mais *você* (93,3%), a esposa utilizou *tu* em 88,9% dos casos. Os outros remetentes que se enquadram nas relações simétricas, como os irmãos e amigos, seguem o que estava previso nas hipóteses formuladas a partir de estudos anteriores. Entre irmãos, principalmente homens, prevaleceu o *tu* de intimidade com 100% de frequência e, entre amigos, houve maior uso de *você,* com 85,7%.

Quando analisamos as relações assimétricas, quer as descendentes quer as ascendentes, fica nítido que outros fatores externos, como as relações de Poder e Solidariedade, devem ser considerados para que se compreenda as escolhas tratamentais da família FB.

Nas relações assimétricas ascendentes observadas no *corpus,* houve maior uso de *tu* nas cartas de mãe-filho/a (100%) e maior emprego de *você* nas cartas de pai-filho/a

(97,5%). Esses resultados são interessantes, porque reiteram a mesma polarização entre *tu* e *você* identificada nas relações simétricas entre casais. Esses indivíduos, no caso (E) e (W), mantiveram o mesmo comportamento linguístico quanto às formas de tratamento empregadas em suas cartas, ou seja, a remetente feminina (E) emprega *tu* enquanto o remetente masculino (W) usa preferencialmente *você* independentemente da relação de poder estabelecida entre eles e seus destinatários.

Aparentemente, a filha (A) também mantém o mesmo comportamento. Ela emprega prioritariamente *você,* seja escrevendo ao pai *W* (80%) ou à mãe *E* (88,9%), nas relações ascendentes. Os outros autores das cartas que estabelecem com seus destinatários relações assimétricas ascendentes, entretanto, seguem os modelos previstos para o tipo de relação controlado, empregando formas de distanciamento social como *o/a senhor/a*, nas cartas trocadas entre neto-avô (67,7%) e sobrinho-tia (100%), ou formas não-marcadas (terceira pessoa), nas cartas de sobrinho-tio (100%).

4.3 Interpretação dos resultados gerais à luz das redes sociais

O controle das relações de Poder e Solidariedade, com base nos parentescos, além de confirmar algumas de nossas hipóteses sobre o avanço de *você* em relações pessoais mais íntimas a partir da segunda metade do século XX, evidenciou comportamentos individuais de alguns remetentes, principalmente da remetente mais velha (E). O conservador *tu* presente, de maneira inusitada, nas cartas de (E), seja naquelas destinadas ao marido (W) seja nas cartas à filha (A), além de destoar do comportamento dos demais membros da família e da sociedade de então, suscita algumas questões a serem respondidas a partir dos postulados da SH aliado ao conceito de redes sociais:

i. Por que no círculo familiar pai-mãe-filha só uma remetente, a esposa-mãe (E), ainda emprega *tu* em suas cartas? Seria um uso relacionado a um comportamento de uma pessoa mais velha nascida provavelmente no início do século XX? Tal uso estaria relacionado à faixa etária da remetente? Se a resposta for positiva, por que seu marido, de idade equivalente, emprega majoritariamente *você* e não *tu*?
ii. Que aspectos podem ajudar a compreender esses resultados da remetente (E) que apresentou um comportamento divergente das mulheres adultas da época, difusoras da forma *você,* como mostrou Souza (2012)?
iii. Como a identificação do perfil social dos autores dessas cartas e das redes sociais das quais participavam podem elucidar tais questões?

É fato que não dispomos de informações precisas sobre a posição e as características sociais desses autores desconhecidos, que foram identificados apenas como família FB. Nesse caso, urge recuperar algumas informações textuais do material e reconstruir a história social da família a partir de preceitos peculiares à SH.

Para responder as questões formuladas, retomamos as observações de Bergs (2005) do alcance dos fatores externos (faixa etária, gênero, escolaridade) na interpretação da difusão de uma mudança a partir de fontes pretéritas. No caso em análise, é preciso tentar explicar a diferença de comportamento linguístico da remetente (E) em relação a outros membros do seu círculo familiar, lembrando que os três remetentes participam de diferentes *redes sociais pessoais*, como pretendemos mostrar (cf. Milroy1992).

Em termos etários, embora não se saiba exatamente a idade da remetente (E), é presumível que ela fosse uma mulher adulta (quase idosa) quando começou a se corresponder com sua filha, em 1970. O seu marido (W) era um militar da Marinha brasileira e suas primeiras cartas são da década de 1960. Os pais pertenciam à elite carioca de sua época. Moravam em um apartamento próprio em Copacabana, tinham outro imóvel no bairro da Tijuca e uma casa de férias em São Lourenço. Os dois compartilhavam de algumas características sociolinguísticas tradicionais, como idade, local de residência e classe social, logo a diferença de comportamento linguístico, no que se refere às escolhas tratamentais (ela com o emprego de *tu* e ele com o uso de *você*), não pode ser atribuída simplesmente à faixa etária, pois ambos certamente deviam ter idades equivalentes.

Ainda que estivessem no mesmo escopo etário, os estilos de vida (as suas redes) eram diferentes. O pai era um militar aposentado, bastante letrado. Sua escrita é mais rebuscada e suas escolhas lexicais são precisas quando escreve à esposa: *"Então, se não há o encanto e alegria de sua presença, há o lenitivo, a terna lembrança de que você pensando nós, bem perto de mim está."* Embora tenha sido preso no período da Ditadura, (W) foi um comandante de alta patente nas forças armadas. Era um homem viajado e tinha prestígio social por sua atuação militar. Esses fatores podem favorecer o uso de formas inovadoras (não estigmatizadas), pois a mobilidade espacial acaba por propiciar o contato com membros de outras redes sociais e, consequentemente, com outras normas de uso que influenciam a sua variedade linguística (cf. Tieken-Boon Van Ostade 2006).

Em termos do tratamento utilizado, o pai empregava preferencialmente formas de *você* sem "mesclar" com formas do paradigma de *tu*. Em suas cartas, é possível encontrar, por exemplo, o emprego do clítico *-a-* na função acusativa e o *-lhe-* no dativo, ambos se referindo à segunda pessoa e seguindo o *padrão canônico* de uniformidade tratamental *você-o-lhe-com você-seu*, como se vê em (1):

(1) "Tenho pensado muito em *você* e estou ansioso para voltar e então terei um mundo de carinho para envolvê-*la* ternamente. (...) Espero, tudo esteja correndo normalmente *com você*, *sua* saude e também com nossos filhos". (Carta do marido W para a esposa E, Rio, 4. 21-10-1960)

(2) "*Seu* sempre, q *lhe* beija" (Carta do marido W para a esposa E, Rio, 01-08-1978)

A esposa e mãe (E) tinha um perfil distinto. Era uma dona de casa reservada, cristã devota que vivia em função dos filhos e de sua fé. Suas cartas tratam de questões mais cotidianas do que as cartas de seu marido e trazem aspectos próprios da modalidade falada (maior proximidade comunicativa), nos termos de Koch e Oesterreicher (2007). Seu texto apresenta ideias ou frases enumeradas, pouco planejamento textual, vocabulário simples, presença de operadores discursivos típicos da fala, ausência eventual de sinais de pontuação e palavras com desvios grafofonéticos (*quasi* por *quase*, *faser* por *fazer*, *dis* por *diz,* ilustrados nos exemplos 3 e 4): aspectos esses que evidenciam seu relativo contato com modelos de escrita (cf. Barbosa, 2005):

(3) "*Olha* vou chegar no Rio dia 20 às 5 da manhã, não vou levar *quasi* nada só alguns perfumes, porque as coisas aquí fedem a alho" (Carta da esposa E para o marido W, Paris 14-03-1980)

(4) "Ela é muitíssimo malcriada, autoritária e não aceita conselhos se se der mal ela diz que pagará e com juros e correção monetária. Isso não é nosso, o mundo é passagem vamos *faser* a vontade dela ela confia na amiga e no amante da amiga *dis:*" (Carta da esposa E para o marido W, Paris 14-03-1980)

Há dados adicionais que reiteram as dificuldades de (E), por exemplo, com a língua francesa. Em (5), destacamos a passagem em que a mãe (E) comenta como se sente constrangida por sua "ignorância" diante da neta. A sua devoção a Deus, sua religiosidade, pertencimento a grupo de amigas da igreja, idas à missa, estão registradas também em (6):

(5) "A G. quan*do se fala em Deus* só diz: _ Oui, Oui... e depois nada mais. Ela fala fala francês, mas *eu não entendo nada*, ela até ri da *minha ignorância* e fica falando mais do q*ue* ela sabe para me confundir mesmo. É até engraçado (...)" (Carta da esposa E ao marido W, Paris, 14-03-1980)

(6) "…. mas confio no "Senhor e deposito esta criança em *suas Benditas* mãos sei que *Ele fará o que quiser*, digo aqui na terra. W na vida nada somos, como passageiros, pena não possa estar aí perto de ti, como gostaria. Cada dia que passa sinto mais saudades, desejava ir a praia*, ver as minhas amigas que tanto amo, o pessoal da carismática,* tanta coisa me passa pela..." (Carta da esposa E ao marido W, Paris, 20-01 1980)

Esses e outros fragmentos das cartas de (E) ilustram que seu círculo de amizades e as *redes sociais* das quais participava eram mais restritas e densas se comparadas, por exemplo, ao que se observa com seu marido e, principalmente, com sua filha que era uma mulher cosmopolita. A esposa-mãe (E), nascida provavelmente nas primeiras décadas dos novecentos, teve acesso às primeiras letras, mas não deve ter avançado tanto nos estudos, como era fato entre as mulheres de sua época. Assim, (E) trouxe para sua escrita formas mais convencionalizadas do século anterior ao empregar de maneira bastante regular e conservadora as formas de *tu*. Há dois aspectos ainda que reiteram esse conservadorismo nas suas escolhas tratamentais. O primeiro deles diz respeito à preservação de formas do mesmo paradigma. Nas funções de complemento acusativo, em (7), e dativo (de 8 a 18), a maioria quase absoluta dos dados (mais de 90% dos casos) de (E) é do clítico *te*. Houve apenas uma ocorrência do dativo "*a você*" em (18):

(7) "Não *te* cuques nem te apoquentes" 102 - 13-01-1980
(8) "Desejo [*te*] escrever pouco, para não te dizer o q sinto, porem tenho q escrever, aqui vai a lenga, lenga". (Paris, 25-03-1977)
(9) "A A. ainda esta dura e forte ontem ela falou 1/2 hora com o At. q. [*te*] explicará" (Paris, 01-01-1980)
(10) "W. eu [*te*] peço assine, faça o q. esta besta quer, se perder fica vêr navios". (Paris, 13-01-1980)
(11) "o D. eufórico com a vinda do bebê me fasia ou melhor ligou o telefone para [*te*] participar".(Paris, 27-01-1980)
(12) "hoje estou [*te*] escrevendo são 5 da madrugada e acordei e achei de melhor escrever" (Paris, 27-01-1980)
(13) "Estou [*te*] escrevendo, esta carta, vai por obsequio" (Paris, 04-02-1980)
(14) "Hoje coloquei uma carta no correio para ti e vai essa para [te] esclarecer sobre o empréstimo" (Paris, 04-02-1980)

(15) "se não [*te*] encontrar em casa já sei q. estás fora" (Paris, 04-02-1980)
(16) "Wa., só [*te*] peço q. não consertes nada antes de eu chegar" (Paris, 13-03-1980)
(17) "não estou falando sobre o Wi., mas depois [*te*] digo de quem". (Paris, 13-03-1980)
(18) "Ela disse que é a única coisa que pediu *a voce* e espera que este negócio seja feito." (Paris, 13-03-1980)

O segundo aspecto do conservadorismo de (E) está na confusão que a remetente faz com as formas verbais de segunda pessoa. Em suas cartas, a remetente emprega no pretérito perfeito a desinência de segunda pessoa do plural *-stes* no lugar de *-ste*, o que tanto pode evidenciar um uso arcaizante do tratamento *vós* de deferência, quanto a confusão bastante frequente que se faz entre a desinência número-pessoal e a modo-temporal, principalmente no português europeu[5]:

(19) "Sei que fostes a São Lourenço ***bebestes*** umas aguinhas e ***tivestes*** tranquilidade." (Mulher-marido, Paris, 25-07-1980)

Em termos da distinção de gênero, questão (ii), os usos linguísticos da remetente feminina (E) não estão afinados com o que fora observado em documentos de mulheres que escreveram no mesmo período (cf. Souza, 2012, Lopes *et al*., 2018, entre outros). Nas cartas trocadas entre (E) e sua filha (A), notamos que as duas utilizam estratégias diferentes de tratamento: enquanto (E) sempre trata todos por *tu*, a filha (A) emprega quase categoricamente o *você* nas suas correspondências aos pais, irmão e tios.

Como o "estilo" de vida da filha (A) pode ter influenciado nos usos linguísticos identificados em suas cartas? Diferentemente da mãe e dos irmãos, a filha (A) participava de *redes sociais* bastante amplas e diversificadas. Morou, por um longo tempo, na Cidade Luz, onde levava um estilo de vida extremamente elevado: viajava por toda a Europa, convivia com amigos da alta sociedade, entre eles filhos de embaixadores e outras pessoas de alto nível social. Além de francês, sabia inglês e um pouco de alemão. Lia muito e fazia questão de citar obras e autores nas cartas que escrevia aos pais:

[5] Como o morfe mais geral de 2SG é o -s- e não há desinência específica para o pretérito perfeito, é comum que se associe a marca -ste- ao tempo e não à pessoa. Nesse caso é frequente que se interprete -ste- como marca do pretérito e se acrescente -s- para marcar o singular e não o plural.

(20) "Divido-me também minhas horas de lazer, outrora gastas com praias (que aqui infelizmente não as tenho) e bavardices(sic) tolas com a riqueza do clássico Stendhal onde o conheço atraves de um personagem pleno que personalidade chamado Lucien Leuwen e onde já tive o prazer de ler seu clássico Rouge et noir. Bom, prometi a meu médico durmir cedo e já são 2 hrs. É realmente incrível como, qdo se quer fazer muitas coisas, o tempo urge." (Carta da filha A para os pais W e E, Paris, 02-06-1970)

Além disso, a filha (A) trabalhou em várias atividades diferentes na França e na fase adulta enveredou pela carreira cinematográfica com seu marido (D). Embora estivesse sempre em busca de trabalho na França, não havia grandes preocupações suas com questões financeiras porque ela sabia que poderia contar com uma ajuda de custo dos pais que sempre lhe enviavam dinheiro.

A remetente filha (A) era uma cidadã do mundo e pertencia a redes sociais mais abertas e frouxas. Na esteira de Tieken-Boon van Ostade (2006, p. 251), pode-se conjecturar qua a sua "mobilidade espacial a coloca em contato com outras redes sociais e, portanto, com diferentes normas de fala que podem influenciar sua própria língua e a dos que a cercam". Em perfis como os da filha (A), a pressão normativa seria mais fraca, pois os laços entre os falantes são menos densos: praticamente uma pessoa conhece muitas pessoas, mas em um único papel social, como afirmam os autores. Nesses casos, há favorecimento para o uso de formas mais inovadoras e a filha (A) se enquadra nesse perfil. Com relação a suas escolhas tratamentais, ela é a única do círculo familiar (pai-mãe-filha) que adota o padrão *híbrido* (*você~tu*) vigente no português falado atual do Rio de Janeiro. Os exemplos (21) e (22) mostram a presença do clítico *te* na função acusativa (***te*** *disser*, ***te*** *levará,* ***te*** *procuraremos*) ao lado de *você* na posição de sujeito (***você*** *diga*, *quando* ***você*** *chegar*):

(21) "quando êle te disser *você* diga (não se esqueça) à pessoa que ***te*** levará a estação que deve ser M. K. (Carta de A para a mãe E, Paris, 23-02-1977)

(22) "Quando *você* chegar a Paris desça do trem com as malas e fique parada na plataforma nós ***te*** procuraremos. (Carta de A para a mãe E, Paris, 23-02-1977)

Como se vê nos dados das cartas da filha (A), a forma *você* predominou na posição de sujeito como ocorre nas cartas do seu pai, mas, nas funções acusativas e dativas, a

filha (A) foge do padrão canônico ***você-o-lhe-com você*** utilizado por ele. Suas cartas já apresentam o novo *paradigma híbrido* de 2SG que é vigente no Rio de Janeiro atual: *você -te-com você.*

5 Considerações finais

Os resultados obtidos são complexos no que diz respeito aos usos linguísticos que caracterizam à comunidade carioca da época, em particular, alguns membros da família FB. A frequência de uso de *tu* e *você* parece estar associada aos traços individuais de cada autor, pois o comportamento dos remetentes da amostra não é uniforme, visto que alguns adotam um subsistema tratamental mais inovador (o paradigma híbrido atual) e outros, por outro lado, ainda empregam formas que eram frequentes em fins do século XIX e início do XX.

Em termos objetivos, o padrão canônico *tu-te-ti-contigo,* mais conservador, ficou patente nas cartas da matriarca da família (E) e as formas *você-o-lhe-com você* prevaleceram nas cartas do pai (W). A primeira (a esposa) pertencia a redes sociais mais densas do que o segundo (o marido). Diferentemente dos pais, a filha (A) -- a protagonista mais jovem da amostra, uma "cidadã do mundo" pertencente a redes mais difusas -- adota o *padrão híbrido* e já vigente no português falado atualmente no Rio de Janeiro.

Foi possível comprovar que o comportamento linguístico mais ou menos inovador de alguns indivíduos parece ser consequência do pertencimento a redes sociais mais ou menos abertas e não ao que habitualmente se detecta em um grupo social concreto do período em análise. Os aspectos biográficos dos remetentes-destinatários, detectados a partir da análise de pistas textuais das próprias cartas analisadas, se mostraram relevantes para a análise das estruturas prototípicas da rede (*unidas vs. soltas*; *laços fracos vs. fortes*) de cada membro da família FB, favorecendo, assim, seus usos linguísticos em termos das escolhas tratamentais, conforme defendem Bell (1984); Moreno Fernández (1998); Schilling-Estes (2002)-(*apud* Conde Silvestre, 2007).

Estudos posteriores com novos materiais a serem descobertos e editados podem ajudar a confirmar essa linha de interpretação fundamentada no controle das redes sociais de cada indivíduo.-

Referências

BARBOSA, Afrânio Gonçalves. Tratamento dos corpora de sincronias passadas da língua portuguesa no Brasil: recortes grafológicos e linguísticos. *In:* LOPES, Célia Regina dos Santos (Org.). *A norma brasileira em construção: fatos linguísticos em cartas pessoais do século XIX*. Rio de Janeiro: UFRJ, FAPERJ, 2005. p. 25-43.

BERGS, A. Social networks and historical sociolinguistics: Studies in morphosyntactic variation in the Paston letters (1421–1503). *Language in Society*, v. 37(2), 2005. p. 316-317.

BROWN, R.; GILMAN, A. The pronouns of power and solidarity. *In:* SEBEOK, T. (Ed.). *Style in language*. Cambridge: MIT Press, 1960. p.253-276.

CONDE SILVESTRE, Juan Camilo. *Sociolinguística histórica*. Madrid, Gredos, 2007.

HERNÀNDEZ-CAMPOY, Juan; CONDE SILVESTRE, J. Camilo. *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, 2012.

KOCH, Peter; OESTERREICHER, Wulf. Oralidade y escrituralidad a luz de la Teoria del Lenguage. *In: Lengua Hablada en La Romania: español, francés, italiano.* Madrid; Editorial Gredos, 2007. p. 20-42.

KOCH, Peter. Tradiciones discursivas y cambio linguístico: el ejemplo del tratamiento vuestra merced en español. *In:* Kabatek, Joahnnes. *Sintaxis histórica del español y cambio linguistico:* Nuevas perspectivas desde las Tradidiciones Discursivas. Iberomaricana/ Vervuert, Madrid/Frankfurt, 2008.

LOPES, C.R.S.; MARCOTULIO, L. L.; OLIVEIRA, T. L. Forms of address from the Ibero-Romance perspective: A brief history of Brazilian voceamento. *In:* Célia Lopes; Martin Hummel. *Address in Portuguese and Spanish: Studies in Diachrony and Diachronic Reconstruction.* 1ed. Berlin, Boston: De Gruyter, 2020. p. 111-154.

LOPES, Célia Regina dos Santos *et al.* A reorganização no sistema pronominal de 2a. Pessoa na história do português brasileiro: outras relações gramaticais. *In:* Célia Regina dos Santos Lopes. (Org.). *História do português brasileiro: mudança sintática das classes de palavra: perspectiva funcionalista*. 1ed.São Paulo: Editora Contexto: 2018. v. 4 ,p. 106-189.

LOPES, Célia Regina dos Santos; SOUZA, Janaína Fernandes. Formas de tratamento em cartas do Rio de Janeiro. *In:* Célia Regina dos Santos Lopes. (Org.). *História do português brasileiro: mudança sintática das classes de palavra: perspectiva funcionalista*. 1ed.São Paulo: Editora Contexto, 2018. v. 4, p. 46-67.

MILROY, J. *Linguistic variation and change.* Oxford: Blackwell: s.n., 1992.

MILROY, L., Milroy, J. Social network and social class: Toward an integrated sociolinguistic model. *Language in Society*, 21(1), 1-26.: s.n., 1992.

NEVALAINEN, Terttu. 'What are historical sociolinguistics?' *Journal of Historical Sociolinguistics*, v.1, n. 2, p. 243–269. 2015. https://doi.org/10.1515/jhsl-2015-0014

OLIVEIRA, T. L. *Entre o linguístico e o social: complementos dativos de 2ª pessoa em cartas cariocas (1880-1980)*. 2014. Dissertação (Mestrado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014. https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popup=true&id_trabalho=177523

RUMEU, M. C. B. *A implementação do Você no Português Brasileiro Oitocentista e Novecentista: um estudo de painel*. 2008, 276 f. Tese (Doutorado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

RUMEU, Márcia Cristina de Brito; CALLOU, Dinah Maria Isensee. A implementação do você no português brasileiro: evidências da língua escrita, *Sintaxe diacrônica* ,v. 22, p. 192-211. 2021. https://periodicos.ufsc.br/index.php/workingpapers/issue/view/3180

SOUZA, C. D. *Eu te amo, Eu lhe adoro, Eu quero você: a variação das formas acusativas de 2ª pessoa em cartas pessoais (1880-1980)*. 2014. Dissertação (Mestrado em Língua Portuguesa) – Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2014. https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popup=true&id_trabalho=1310348

SOUZA, Janaina Pedreira Fernandes. *Mapeando a entrada do Você no quadro pronominal: análise de cartas familiares dos séculos XIX-XX*. 2012. Dissertação (Mestrado em Língua Portuguesa). Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2012.

TIEKEN - BOON VAN OSTADE, I. Social network analysis and the history of English. *European Journal of English Studies*, v. 4, n. 3, p. 211-16. 2006. 10.1076/1382-5577(200012)4:3;1-S;FT211

CC BY NC

http://doi.org/10.5007/1984-8420.2025.e101248

A EXPRESSÃO PRONOMINAL DE POSSESSIVOS DE SEGUNDA PESSOA EM CARTAS DE AMOR DO SÉCULO XX

THE PRONOMINAL EXPRESSION OF SECOND-PERSON POSSESSIVES IN 20TH CENTURY ROMANCE LETTERS

Cristiane Namiuti | Lattes
cristianenamiuti@gmail.com
Universidade Estadual do Sudoeste da Bahia

Raiana Cristina Dias da Cruz | Lattes
raianacristinadias@gmail.com
Universidade Estadual do Sudoeste da Bahia

Carlos Wilson de Jesus Pedreira | Lattes
carloswilsonpedreira@yahoo.com.br
Universidade Estadual do Sudoeste da Bahia

Mireia Lêda Silva Santos | Lattes
mireialeda12@gmail.com
Universidade Estadual do Sudoeste da Bahia

Isabel dos Santos Magalhães Gomes | Lattes
dossantosmagalhaesi@gmail.com
Universidade Estadual do Sudoeste da Bahia

Gabriela Santos Barros | Lattes
gabrielabarros852@gmail.com
Universidade Estadual do Sudoeste da Bahia

Daniela Soares de Oliveira | Lattes
danielasoares2806@gmail.com
Universidade Estadual do Sudoeste da Bahia

Resumo: Apresentamos aqui, em perspectiva diacrônica, a descrição e análise do uso das formas possessivas de segunda pessoa em cartas de amor do século XX, escritas por dois casais de missivistas, de diferentes gerações, residentes no Sertão do Pajeú, em Pernambuco. Os resultados apontam para diferentes usos dos pronomes possessivos, registrando-se a preferência das formas *teu/tua* nas cartas de amor dos missivistas que se corresponderam nos anos 1950 (1956 a 1958), coexistindo com o uso do pronome *você,* em variação com o pronome *tu,* na função sujeito e com o pronome *lhe* em variação com *te, expressando o dativo e o acusativo de segunda pessoa,* enquanto que nas cartas do casal da década de 1970 (1972 a 1977) atesta-se preferencialmente o uso das formas possessivas *seu/sua* coexistindo com o pronome *você* na função sujeito e o pronome *lhe* como a única opção do dativo e acusativo de segunda pessoa. O artigo questiona a hipótese, amplamente aceita, de que as mudanças no sistema pronominal e de concordância do PB seriam consequências da gramaticalização de *Vossa Mercê* em pronome de segunda pessoa *você,* e ainda postula que as alterações no uso das formas possessivas para expres-

sar a segunda pessoa nas cartas de amor do Sertão do Pajeú inserem-se em um quadro de outras alterações no sistema pronominal e de concordância, fatos que sugerem mudança relacionada a perda da marcação morfológica de traços morfossintáticos, morfossemânticos e discursivos.

Palavras-chave: Sistema pronominal. Concordância. Variação. Mudança.

Abstract : We present here, from a diachronic perspective, a description and analysis of the use of second person possessive forms in 20th century love letters written by two couples from different generations living in the Sertão do Pajeú region of Pernambuco. The results point to different uses of possessives pronouns, with a preference for the forms 'teu/tua' in the love letters of the writers who corresponded in the 1950s (1956 to 1958), coexisting with the use of the pronoun 'você', in variation with the pronoun 'tu', in the subject function and with the pronoun 'lhe' in variation with 'te', expressing the dative and accusative of the second person, while the couple's letters from the 1970s (1972 to 1977) preferentially show the use of the possessive forms 'seu/sua' coexisting with the pronoun 'você' in the subject function and the pronoun 'lhe' as the option for the second person dative and accusative. The article contests the hypothesis, widely accepted, that the changes in the pronominal and agreement system of BP are a consequence of the grammaticalization of "Vossa Mercê" into the second person pronoun "você", and also postulates that the changes in the use of possessive forms to express the second person in love letters from the Sertão do Pajeú are part of a scenario of other changes in the pronominal and agreement system, facts that suggest change in the morphological marking of morphosyntactic, morphosemantic and discursive features.

Keyword: Pronominal System. Agreement. Variation. Change.

1 Introdução

De acordo com Menon (1995), Silva (1996), Lopes e Machado (2005), Arduin (2005), Arduin e Coelho (2006), Martins e Vargas (2014), e com muitos outros trabalhos, as alterações no sistema pronominal do Português Brasileiro (PB) estão relacionadas com a entrada, no século XVIII, da forma *você*, originária do sistema de tratamento da língua portuguesa, no paradigma de pronomes pessoais de segunda pessoa com concordância de terceira. Esse fato teria alterado o sistema de concordância e afetado outras formas pronominais, incluindo o sistema de possessivos (Silva, 1982, 1996; Cerqueira,

1996, 2018; Perini, 1985; Galves, 1998; Lopes; Machado, 2005; entre outros). O pronome possessivo de terceira pessoa *seu* teria passado a ser utilizado como pronome possessivo de segunda pessoa (ex. 1 e 2), como um dos resultados das modificações ocorridas no paradigma pronominal, decorrentes da inclusão da forma *você* no quadro de pronomes pessoais, tornando o possessivo *seu* altamente ambíguo. Segundo Silva (1996), duas estratégias entram em ação para a dissolução da ambiguidade na língua brasileira: (i) a combinação de tratamento *você/tu*, ou seja, o uso do pronome *você* na função sujeito ou objeto juntamente com o pronome possessivo *teu* (ex. 3); (ii) o uso do genitivo *dele* tomando, cada vez mais, o lugar de *seu* na 3ª pessoa semântica, principalmente na língua oral (Silva, 1996 p. 172), conforme pode ser observado no ex. (5).

(1) Aquele ***seu*** amigo esteve aqui procurando ***você***.
(2) Aquele ***seu*** amigo esteve aqui ***te*** procurando.
(3) Aquele ***teu*** amigo esteve aqui procurando ***você***.
(4) Aquele ***teu*** amigo esteve aqui ***te*** procurando.
(5) Aquele amigo ***dele*** esteve aqui procurando ***você*** (falando do amigo de 'José', uma terceira pessoa)

Marcotulio (2010), Martins e Vargas (2014), Lopes *et al*. (2018) e Lucena (2019) também encontraram esse uso combinado de formas de terceira pessoa com formas de segunda pessoa, referindo-se à segunda pessoa do discurso, no funcionamento do sistema pronominal em dados históricos em perspectiva diacrônica, tendo incidência nos séculos XVIII, XIX e XX, sugerindo que o sistema de concordância que mantém a uniformidade no sistema de tratamento, com o uso de pronomes concordantes, já teria sofrido alguma alteração no século XVIII.

O exemplo (6), extraído da carta do missivista J. R. S, *fac-símile 12* do *corpus* da presente pesquisa e datada de 18 de janeiro de 1958, também ilustra esse uso combinado das formas dos possessivos *teu, seu* e o pronome *você*.

(6) Recebi a ***tua*** cartinha fiquei jubilada em saber notícias ***tuas***. Maria domingo eu fui com Dé até la a casa de seu Artur eu ia até lá em ***sua*** casa; mas desconfiei que ***você*** não estava em casa e voltei. Lamentei quando soube que ***você*** estava mas não tem nada fica para quando nós se avistar então o meu coração irá passar momentos feliz ao ***teu*** lado. (J. R. S fac-símile 12, 18/02/58)[1]

[1] Optamos por uma transcrição conservadora do *fac-símile* do manuscrito original, mantendo a ortografia e segmentação de palavras tal como no documento original. Os exemplos trazidos no artigo seguem essa transcrição.

De fato, o PB coloquial apresenta variação na expressão da posse nominal de segunda e terceira pessoa. As formas pronominais *seu* e *sua* podem expressar posse de segunda ou terceira pessoa em variação com *teu, tua* para segunda pessoa e *dele, dela* para terceira pessoa e este fato remonta fases mais antigas da língua.

Silva (1984) estuda a distribuição de *dele/seu* (de terceira pessoa) e constata que a forma *dele* é a preferida no registro oral, assinalando que essa forma está completamente estabelecida na modalidade oral.

Segundo Menuzzi (1999), o sistema pronominal do PB está passando por mudanças que fazem com que *seu* se torne desfavorecido como forma anafórica para antecedentes referenciais de 3ª pessoa. Este seria o traço inovativo do PB, diferenciando-o de outras línguas românicas.

Partindo de uma perspectiva histórico-diacrônica, baseada em pressupostos gerativistas para modelar a variação e explicar a mudança gramatical (Kroch, 1989; 2001), no presente artigo, apresentamos o resultado da descrição e análise dos usos de *seu* e *teu* em dois conjuntos de cartas de amor de dois casais de missivistas que residiram no sítio Brejinho – comunidade rural do município de Triunfo-PE, que está localizado no Sertão do Pajeú.[2]

O primeiro conjunto de cartas analisado é de um casal que se correspondeu na década de 1950 e o segundo de um casal que se correspondeu na década de 1970. De maneira semelhante ao que se observa em outros trabalhos sobre o mesmo tema (Martins; Vargas, 2014; Lucena, 2019), verificamos a variação entre *seu* e *teu* em termos quantitativos e sua relação com conteúdo discursivo-pragmáticos e, ainda, sua distribuição em relação às demais formas pronominais não possessivas (os pronomes sujeito e objeto).

Observamos diferenças interessantes entre os dois recortes temporais apresentados e, aqui, destacamos duas. A primeira é quantitativa, em que há predominância de uso da forma *teu* para segunda pessoa possessiva nas cartas da década de 1950 (73%) e de *seu* nas cartas da década de 1970. Esta diferença é interessante, pois trabalhos que ana-

[2] O Sertão do Pajeú em Pernambuco possui uma área de 8.689,7 km² e é formada por 17 municípios. A ocupação dessa região pelo colonizador europeu teve início ainda no século XVI e teve como ponto de partida a cidade de Salvador-BA que através de expedições de reconhecimento das terras e aprisionamento dos povos indígenas, promovidos pela Casa da Torre, Bahia, sob a administração de Garcia D'Ávila, iniciou-se o processo de povoação desta região, marcada pela exploração colonial e posterior incentivo a criação de gado e cultivo de alimentos que deram origem a pequenos núcleos populacionais que se desenvolveram em cidades e vilas (Abreu, 1930). De acordo com o censo demográfico 2010 do IBGE, vive uma população de 314.603 habitantes, sendo 199.726 habitantes na área urbana e 114.877 habitantes na zona rural. Na agricultura, além do milho e feijão, a região cultiva a cana-de-açúcar utilizada por cerca de 100 engenhos que produzem mel, rapadura e cachaça. Abrangência: Serra Talhada, Sertânia, Afogados da Ingazeira, São José do Egito, Custódia, Flores, Santa Terezinha, Carnaíba, Verdejante.

lisaram a variação, nesse mesmo espaço temporal, mesmo tipo de documento e mesma relação de missivista, embora com um recorte geográfico diferente, não atestaram essa diferença quantitativa nos usos de *teu* e *seu*. Lucena (2019), por exemplo, ao analisar cartas pessoais de missivistas com o mesmo tipo de relação (relação simétrica Noivo/Noiva, Namorado/Namorada), para o período equivalente ao nosso recorte temporal (1940-1970), registrou 91% de uso da forma *seu* como possessivo de segunda pessoa, sendo o pronome *teu* bastante marginal no *corpus* desse período. A segunda é que há um fator discursivo/pragmático orientando o uso dos possessivos *seu* e *teu* que está presente somente nas cartas da década de 1950. Os dados extraídos das cartas da década de 1970 sugerem que este fator não está mais presente, como veremos na seção 3 desse artigo.

Uma vez que o uso de *seu/sua*, como pronomes possessivos de segunda pessoa, insere-se em um quadro de outras alterações no sistema pronominal, como a alteração do sistema de clíticos, o uso do clítico *lhe*, para expressar a segunda pessoa dativa e também a segunda pessoa acusativa, a gramaticalização de *você* como pronome de segunda pessoa, além de outras alterações na gramática como a alteração no sistema de concordância, fatos descritos em muitos trabalhos a exemplos dos citados neste artigo, questionamos como se dá a variação no uso das formas pronominais de segunda pessoa possessiva nas cartas de amor do Sertão do Pajeú e, ainda, qual seria o estatuto gramatical dela, ou seja, o uso variável das formas encontradas nas cartas de 1950 é produzido pelo mesmo sistema que produziu o uso variável nas cartas de 1970?

O presente artigo está estruturado da seguinte forma: nesta primeira seção, foi apresentado o problema juntamente com uma síntese do estado da arte. Na seção 2, apresentam-se o *corpus* e a metodologia. A seção 3 está dedicada à apresentação dos resultados e, por fim, na seção 4, apresentam-se as considerações finais.

2 Corpus e metodologia

Partindo de uma perspectiva histórico-diacrônica, baseada em pressupostos gerativistas para modelar a variação e explicar a mudança gramatical (Kroch, 1989; 2001), buscou-se descrever e analisar a variação das formas de expressão pronominal da posse de segunda pessoa *teu/tua vs seu/sua*, em cartas pessoais produzidas por dois casais de missivistas de diferentes gerações que trocaram cartas na segunda metade do século XX (1956-1977). Os dados, aqui analisados, foram extraídos de 122 cartas pessoais, do subgênero carta de amor, pertencentes a dois casais que se corresponderam em sua juventude. O primeiro casal trocou cartas na década de 1950 e o segundo casal, na década de 1970.

Tais cartas foram doadas ao Laboratório de Documentação Linguística de Pernambuco (LEDOC) e hoje fazem parte do banco informatizado de textos do LEDOC que as cedeu ao Laboratório de Pesquisa em Linguística de *Corpus* (LAPELINC) para processamento, seguindo a metodologia de construção de corpora do LAPELINC (Santos; Namiuti, 2019).

Os *fac-símiles* das cartas foram numerados e o texto de cada carta transcrito em meio digital, em documento de texto identificado com o número do *fac-símile*. Os dados foram extraídos manualmente e organizados em planilha para descrição, tabulação, quantificação e análise. Foram identificados os pronomes possessivos, pessoais e oblíquos, com valor de segunda pessoa, utilizados nas cartas e observado o funcionamento do sistema pronominal em cada carta para cada missivista. Os dados são identificados com o número do *fac-símile*, as iniciais do nome do missivista e a data.

3 O funcionamento e a distribuição das formas pronominais possessivas, encaixadas no sistema pronominal, em cartas de amor das décadas de 1950 e 1970

As cartas analisadas apresentam um retrato da variação no emprego dos possessivos de 2ª pessoa, compondo as marcas linguísticas de missivistas que representam duas gerações subsequentes. O primeiro casal, composto de homem e mulher nascidos ainda na primeira metade do século XX, escreve na década de 1950, e o segundo casal, composto de homem e mulher nascidos já na segunda metade do mesmo século, escreve na década de 1970. O resultado da distribuição quantitativa do uso de *teu(s)/tua(s)* e *seu(s)/sua(s)* aponta para uma diferença significativa entre os dois conjuntos, conforme se vê na Tabela 1.

Tabela 1 – Pronomes com valor de segunda pessoa possessiva utilizados nas décadas de 50 e 70

Formas Pronominais	**Década de 50**		**Década de 70**	
	22 cartas		**99 cartas**	
teu(s)/tua(s)	45	**73%**	71	33%
seu(s)/sua(s)	17	27%	143	**67%**
Total	**62**		**214**	

Fonte: elaborada pelos autores.

O resultado quantitativo da análise das cartas da década de 1950 se aproxima mais do resultado de Lucena (2019) para os documentos do início do século (1900-1939) (com

95% de uso de *teu* na relação Namorado-Namorada) que para os documentos contemporâneos a elas (1940-1979) (com 9% de uso de *teu* na relação Namorado-Namorada). Já o resultado quantitativo das cartas da década de 1970 ainda apresenta um uso de *teu* em torno de 33% enquanto esse uso já aparece quase marginal nos documentos analisados por Lucena (2019) e que compreendem também esse período de 1970.

Todavia, um aspecto discursivo-pragmático foi observado no uso das formas *teu* e *seu*, nas cartas da década de 1950, que não está exata ou diretamente relacionado a graus de parentesco, como demonstraram Lucena (1999) e outros autores, ou a graus de intimidade, como demonstraram Martins e Vargas (2014) e outros trabalhos, mas mais a um funcionamento discursivo de aproximação e distanciamento entre os missivistas, que pode sim estar relacionado ao rompimento momentâneo da relação simétrica envolvendo o mesmo casal em uma forma de observação mais qualitativa do texto.

De acordo com Andrade e Namiuti (2019), o PB teria uma situação de diglossia entre um sistema semelhante ao uso do Português Clássico (PC), com preferência pelo uso de *seu/sua* para a terceira pessoa, e um sistema reinterpretado (novo) em que *seu/sua* têm valor variável em oposição à anáfora de 3ª pessoa *dele* e o pronome possessivo de 2ª pessoa *teu*. Segundo os autores, a diglossia impede que se tenha uma visão clara do sistema, havendo uma preferência pelas formas *seu/sua* quanto maior for o nível de formalidade.

O funcionamento pragmático relacionado à aproximação e distanciamento entre 'falante' e 'ouvinte', como motivador dos tipos de uso das formas pronominais de segunda pessoa, pôde ser observado nas cartas de amor da década de 1950. Os missivistas apaixonados se tratavam por *tu*, *te*, *ti*, *teu/tua*, como pode ser visto no exemplo (7), contendo um trecho destacado de uma carta em um momento de reaproximação dos enamorados, datada de 18 de novembro de 1956. Mas, em alguns momentos, em que há um distanciamento entre eles, provocado pela ameaça de nunca mais poderem se aproximar e viver seu amor, o tratamento se dá pela forma *você*, associada ao possessivo *seu/sua* e ao pronome obliquo *lhe*, como no exemplo (8) que contém um trecho destacado da carta datada de 20 de janeiro de 1956.

(7) [...] julgava esquecido por ***ti*** mais em horas quando uma feliz noticia me desperta trazendo grande alegria e satisfação pois já pensava em esta esquecido por ***ti*** amando sem ser amado no momento em que de ***ti*** tenho uma noticia, meu coração cada dia ***te*** ama mais amas***te*** sem fingimento, meu nasceu somente para ***ti***, és de toda minha consideração, não ***ti*** esqueço um só momen-

to, fico horas a pensar e meditar em que braços andara aquela querida que consagrei todo meu amor puro e declarado pois é claro que ***tua*** simpatia me domina feliz me considero somente em ***ti*** amar apesar das minhas fraquesas [...] (fac-símile 3 RJS, 20/01/1956)

(8) Levo ao ***seu*** conhecimento que recebi a ***sua*** carta que veio me traser uma grande surpresa porque ***você*** sendo noiva eu pensava que ***você*** tinha se esquecido de mim eu lhe digo que não esquecerei ***você*** nunca pois o amor é o mesmo. Maria quanto ao assunto do ***seu*** casamento eu tenho a diser-lhe o seguinte: ***você*** está para se casar com um rapaz que tem recurso e é do gosto de ***seus*** pais e eu sou pobre e nada tenho [...] (fac-símile 3, RJS, 20/01/1956)

Já nas cartas da década de 1970, o uso do pronome possessivo *seu* é quase generalizado e independe das situações pragmáticas, estando sempre associado ao pronome *você*. O uso de *tu* é bastante marginal e o possessivo *teu/tua* ainda é atestado, mas, com uma frequência muito menor. A Tabela 2 apresenta a distribuição dos pronomes com valor de segunda pessoa utilizados em cada uma das gerações e demonstra esses fatos.

Tabela 2 – Pronomes com valor de segunda pessoa utilizados nas décadas de 1950 e 1970

Formas Pronominais	**Década de 50**		**Década de 70**	
	22 cartas		**99 cartas**	
tu	36	**14%**	5	1%
teu(s)/tua(s)	45	**18%**	71	8%
seu(s)/sua(s)	17	7%	143	**17%**
te	73	**29%**	112	13%
ti	29	**12%**	18	2%
você	33	13%	415	**49%**
lhe	17	7%	75	**9%**
Total	**250**		**839**	

Fonte: elaborada pelos autores.

A Tabela 2 demonstra um aumento no uso das formas *seu/sua* como pronomes de segunda pessoa (passando de 7% dos dados na década de 1950 para 17% dos dados na década de 1970), associado a um aumento do uso do pronome *você* (passando de 13% dos dados na década de 1950 para 49% na década de 1970) e do pronome *lhe* representando a segunda pessoa dativa ou acusativa (passando de 7% dos dados na década de 1950 para 9% na década de 1970).

A relação pragmática de afastamento/aproximação como motivação não é observada nas cartas da década de 1970. O uso marginal de *tu* foi atestado em um contexto de reclamação/afastamento entre o casal, como ilustrado no exemplo (9), enquanto o uso de *você* foi encontrado em momentos de aproximação do casal, como ilustra o exemplo (10).

(9) [...] pois eu queria muito que ***tu*** continuasse como eras antes [...] (fac-símile 51, C, 02/06/1975)

(10) "... palavras tão lindas que ***você*** fez com o aniversario..." [...] (fac-símile 01, JGSR, "s.d."

O paradigma flexional e o sistema de concordância se associam às alterações do sistema de pronomes da língua em alguma medida (Galves, 1993; entre outros). A concordância de *tu* com uma morfologia de 3ª pessoa, presente no dado em (9), nos traz mais um elemento para analisarmos a mudança gramatical associada às alterações do sistema pronominal, incluindo os pronomes possessivos de 2ª e 3ª pessoas.

Esses fatos podem sugerir uma acomodação das formas presentes nas cartas da década de 1970, no Sertão do Pajeú, em um novo sistema também revelado pela oscilação da concordância de segunda pessoa associada ao uso do pronome *tu* presente no exemplo (9), cujo sistema já é bastante descrito para o PB. Os dados das cartas de 1950 não estão totalmente acomodados nesse sistema, não há uso de *tu* sem concordância e seu uso não é marginal; é, inclusive, superior ao uso de *você*, além de a preferência pelo uso de *teu* e pelo uso de *seu*, como segunda pessoa, estar condicionada ao afastamento/mudança na simetria da relação, o que não se observa.

Os resultados apontam, claramente, para um aumento do uso das formas *seu/sua* como pronome possessivo de segunda pessoa, na década de 70, associado a uma mudança na especificação pragmática da 3ª pessoa morfossintática, com referência à 2ª pessoa do discurso, ou seja, o uso da 3ª pessoa morfossintática com referência à 2ª pessoa do discurso passou da esfera da formalidade para a esfera da intimidade (ex. 10).

Nas cartas de 1950, é importante apontar que não apenas o pronome possessivo muda de forma em uma situação de distanciamento entre os missivistas, mas também os pronomes sujeito e objeto. Os dados em (8) ilustram bem esse uso na esfera da formalidade dos pronomes da 3ª pessoa morfossintática, com referência à 2ª pessoa do discurso, em que RJS escreve: "recebi a **sua** carta" e "não esquecerei **você**" ao invés de "recebi a **tua**

carta" e "não **te** esquecerei", formas comumente utilizadas por esse missivista que usa exclusivamente as formas *teu/tua* e *te* nas cartas de amor, com exceção das cartas escritas após um momentâneo distanciamento afetivo, momento de rompimento do casal, como é o caso da carta em que foram atestados os dados em (8).

Nas cartas de 1970, de maneira oposta, são os pronomes *seu/sua* os mais comumente atestados e em qualquer situação, apontando para uma generalização dessas formas de 3ª pessoa morfossintática para representar a 2ª pessoa do discurso. Esses dados também apresentaram concordância variável da pessoa verbal, como ilustrado em (9): "tu continua**sse** como era**s**", com a presença tanto do morfema zero da 3ª pessoa verbal, como o morfema *s* da segunda pessoa verbal, foram atestados em concordância com o pronome sujeito *tu* (2ª pessoa morfossintática).

A associação entre a alteração do sistema de concordância e a generalização do uso dos pronomes possessivos *seu/sua* para instanciar a 2ª pessoa do discurso pode ser feita, mas muito há ainda para se investigar, uma vez que os gatilhos para as alterações podem estar em zonas de interface com conteúdos informacionais/discursivos como a topicidade, conforme argumentaram Andrade e Namiuti (2019) para a variação *seu/sua vs dele/dela* para a 3ª pessoa do discurso no PC. Neste trabalho, verificamos que o uso de *seu/sua* deixou de ser sensível a questões pragmáticas como formalidade nas cartas de 1970 e que este fato vem junto com uma alteração no sistema de concordância e outras alterações no sistema pronominal, como já observam em outros trabalhos.

4 Considerações finais

O contraste presente nos dados extraídos das cartas de amor de casais que representam duas gerações adjacentes no tempo e pertencentes a um mesmo espaço geográfico, o Sertão do Pajeú em Pernambuco, sugere uma mudança que generaliza e especifica as formas pronominais possessivas *seu/sua* (3ª pessoa morfossintática) para o referente de 2ª pessoa discursiva. Postulamos, seguindo outros autores, que essa alteração tem origem em uma mudança gramatical profunda que afeta várias alterações superficiais observadas em variação nos dados (Kroch, 1989; 1994; 2001) e que tal mudança está localizada no sistema de concordância da língua.

Muito há ainda para se investigar e discutir, principalmente em relação **à**s associações e relações entre as alterações encontradas nos textos e associadas à mudança gramatical do PB e suas possíveis causas. Começamos este artigo listando alguns trabalhos publicados entre finais do século XX, primeira e segunda década do século XXI, e, em todos, a gramaticalização de *você* aparece ou de forma explicita e direta, ou sugerida, como a cau-

sa das demais alterações. Esta parece ser uma verdade absoluta e inquestionável. Todavia, qual alteração veio primeiro e se é causa ou consequência de uma mudança profunda é uma questão aberta, e, talvez, de 'ovo e galinha'. No caso especifico da hipótese da entrada da forma *você* como pronome de segunda pessoa no paradigma pronominal do Português como gatilho/causa/origem da mudança gramatical, uma questão simples (e até ingênua) que se coloca para uma explicação simples é porque as alterações observadas, no Brasil, não foram observadas em Portugal, uma vez que a gramaticalização de *você* se deu também no Português Europeu (PE). O uso de formas morfossintáticas de 3ª pessoa para se referir **à 2ª pessoa do discurso é antigo está presente no** PE e não foi suficiente para desencadear mudança no sistema de concordância da língua. Por outro lado, a mudança gramatical pode ser explicada por alterações em zonas gramaticais de interface associadas a conteúdos informacionais/discursivos, como a topicidade, que pode explicar a variação *seu/sua vs dele/dela* no PC (Andrade; Namiuti, 2019) e, ainda, a concordância com um tópico discursivo no PB ("*O carro* furou o pneu") (Galves, 2001). Mas estas são questões em aberto para serem desenvolvidas em trabalhos futuros.

Desse modo, concluímos o artigo com uma perspectiva futura de aprofundamento da investigação das alterações das formas possessivas em variedades do PB, sendo esta contribuição inicial o direcionamento dos caminhos de investigação que buscamos trilhar.

Referências

ABREU, J. C. *Caminhos antigos e povoamento do Brasil*. Rio de Janeiro: Sociedade Capistrano de Abreu, 1930.

ANDRADE, Aroldo L. de; NAMIUTI, Cristiane. *A variação nas marcas de posse nominal de 3ª. pessoa no português antigo e clássico*. Apresentação no Congresso Internacional da ABRALIN 50 anos. 2019.

ARDUIN, J. *A variação dos pronomes possessivos de segunda pessoa do singular teu/seu na região sul do Brasil*. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, UFSC, 2005.

ARDUIN, J.; COELHO, I. L. A variação dos possessivos teu e seu e suas implicações estilísticas. *In:* VANDRESEN, P. (Org.). *Variação, mudança e contato linguístico*. Pelotas: Editora da UCEPel, 2006, p. 185-203.

CERQUEIRA, Vicente C. A forma genitiva "dele" e a categoria de concordância (AGR) no português brasileiro. *In:* ROBERTS, I.; KATO, M. (Orgs.) *Português brasileiro*: uma viagem diacrônica. Campinas: Editora Contexto, 2018, p. 105 – 128.

CERQUEIRA, Vicente C. A sintaxe do possessivo no português brasileiro. 1996. Tese (Doutorado em Letras) - Universidade Estadual de Campinas, Campinas-SP.

CERQUEIRA, Vicente C. A forma 'dele' na indicação de posse e as mudanças no sistema pronominal do português brasileiro. Manuscrito. 1990.

GALVES, Charlotte. *Ensaios sobre as gramáticas do português*. Campinas: Editora da UNICAMP, 2001.

GALVES, Charlotte. O enfraquecimento da concordância no português brasileiro. *In:* KATO, Mary A.; ROBERTS, Ian (Org.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora Contexto, 2018. p.301 – 316.

GALVES, Charlotte. Tópicos, sujeitos, pronomes e concordância no português brasileiro. Cadernos de Estudos Linguísticos, Campinas, v.34, p.19-31, 1998.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Censo Demográfico 2010: Brasil. Rio de Janeiro: IBGE, 2012.

KATO, M. A. *A complementaridade dos possessivos e das construções genitivas no português coloquial: réplica a Perini.* D.E.L.T.A., São Paulo, v. 1, n. 1-2, p. 107-120, 1985.

KROCH, Anthony. Morphosyntactic variation. *In:* BEALS, K. et al. (Ed.). Papers from the 30th Regional Meeting of the Chicago Linguistics Society. v.2: *The Parasession on variation and linguistic theory*, Chicago, v. 2, The University of Chicago Press, 1994.

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

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

LOPES, C. R. dos S.; MACHADO, Ana C. Morito. Tradição e inovação: indícios do sincretismo entre a segunda e a terceira pessoas nas cartas dos avós. *In: A norma brasileira em construção, fatos linguísticos em cartas pessoais do século XIX*. FAPERJ/UFRJ. Rio de Janeiro, 2005.

LOPES, Célia R. dos Santos; MARCOTULIO, Leonardo; OLIVEIRA, Rachel; OLIVEIRA, Thiago; SOUZA, Camila. A reorganização do sistema pronominal de 2ª pessoa na história do português brasileiro: outras relações gramaticais. *In: História do português brasileiro:* Mudança sintática das classes de palavra: perspectiva funcionalista. São Paulo: Contexto, 2018. p. 142-185.

LUCENA, Rachel de O. Pereira. O estudo da variação 'teu'/ 'seu': atuação do fator grau de parentesco. *In:* LaborHistórico, Rio de Janeiro 5 (Especial), p.83-103, jan./jun. 2019.

MARCOTULIO, L. L. *Língua e história*: o 2º marquês do Lavradio e as estratégias linguísticas da escrita no Brasil Colonial. Rio de Janeiro: Ítaca, 2010.

MARTINS, M. A.; VARGAS, M. R. M.; Os possessivos de segunda pessoas do singular em cartas de leitores de jornais brasileiros dos séculos XIX e XX. *Polifonia*, Cuiabá, MT, v. 21, n. 29, p. 369-395, 2014.

MENON, O. P. S. O sistema pronominal do português do Brasil. *Letras*, Curitiba, n. 44, p. 91-106, 1995.

MENUZZI, S. 3rd Person Possessives in Brazilian Portuguese: On the Syntax-Discourse Relation. *In:* Botley S, editor. Proceedings of the Discourse Anaphora and Anaphora Resolution Colloquium, Working Papers collection. Inglaterrra: Lancaster University, Centre for Computing and Corpus Research on Language, UCREL Technical Papers, 1999. v.8, p. 191-210.

PERINI, Mário. *O surgimento do sistema possessivo do Português coloquial: uma* interpretação funcional. Revista DELTA, São Paulo, v. I, n. 1 e 2, p. 1-16, ago., 1985.

SANTOS, J. V.; NAMIUTI, C. O futuro das humanidades digitais é o passado. *In:* CARRILHO, E. et al. Estudos Linguísticos e Filológicos oferecidos a Ivo Castro. Lisboa: Centro de Linguística da Universidade de Lisboa, 2019. p. 1381-1404.

SILVA, *G. M. de O. Estudos da regularidade na variação dos possessivos no português do Rio de Janeiro.* Tese (Doutorado em Letras) – Faculdade de Letras, Universidade do Rio de Janeiro, UFRJ, 1982.

SILVA, G. M. de O. *A variação no Sistema Possessivo de Terceira Pessoa. Tempo Brasileiro. n. 78 e 79, 1984* Tese (Doutorado em Letras) – Faculdade de Letras, Universidade do Rio de Janeiro, UFRJ, 1982.

SILVA, G. M. de O. Estertores da forma *seu* na língua oral. *In:* SILVA, Gisele M. Oliveira; SCHERRE, Maria Marta (Orgs.). *Padrões sociolinguísticos*: análises de fenômenos variáveis do português falado no Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996, p.171-180.

http://doi.org/10.5007/1984-8420.2025.e100819

SOBRE A VARIAÇÃO DOS POSSESSIVOS *TEU/SEU* NOS SÉCULOS XIX E XX: ANÁLISE EM CARTAS PESSOAIS

ON VARIATION OF POSSESSIVES TEU/SEU IN 19TH AND 20TH CENTURIES: PERSONAL LETTERS ANALYSIS

Dinah Callou | Lattes | dcallou@gmail.com
Universidade Federal do Rio de Janeiro

Márcia Cristina de Brito Rumeu | Lattes | marcia.rumeu@gmail.com
Universidade Federal de Minas Gerais

Resumo: Neste estudo, apresentamos um panorama da alternância *teu/seu*, com foco no uso do *seu* de 2SG e seus condicionamentos. Embasamos a análise em amostras de cartas pessoais de dois representantes brasileiros (um carioca e um mineiro), de escrita culta, entre fins do século XIX e o início do século XX, considerando as edições propostas por Ferreira Filho (2023) e Luz (2015). Conduzimos o estudo pelos princípios da Sociolinguística Histórica, em relação aos parâmetros de seleção de amostras históricas, e da teoria da mudança – sociolinguística variacionista. Em termos gerais, a produtividade de *teu* é maior, sendo os dados de *seu* motivados pelo contexto linguístico de sujeito de 2ª pessoa (formas tratamentais) e pelo contexto extralinguístico dos subgêneros das cartas pessoais (cartas de amizade e familiares). Ainda que tenhamos em cena um recorte muito específico da escrita culta oitocentista e novecentista, já é possível entrever não só a covariação *teu/seu*, mas também evidências do *seu* de 2SG em sincronias passadas do PB, como discutido por Lucena (2016) para as cartas cariocas à luz da análise de Oliveira e Silva (1982).

Palavras-chave: Possessivos de 2SG. Recorte temporal. Cartas pessoais. Variação e mudança.

Abstract: In this study, we present an overview of the variation of possessive pronouns teu/seu (*your*), focusing on the use of 2nd person *seu* and its conditioning factors. The analysis is based on samples of personal letters from two Brazilian representatives (one

from Rio and one from Minas Gerais), of educated (standard) writing, between the end of the 19th century and the beginning of the 20th century, considering the editions proposed by Ferreira Filho (2023) and Luz (2015). We conducted the study according to the principles of Historical Sociolinguistics, in relation to the parameters for historical samples selection, and the theory of language change – variationist sociolinguistics. In general terms, the productivity of *teu* is greater, and the use of *seu* in our data being motivated by the linguistic context of the 2nd person subject (address forms) and by the extralinguistic context of subgenres of personal letters (friendship and family letters). Even though we have on the scene a very specific time frame of 19th and 20th century educated writing, it is now possible to glimpse not only the covariation of *teu/seu* but also evidences of second person *seu* in past synchronies of BP, as discussed by Lucena (2016) for the carioca letters, following Oliveira e Silva (1982).

Keywords: 2nd-person possessives. Time frame. Personal letters. Variation and change.

Considerações iniciais

Para traçar o percurso histórico de aspectos linguísticos específicos, é preciso analisar todos os dados a que tivemos acesso, mesmo que de forma indireta, a fim de poder identificar a variação e mudança ocorridas numa língua determinada, em séculos anteriores, através de textos representativos de cada época, e, também, de manuais de gramática, que procuram retratar a norma-padrão subjetiva.

É fato que houve sempre duas tendências na elaboração de gramáticas, (i) uma reflexiva/descritiva, como, por exemplo, em relação à língua portuguesa, a de Fernão de Oliveira, de 1536, a de Argote, de 1725, e (ii) outra, de caráter mais pedagógico, como a de João de Barros (1540) e a de Lobato (1770), que deixam transparecer a experiência prévia no ensino de latim. Em busca de elementos que contribuíssem para esse conhecimento prévio, nos deparamos com uma gramática de João de Barros, de 1538, uma gramática do latim (*Grammatices rudimenta*), que apresenta, pela primeira vez – pelo menos, até onde vai o nosso conhecimento – uma representação arbórea da constituição de uma gramática, exposta na iluminura, a seguir, cf. imagem 1.

Imagem 1: Página de rosto da *Grammatices rudimenta* de João de Barros (1538)



Fonte: https://fontesdoportugues.bnportugal.gov.pt/index.php/gramatica-norma-e-ensino?view=tag&id=20:gramatica-norma-e-ensino

Ao inserir a iluminura nosso objetivo foi situar o foco de análise. A organização em forma de uma árvore, construída a partir de categorias morfológicas – *castas,* como os gramáticos antigos denominavam – e não da organização sintática, como a representação gerativa atual, tem por finalidade registrar as várias categorias, classes de palavras, por vários níveis, ('galhos'), a partir de *Litera* (Letra), *Syllaba* (Sílaba), *Dictio* (Dicção), *Coniūctio* (Conjunção), *Interiectio* (Interjeição), *Pirpositio* (Preposição) até chegar ao nível mais alto, da esquerda para a direita, *Adverbiū* (Advérbio), *Verbum* (Verbo), *ininteligível,*[1] *Nomen* (nome), *Pronomen* (Pronome), disposto pouco abaixo. Essa organização evidencia a posição que ocupa o pronome e confirma a intuição daqueles que nos precederam. Sem dúvida, gramáticos e linguistas de hoje não partiram do zero, ao apresentar

[1] Conjectura: é possível que se tenha feito menção à "frase" ou ao "enunciado".

suas hipóteses teóricas e configurações.

Essas observações prévias localizam, pois, dentro dessa concepção de gramática, nosso objeto de estudo. Neste texto, vamos nos debruçar sobre um tipo de pronome – possessivos, mais especificamente, a variação *teu/seu* de referência à segunda pessoa do singular, em cartas escritas no final do século XIX e início do XX, num período de 41 anos (1869 a 1910). Os autores são figuras de renome, considerados cultos, ainda que um deles não tenha chegado à Universidade, um, nascido no Rio de Janeiro e, o outro, em Minas Gerais. Nosso objetivo é mostrar que, após outras modificações no quadro pronominal – a inserção do *você* – o pronome possessivo *seu,* que formalmente se referia à terceira pessoa do singular e terceira do plural, passaria a ser utilizado também em contextos de segunda pessoa (doravante 2SG), ampliando seu domínio.

Para o processo de mudança alcançado pela forma *seu* que, atualmente, também está voltada à 2SG, assumimos o estudo de Oliveira e Silva (1982) que, ao analisar amostras dos séculos XV, XVI e XVII, conjectura que o *seu* tenha sido favorecido pelo possuidor [+humano] e que, a partir do século XVIII, esse possessivo tenha passado a se integrar a núcleos nominais [- humano]. A ideia é a de que o século XVIII tenha sido o momento (inicial) a partir do qual o *seu* teve o seu escopo semântico ampliado, uma vez que passou a fazer referência a núcleos nominais [-humano] na referência à 2SG, em alternância, atualmente, com o *teu*.

Após considerações prévias, este texto está organizado em quatro partes, a primeira, em que detalhamos nossa fundamentação teórico-metodológica, a segunda, em que apresentamos uma breve revisão do tema e os parâmetros de análise, a terceira, em que descrevemos as amostras históricas, seguida da quarta parte, os resultados da análise, à que se seguem as observações finais.

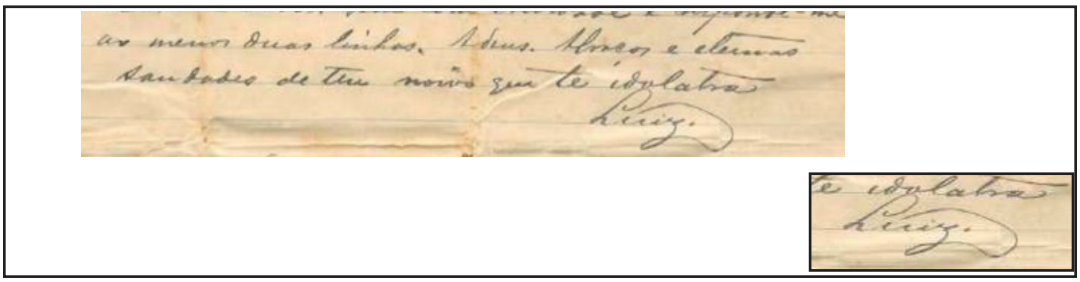
1. Pressupostos teórico-metodológicos: parâmetros da sociolinguística histórica.

As análises embasadas em *corpora* escritos de sincronias passadas estão apoiadas na organização de amostras fidedignas às realidades linguísticas pretéritas (Lima, Marcotulio e Rumeu, 2019). Ao prezarmos pelo conservadorismo das edições com as quais trabalhamos, conduzimo-nos por parâmetros tais como a *autoria*, a *autenticidade* e a *validade social e histórica* das amostras utilizadas (Hernández-Campoy e Schilling, 2012).

A *autoria* dos textos que compõem as amostras históricas é imprescindível à coerência e à veracidade dos resultados dos estudos linguísticos. É necessário trazer à tona o fato de os textos se encontrarem nos acervos públicos e privados, podendo ser autógrafos

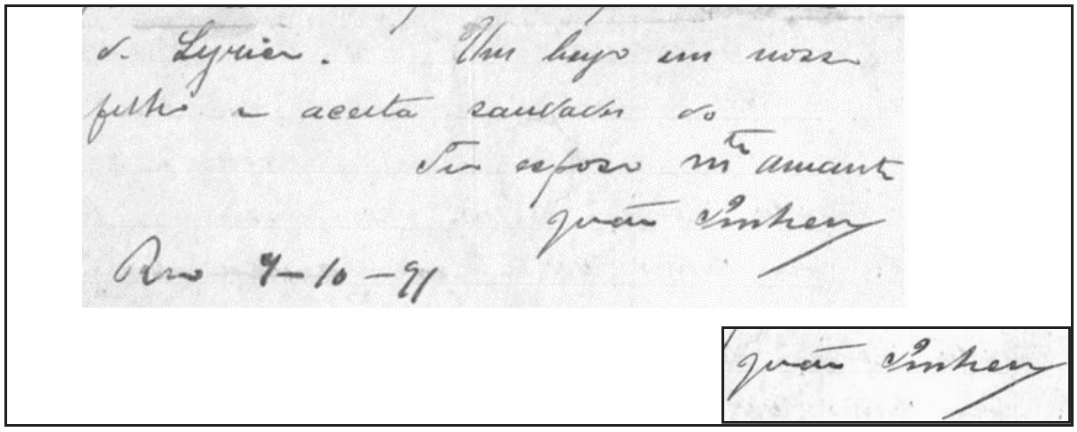
(o redator redige o texto e o assina), apógrafos (o redator redige o texto, mas outro punho o assina) ou ideógrafos (o redator redige o texto, mas há um outro autor (intelectual) que se responsabiliza pela semântica do texto em si. Correspondem a testemunhos históricos distintos, que podem ser viabilizados também através do gênero textual *carta*. A partir da análise do *ductus* em si, uma categoria de análise paleográfica relevante à análise da constituição da escrita, nos termos de Núñez Contreras (1994), a *autoria* dos textos históricos constitui um parâmetro da sociolinguística histórica que depende também da caracterização da materialidade das letras. Neste estudo, são utilizadas cartas pessoais (amor, familiar e amizade) comprovadamente autógrafas, editadas de forma conservadora, e redigidas por dois escreventes que figuram como representantes nascidos em espaços brasileiros de Minas Gerais (João Pinheiro) e do Rio de Janeiro (Gonzaga Duque), conforme ilustrado, através das imagens 2 e 3, a partir de excertos das cartas autógrafas, devidamente assinadas por seus redatores.

Imagem 2: GD. Botafogo, 10.11.1884



Fonte: Ferreira Filho (2023).

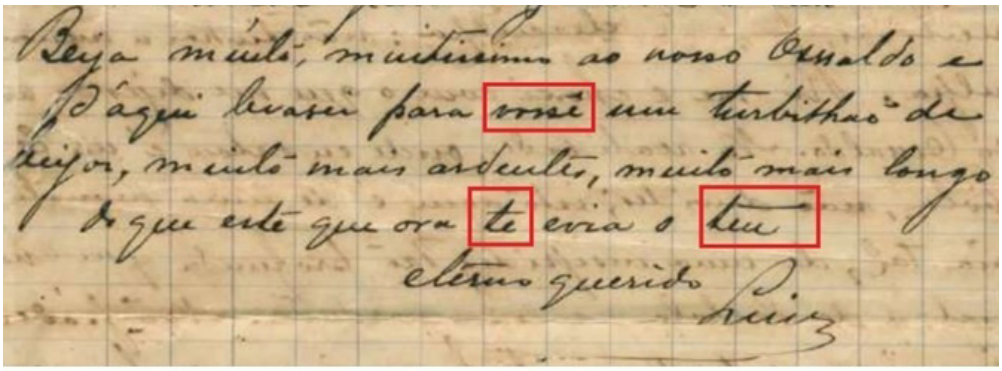
Imagem 3: JP. RJ, 7.10.1891



Fonte: Ferreira Filho (2023).

A *autenticidade* das amostras históricas é outro parâmetro igualmente relevante aos estudos linguísticos de sincronias passadas. Considerando o fato de as evidências linguísticas de sincronias poderem ser resgatadas tão somente a partir da sua expressão escrita – evidenciando a produção de punhos de escreventes escolarizados – é importante demarcar o que é expressão da força da norma-padrão, em relação às evidências de traços que refletem punhos habilidosos, dotados de um evidente domínio da escrita (o "filtro da escrita" (Romaine, 2010 [1982]). Acrescente-se ainda o fato de as fontes históricas nos permitirem passar por evidências de hipercorreção, mistura dialetal e "erros" (Labov, 1994, p. 11), o que pode ofuscar a expressão linguística do vernáculo de sincronias passadas, ainda que estejamos também cientes de haver em cena tão somente indícios positivos relativos aos textos históricos que sobreviveram à força do tempo, no interior dos arquivos. Nesse sentido, as cartas pessoais em análise nos permite mensurar, em certa medida, a expressão vernacular, através da produção escrita mais íntima de brasileiros cultos, entre fins do século XIX e o início do século XX. As imagens 4 e 5 ilustram algumas evidências da convivência de formas do paradigma de *tu* (*te, teu, tua, teus*) nas funções de complementação acusativa ("*te* envia", "*te*-ferir"), genitiva (*teu, tua, teus*) com formas do paradigma de *você* ("levarei *para você*"), inclusive em estrutura do redobro de clítico de 2SG que é retomado na forma do sintagma preposicionado "*a você*" na construção "havia [...] d *te*-ferir assim a v*ocê* [...]".

Imagem 4: GD. Lisboa, 11.08.1889

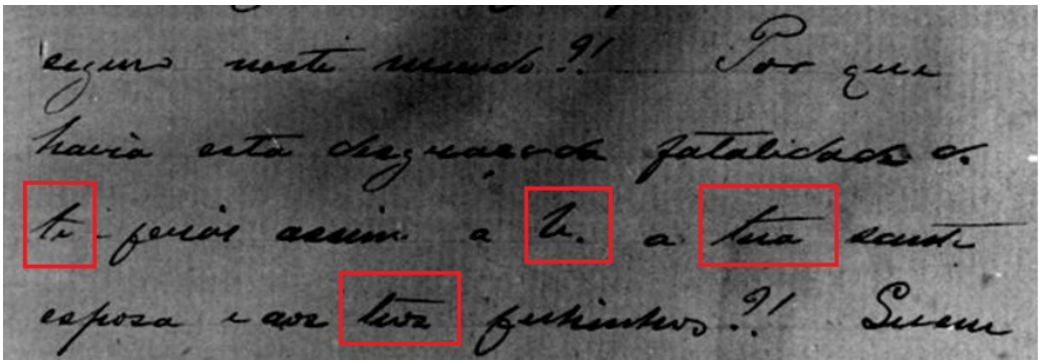


Fonte: Ferreira Filho (2023).

"[...] d'aqui levarei para *vossê* um turbilhão de beijos, muito mais ardentes, muito mais longo do que este que ora *te* evia o *teu* eterno querido Luiz. [...]"

(GD. Lisboa, 11.08.1889.)

Imagem 5: JP. Caeté, MG, 29.12.1896



Fonte: Luz (2015).

> "[...] Por que havia esta desgraçada fatalidade d. *te*-ferir assim a *Você* a *tua* santa esposa e aos *teos* filhinhos?! [...]"
>
> (JP. Caeté, MG, 29.12.1896.)

Nas imagens 4 e 5, é possível observar, através de formas dos paradigmas de *tu* (*te, teu, tua, teus*) e de *você* (*para você*), a expressão da norma de uso do PB, que segue, por sua vez, de encontro à norma-padrão, a partir da quebra do paralelismo formal (em termos tradicionais), embora tenhamos formas pronominais cuja referência semântica corresponde à 2SG. Apesar de haver em cena evidências da produção escrita de redatores familiarizados com as especificidades da norma-padrão e, portanto, habilidosos em sua manifestação escrita, lidamos com a expressão linguística dos redatores em cartas íntimas, o que tenderia normalmente a evidenciar um fluir da pena mais livre da rigidez da norma-padrão.

A *validade social e histórica* das fontes está consolidada na reconstrução da história de vida dos redatores, atendendo não só ao parâmetro da *autoria* (Hernández-Campoy e Schilling, 2012), mas também ao fato de poder ser difícil identificar a camada social a que pertence o redator e a estrutura social da comunidade idiomática[2] (Labov, 1994). Considerando o fato de termos em foco amostras de cartas pessoais, fidedignamente organizadas e editadas, no que se refere à *autoria* (cartas autógrafas) e à *autenticidade* (cartas pessoais) de missivas produzidas por redatores brasileiros, acreditamos estar diante de uma produção escrita legitimamente brasileira e, portanto, válidas (social e historicamente) para estudos linguísticos de sincronias passadas.

[2] "[...] we usually know very little about the social position of the writers, and not much more about the social structure of the community." (Labov, 1994, p. 11)

Para conduzir este estudo, tomamos por base os princípios da mudança de base laboviana (Weinreich *et al.*, 1968; Labov, 1994), com a utilização do programa computacional Goldvarb X. Seguimos, de um modo geral, as hipóteses e resultados da análise de Lucena (2016), voltados a outras cartas pessoais cariocas oitocentistas e novecentistas.

2. As amostras históricas: cartas pessoais de João Pinheiro e de Gonzaga Duque.

Neste estudo, utilizamos trinta e duas (32) cartas pessoais de João Pinheiro (doravante JP) distribuídas entre amorosas (6/32), familiares (7/32) e de amizade (19/32). Trata-se de um redator que nasceu, em 1860, no Serro (MG), morou no Rio de Janeiro, até o ano de 1870 (momento do falecimento do seu pai), concluiu o curso de Direito (Faculdade de Direito de São Paulo), tendo exercido as funções de professor, advogado, político (Secretário de Estado de Minas Gerais, Ministro do Interior do Governo Provisório, Governador de Minas Gerais e Deputado) e empreendedor (dono da fábrica de louças Cerâmica de Caeté), segundo Luz (2005).

No que se refere à Gonzaga Duque (Estrada) – doravante GD –, há, no contexto de vinte e seis (26) cartas pessoais, dezesseis (16) cartas amorosas, seis (6) cartas familiares e duas (2) cartas de amizade. Na sua trajetória profissional, GD deixa vestígios das suas atividades como crítico de arte, jornalista e escritor (Espindola, 2009, p. 14). Nascido no Rio de Janeiro, na 2ª metade do século XIX (1863) e falecido, já no 1º quartel do século XX (1911), concluiu o ensino secundário, mas parece não ter avançado pelo Ensino Superior (Fonseca, 2015, p. 38). Mesmo assim, trata-se de um redator que se mostra engajado no *mundo da língua escrita*, tendo em vista não só a sua atuação na criação dos periódicos *O Guanabara*, em fins do século XIX (1880), *Rio-Revista* e *Galáxia*, também em fins da era novecentista, nos anos de 1895 e 1897, *Mercúrio* e *Fon-Fon*, no início do século XX, 1901 e 1907, respectivamente, mas também através das suas contribuições em variados periódicos brasileiros. Acresce ainda o fato de o redator em questão ser autor de livros (romance, ensaio, diário). Trata-se, pois, de uma figura que pode ser considerada um legítimo "homem de letras" (Fonseca, 2015, p. 42), o que permite entendê-lo como um informante bem treinado em relação aos modelos de língua escrita (Ferreira Filho, 2023).

Considerando o fato de que as línguas humanas estão orientadas pelos fluxos e contrafluxos da variação e da mudança, entendemos que as relações sociais estabelecidas entre os informantes podem nos evidenciar indícios da história social (movimentos externos) que se interceptam com a história linguística (processos internos). Nesse sentido, de acordo com Conde Silvestre (2007) e Bergs (2012), é possível interpretar as redes

sociais como (i) fixas, densas (*close – knit, multiplex nature*) ou (ii) flexíveis, difusas (*loose – knit, uniplex nature*). Em termos gerais, a ideia é a de que informantes envolvidos em redes sociais mais fixas (densas), redes sociais menos diversificadas em relação às possibilidades de contato linguístico, tendem a se mostrar resistentes aos processos de variação e mudança linguística, enquanto as redes sociais difusas (flexíveis), caracterizadas pelo movimento dos informantes em intensas dinâmicas interna e externa ao seu país de origem, anunciam o contato com um maior número de pessoas (Conde Silvestre, Hernández-Campoy, 2004, 2005), o que tenderia a impulsionar a velocidade da implementação de fenômenos linguísticos em processo de mudança linguística (Bergs, 2012, 2005; Conde Silvestre, 2007). A partir dessas premissas, estaríamos, no caso, diante de dois redatores (GD e JP) que, em viagens internas e externas ao Brasil, se enquadram em redes sociais difusas, diversificadas, portanto, em função de distintas interações linguísticas como evidência da ampliação dos contatos linguísticos. Ainda que as redes sociais dos autores das cartas possam ser consideradas difusas, em certa medida, o que tenderia a favorecer a dinâmica de processos linguísticos de mudança, não podemos esquecer que estamos diante de um fenômeno linguístico de variação linguística que teve início em sincronias passadas (Lucena, 2016; Barbosa, 2018; Silva, 2023), mas permanece até hoje (Rocha, 2009; Arduin, 2005; Soares, 1999; Menon, 1997; Neves, 1993; Perini, 1985).

A distribuição das cartas pessoais, produzidas por um redator nascido no Rio de Janeiro (GD), entre os anos de 1869 e 1908, e por um redator mineiro (JP), entre os anos de 1884 e 1909, para a análise da dinâmica *teu/seu*, está exposta a seguir (Quadro 1), em que se registram os subgêneros das cartas pessoais, período e origem dos autores, controlados na análise variacionista como variáveis extralinguísticas.

Quadro 1: As amostras históricas: cartas pessoais dos séculos XIX e XX

Acervos	Redatores	Períodos	Subgênero da carta pessoal			
			Amor	**Familiar**	**Amizade**	Total
APM[3] (MG)	João Pinheiro (JP)	1869 - 1908	6	7	19	32
FCRB[4] (RJ)	Gonzaga Duque (GD)	1884 - 1909	16	8	2	26
Total	2 redatores	1869 - 1909	22	15	21	58

[3] Arquivo Público Mineiro (APM).
[4] Fundação Casa de Rui Barbosa (FCRB).

3. A dinâmica *teu/seu* em cartas brasileiras oitocentistas e novecentistas: uma breve revisão do tema e os parâmetros da análise linguística.

Os trabalhos sobre a variação *teu/seu* embasados em dados linguísticos de sincronias passadas evidenciam o uso progressivo do *seu* na referência à 2SG, em distintos níveis de produtividade. Assim sendo, apresentamos uma breve descrição dos principais resultados de estudos embasados em cartas pessoais oitocentistas e novecentistas (cariocas, baianas e pernambucanas), conforme as discussões de Lucena (2016), Barbosa (2018) e Silva (2023).

A partir de trezentos e sessenta e três cartas confeccionadas por escreventes nascidos e/ou residentes no espaço do Rio de Janeiro, Lucena (2016) estudou a alternância *teu/seu* entre os anos de 1870 e 1970. A conjectura inicial é a de que a referência ao sujeito através das formas *tu* e *você* poderia conduzir ao uso dos possessivos *teu* e *seu*, respectivamente, impulsionando um contexto de harmonização morfossintática entre as formas pronominais nominativas e genitivas, respectivamente, em relação à 2SG. Em termos gerais, a autora constatou a preferência pelo *teu* (76%), tendo em vista o uso do *seu* (24%) em relação a um total de 1376 ocorrências de possessivos pronominais de 2SG nas cartas pessoais. Lucena (2016, p. 77) constata que o *seu* prevalece nas cartas de *você-sujeito* (70%), ainda que se tenha se apresentado com níveis baixos de produtividade nas cartas sem referência explícita (11%, 37/336), nas cartas de *tu-sujeito* (10%, 33/336), nas cartas mistas (9%, 31/336).

Na escrita dos sertanejos baianos, Barbosa (2018) também se voltou à alternância *teu/seu*. A partir da análise de noventa e uma (91) cartas pessoais redigidas entre os anos de 1996 e 2000 por redatores semialfabetizados das áreas rurais baianas dos municípios Riachão de Jacuípe, Conceição do Coité e Ichu, o autor identificou que o possessivo *seu* foi registrado, de um modo geral, em 91% dos dados (168/185), restando ao *teu* a sua produtividade, em tão somente 9% (17/185). A variação entre as formas *teu/seu* evidencia a preferência pelo *seu*, em 90% (161/178) dos dados das cartas sem referência pronominal explícita, o que parece ser um indício de um processo de estabilização do *seu* de 2SG na escrita dos sertanejos baianos.

Em cento e cinquenta e três (153) cartas amorosas pernambucanas, Silva (2023) se voltou para a dinâmica *teu/seu* na produção escrita de três casais da região do Sertão do Pajeú, entre as décadas de 50 e 90 do século XX. De um modo geral, o *seu* liderou em 65% dos dados (202/313) ao lado do *teu*, em 35% (111/313). Como expressão de um quadro simétrico observou o autor que o *teu* se deu preferencialmente, nas cartas de *tu-sujeito*

(75%, 70/93), enquanto o *seu* predominou nas cartas de *você-sujeito* (179/221). Apesar de o *seu* ter prevalecido nas cartas de *você-sujeito*, Barbosa confirmou a sua hipótese de que o *seu* também tenha se deixado evidenciar nas cartas de *tu-sujeito* (25%, 23/93) das cartas pernambucanas novecentistas (Souza, 2023, p. 15).

A variação entre os possessivos *teu* e *seu* é analisada a partir de parâmetros linguísticos[5] tais como o pronome-sujeito; a animacidade do sintagma possessivo; o tipo de posse – alienável ou não, o tipo de posse inalienável (parentesco, partes do corpo, outros inalienáveis) e a presença ou ausência de artigo diante do possessivo, tema muito recorrente na literatura. Nesse caso, nossa hipótese era a de que o uso do artigo seria mais frequente diante de *seu,* com o objetivo de acrescentar o traço [+ definido] ao utilizá-lo de referência à segunda pessoa, hipótese referendada por Palladino Netto (2009) em sua tese de doutorado. Em relação aos extralinguísticos, controlamos o subgênero das cartas pessoais, a seção, o período e a origem dos autores das cartas.

Em relação ao subgênero da carta (*amorosa*, *familiar* e *de amizade*) e à seção da carta (*saudação inicial, núcleo, saudação final*, *post scriptum*), os objetivos seriam reconhecer se algum subgênero da carta pessoal e se alguma das suas seções teriam ou não fomentado a produtividade do *seu.* Controlamos, ainda no âmbito dos fatores também extralinguísticos, o *período* e a *origem do redator da carta*. Os períodos em que as cartas pessoais foram produzidas (entre os anos de 1869 e 1910) e a origem dos redatores (um redator carioca e um redator mineiro) também foram controlados, para que monitorássemos se o momento entre o fim da era oitocentista e o início dos novecentos no Brasil teriam ou não influenciado os redatores carioca e mineiro em relação ao uso do *seu* para a 2SG.

No que diz respeito ao pronome-sujeito de 2SG, partimos da hipótese de que o *tu-sujeito* tenderia a estimular o uso do *teu*, enquanto o *você-sujeito* tenderia a promover o *seu*, configurando um quadro sistemático e simétrico, em relação às formas dos paradigmas de *tu* e de *você*, respectivamente. Para averiguarmos a força do contexto de sujeito de 2SG (*tu* e *você*) a impulsionar ou não o uso dos possessivos *teu* e *seu*, correlacionamos os dados aos contextos de nominativo (*tu* e *você*) e de genitivo (*teu* e *seu*), aspecto já averiguado por Lucena (2016, p. 174), em relação às cartas do Rio de Janeiro.

> (01) Dados de *teu* e *tua* de *tu-sujeito*:
> a. "Pimentel. É resposta da *tua* carta d. 5. Podias$_{tu\text{-}suj}$ e devias$_{tu\text{-}suj}$ contar commigo para o negocio [...]" (JPS. Caeté, 08.12.1899)
> b. "[...] O que é feito dessas alegres creaturas que o *teu* espirito aparvalhado arrancou do pé de nós que hatresmezes choramos sem descanso? E

[5] Para o controle do contexto de pronome-sujeito de 2SG em correlação aos possessivos (*teu/seu*) voltamo-nos às cartas de "tu" exclusivo, às cartas de "você" exclusivo, às cartas de *tu/você* (mistas) e às cartas com formas nominais de tratamento ("cara madrinha").

o caramanchão como um inquisidor-mór me disse: Estou vingado, peralvilho, has$_{tu\text{-}suj}$ de chorar como eu a ausencia destas duas creanças amorosas e puras eu vi crescer [...]" (GD. RJ, Botafogo, 08.11.1884. C2)

(02) Dado de *seu* em contexto de *você-sujeito*:
"[...] Accuso o recebimento sobre a vinda d. Doutor Gorceix e desejos d. visitar a minha fabrica e tambem ultimamente o *seu* livro sobre Minas. [...] Queira você$_{você\text{-}suj}$ por mim transmittir ao Doutor Gorceix o meu desvanecimento pela sua visita á fabrica [...]" (JPS. Caeté, 30.12.1904)

(03) Dados de *tua* e *seu* em contexto de sujeito misto (*tu/você*):
a. "[...] Ainda outro dia em *tua* casa, ella tão feliz com seus filhinhos, juncto d. ti [...] Vai indo esta carta desconexa, mas vai sendo escripta com o coração. Precisas$_{tu\text{-}suj}$ viver; acho conveniente que a Dona Nicota e o Neusinho venhão ficar uns tempos em *tua* casa cuidando dos meninos e que Você$_{você\text{-}suj}$ venha passar uns tempos commigo [...]" (JPS. Caeté, 29.12.1896)
b. "[...] Senti, hoje, saudades de nossos filhinhos, creio, porém, que minhas nem vossê$_{você\text{-}suj}$ nem eles sentirão. Vai com esta uma carta que ahi porás$_{tu\text{-}suj}$ em enveloppes[...] Seria enorme favor que prestarias$_{tu\text{-}suj}$ se arranjasse dinheiro com *tua* mãe para me comprar na rua do Theatro – Casa Cavallier [...]" (GD. Porto Grande, 08.01.1889)

No que diz respeito ao traço de animacidade do referente do sintagma possessivo, assumimos a proposta de Oliveira e Silva (1982) em relação às amostras do português quinhentista, seiscentista e setecentista. Nesses *corpora* históricos, a autora constata o fato de o *seu* de 3SG ter sido fomentado pelo possuidor [+animado], tendo passado, a partir do século XVIII, a ser estimulado pelo possuidor [- animado]. Constata, pois, a autora que o uso do *seu* em relação a um referente [- animado] parece se espraiar, a partir da era setecentista do português (século XVIII). Nas cartas cariocas, Lucena (2016, p. 98) constatou "o crescimento gradual do pronome *seu* nas construções com traço [inanimado], ao passo que as construções [+ humano] com o pronome diminuem". Com base nos resultados de Lucena (2016), prevemos que o *seu* já se apresente, nas cartas em análise, como evidência da relação "possuidor-possuído", voltado aos referentes [- animado]. Os exemplos (04) e (05) ilustram casos de *teu* e *seu* com referentes [+ animado] e [- animado], respectivamente.

(04) Dados de *teu* e *seu* com referente [+ animado]:
a."[...] Adeosmeo Edmundo, recebe o coração d. *teo* amigo [...]" (JPS. Caeté, 29.12.1896.)
b. "[...] Sou *Seo* Sobrinho que te ama de coração João Pinheiro [...]" (JPS. OP, 21.12.1869.)

(05) Dados de *teu*, *tua*, *seu* e *sua* com referente [- animado]:
a. "[...] Recebi o *teu* cartão d. parabens e o resultado da eleição no Piranga, muito obrigado por tudo. [...]" (JPS. Caeté, 25.02.1905.)
b. "[...] nem sei que relações você mantem com o mesmo: não descuides

> disto [...] Recebi *seo* telegramma pedindo papel [...] mas tendo recebido o *seo* telegramma dizendo que o Laemert vendia a 6:200, voltei ao homem que fez o favór d. deixal-o por 7.000. [...]" (JPS. RJ, 25.10.1891)
> c. "Julinha, Recibi ha dias a *tua* cartinha, e, se me não engano, foi escritpta em resposta a que te enviei da Bahia. (GD. Lisboa, 26.07.1889)
> d. "Meu caro Compadre e Amigo Aqui tenho estado para me informar da *sua* saúde e isso por me não possivel ir á sua casa [...]" (GD. 05.07.1907)

No que diz respeito à semântica da posse, assumimos à luz de Neves (2000, p. 476-477) a intepretação de que "o possessivo remete ao possuidor; o substantivo indica o possuído" (Neves, 2000, p. 476). A partir dessa noção de posse prototípica consideramos a ideia de que ao transferirmos o "algo possuído", temos a posse alienável (transferível), ao passo que é possível evidenciarmos contextos em que o "algo possuído" é inseparável do possuidor (intransferível). Na análise de Lucena (2016), a proposta é a de ampliação da perspectiva de análise de Neves (2000), ao incluir entre os *alienáveis* e *os inalienáveis*, em extremos opostos, a *extensão de posse*. O conceito de *extensão de posse* abarca as evidências de *posse não-canônica*". No âmbito da *extensão de posse*, estão as evidências não só nos contextos dos substantivos deverbais, mas também nas relações interpessoais não-consanguíneas materializadas nas dinâmicas de interação entre o remetente e o destinatário. Os exemplos (06), (07) e (08) apresentam evidências da semântica da posse em relação aos traços de *alienabilidade* ("tua carta", "sua carta"), de *extensão de posse* ("seu pedido", "teu esposo") e de *inalienabilidade* ("seo carinho", "tua lembrança", "tua imagem, "sua saúde"), respectivamente.

> (06) Dados de *tua* e *sua* em contexto de posse *alienável*:
> a. "[...] Feca [...] Sobre o assumpto da *tua* carta, me escreveo o Doutor Alvim, mas falando no estabelecimento do ensino technico d. agricultura em Juiz d. Fora [...]"(JPS. Caeté, 03.05.1903)
> b."[...] Você em *sua* carta teve a habilidade d. bulir com todos os pontos fracos da minha industria. [...]" (JPS. Caeté, 28.01.1901)

> (07) Dados de *teu* e *seu* em contexto de extensão de posse (substantivos deverbais e relações pessoais não-consanguíneas):
> a. "[...] Ao conceder-lhe a exoneração pedida, cumpre-me agradecer-lhe os inestimaveis serviços que acaba de prestar ao meu governo [...] Não annuiria ao *seu* pedido, si não tivesse certeza de que, na industria que tão competentemente vai dirigir, prestará ao Estado de Minas serviços tão relevantes [...]" (JPS. BH 08.06.1907)
> b. "[...] Podes despachar a bagagem para Ouro Preto [...] Um beijo em nosso filho e aceita saudades do *Teu* esposo muito amante João Pinheiro. [...]" (JPS. RJ, 09.10.1891)

(08) Dados de *tua* e *seu* em contexto de posse *inalienável*:
a. "[...] Dirijuli amado tio minhas felicitações, como a pessoa, que tantas vezes me há manifestado *seo* carinho. Dezejo-lhe completa ventura [...]" (JPS. OP, 21.12.1869)
b. "[...] mas só a *tua* lembrança minha Helena era capaz de consolar-me. [...]" (JPS. RJ, 14.02.1891)
c. "[...] temo ser te infiel, e vivo com a *tua* imagem como os martyres viviam com a de Christo. Farás o mesmo? Aposto em como já me esqueceste. Fazes bem. [...]" (GD. Lisboa, 10.09.19.. C26)
d. "Meu caro Compadre e Amigo Aqui tenho estado para me informar da *sua* saúde e isso por me não possivel ir á sua casa. [...]" (GD. 05.03.1907. C24)

No que se refere à posse inalienável, identificamos, com base em Lucena (2016, p. 83), à luz de Neves (2000), a posse marcada pelos traços parentesco ("tua mamãe", "sua filha", "seo sobrinho", "sua sobrinha"), de partes do corpo humano ("teu cérebro", "teus lábios") e *outros inalienáveis* como expressão, por exemplo, emoções ("tuas saudades"), traços idiossincráticos ("teu esforço", "sua companhia", "tua lembrança", "sua opinião"), justamente por se mostrarem inseparáveis do seu possuidor, conforme ilustrado de (09) a (11), respectivamente.

(09) Dados de *tua, seu* e *sua* em contexto de posse inalienável (parentesco):
a. "[...] Julinha vê se consegues de *tua* mamãe a volta breve para a côrte. [...]" (GD. 06.11.1884. C1)
b. "[...] Agora o que eu peço a senhora [...] que traga Julia o mais breve possivel. A senhora procedendo assim [...] entregará a *sua* filha á felicidade, será ainda mais boa do que é. [...]" (GD. 06.11.1884. C1)
c. "[...] Sou *Seo* Sobrinho que te ama de coração João Pinheiro [...]" (JPS. OP, 21.12.1869)
d. "[...] Helena de Barros é o nome da *sua* sobrinha [...]" (JPS. OP, 04.03.1890)

(10) Dados de *teu* em contexto de posse inalienável (partes do corpo):
a. "[...] como esforçares te por impôrao *teu* cérebro o que elle tem preguiça de adquiri? [...]" (GD. RJ, 22.07.1906.C23)
b. "[...] Mil beijos no meu Oswaldo 110 e um ardente beijo nos *teus* labios... (mentirosos?) que te envia entre saudades e amor o teu eterno amante. [...]" (GD. Lisboa, 10.09.19.. C26)

(11) Dados de *teu, tua* e *sua* em contexto de *outros inalienáveis*:
a. "[...] Mas para Você, meo infeliz amigo, ella vive d. certo na tua lembrança, nas tuas *saudades* nos teos filhinhos, d.' ond. ella está, ella te-continúa amar [...]" (JPS. Caeté, 29.12.1896.
b. "[...] Não percebes que do *teu* esforço, esforço de acordo com a tua idade, é que virá o bom resultado dessa conquista? Que aspiras tu? [...]" (GD. RJ, 22.07.1906)

c. "[...] Aqui, nestas tiras, vão as immensas saudades que eu tenho sentido da *sua* companhia e meu abraço todo o agradecimento do seu velho amigo GD" (GD. 09.11.1909)
d. "[...] mas só a *tua* lembrança minha Helena era capaz de consolar-me. (JPS. RJ,14.02.1891)
e. "Como verá pelos termos da carta ao Presidente do Espirito Santo, a missão tem um caracter inteiramente reservado [...] Em todo caso convem ler o memorial... e dar-me a *sua* opinião. [...]" (JPS. BH, 20.08.1907)

4. Análise dos resultados: a variação *teu/seu* de 2SG.

Com base nos resultados da análise, através dos programas computacionais, foi possível concluir que apenas dois dos grupos de fatores controlados se mostram relevantes para a implementação do possessivo *seu*: (i) pronome-sujeito de 2SG correlacionado ao possessivo "seu" (Tabela 1); e (ii) subgênero da carta pessoal (Tabela 2).

De um modo geral, há indicações de que o contexto em que o possessivo *seu* começa a ocupar o lugar de *teu* é aquele em que o sujeito corresponde a uma forma nominal de tratamento (*Dona Mariquinhas, Doutor Alvim...*) e, nas cartas amorosas, o uso de *teu* ainda é bloqueado. Embora só haja quatro ocorrências em que o sujeito é exclusivamente *você* (apenas quatro ocorrências nas cartas mineiras de JP), o uso do possessivo *seu* é categórico (100%).

Tabela 1: Distribuição dos dados de *seu* de 2SG em relação ao tratamento do interlocutor nas cartas cariocas e mineiras em análise

Seu (possessivo de 2SG)			
Nível de seleção			
Pronome-sujeito de 2SG	**Oco/total**	**%**	**P.R**
Formas nominais de tratamento	**42/46**	**91.3%**	**0.954**
Cartas de *tu*-sujeito exclusivo	13/109	11.9%	0.313
Cartas mistas (*tu/você*)	9/48	18.8%	0.246
Total	64/203	31.5%	**Input: 0.235**

A correlação entre as funções de nominativo e de genitivo nas cartas brasileiras analisadas (cariocas e mineiras) evidencia que as formas nominais de tratamento tendem a impulsionar o possessivo *seu* de 2SG, contrariando a conjectura inicial de o uso do *você-sujeito* ser o deflagrador, ao menos nas cartas brasileiras em análise, do possessivo *seu*. Na verdade, com base nos pesos relativos de 0.954 e frequência de 91.3%, é possível verificar que é nas cartas com formas tratamentais que o uso do *seu* prevalece, pelo menos, nas

amostras de cartas brasileiras em análise. As formas tratamentais acionam a concordância com formas de 3ª pessoa, o que, em termos formais, parece se refletir sintomaticamente no uso do *seu,* cuja função etimológica nos remete à 3SG, ainda que a referência semântica passe a ser, no decorrer do tempo, a de 2SG. Os exemplos (12) e (13) evidenciam evidências do *seu* (*sua*) em cartas de formas nominais de tratamento exclusivo.

> (12) "[...] *Dona Mariquinhas* [...] Conversando depois sobre a *sua* vinda estivemos de accordo que *a senhora* deveria voltar n'esse mez não só por causa dos *seus* negócios como tambem por causa do mez de Dezembro que se avesinha e que é de grandes febres ahi. [...]" (GD. 08.11.1884. C2)

> (13) "*Doutor Alvim* [...] E como vai tudo correr com a *sua* responsabilidade o *Doutor* vera o que mais convem. [...]" (JPS. OP, 18.02.1892)

Tabela 2: Distribuição do possessivo *seu* de 2SG em relação aos subtipos de cartas pessoais cariocas e mineiras em análise

Seu (possessivo de 2SG)			
	Nível de seleção		
Subgênero da carta pessoal	**Oco/total**	**%**	**P.R**
Amizade	**34/71**	**47.9%**	**0.799**
Familiar	**32/56**	**57.1%**	**0.602**
Amorosa	3/80	3.8%	0.180
Total	**69/139**	31.5%	**Input: 0.235**

O subgênero da carta pessoal foi a 2º variável selecionada como relevante e corresponde às cartas de amizade (0.799) e familiares (0.602) os contextos propulsores do inovador uso do possessivo *seu*, de referência à 2SG: inversamente, são, as cartas amorosas (0.180), que ainda mantêm o do possessivo *teu,* fato respaldado pelos pesos relativos ilustrados. Se, por um lado, o número de cartas analisadas aqui é baixo, por outro, não se pode esquecer que há evidências do possessivo *seu* de referência à 2SG em outras análises (Oliveira e Silva, 1982; Lucena, 2016).

Outro ponto importante é que a variação *teu/seu* é praticamente a mesma, no período de 41 anos, confirmando que mudanças morfossintáticas ocorrem lentamente. A variável "origem do autor da carta", isoladamente, apresenta significância 0.000, opondo Rio de Janeiro a Minas Gerais (Tabela 3).

Tabela 3: Distribuição do possessivo *seu* de 2SG em relação às cartas cariocas e mineiras

Seu (possessivo de 2SG) Nível 1[6]			
Origem do autor	**Oco/total**	**%**	**P.R**
Rio de Janeiro	26/121	21.5%	0.369
Minas Gerais	43/86	**50.0%**	**0.681**
Total	**69/139**	**33.2%**	**Input: 0.319**

À medida que essa variável interage com outras variáveis, tais como os tipos de sujeito e de carta pessoal, a relevância diminui, mas é digno de nota o fato de, (i) em Minas Gerais, o possessivo *seu,* das últimas décadas do século XIX até, pelo menos, a 1ª década do século XX, passar a concorrer, percentualmente, em igualdade de condições, com o possessivo *teu* (50% *versus* 50%); enquanto, (ii) no Rio de Janeiro, esse espraiamento parece ser ainda bloqueado, no início do século XX (21.5% *versus* 78.5%). Resta confirmar, com a ampliação da amostra, se há uma diferenciação de uso nesses dois espaços da região sudeste brasileira.

Considerações finais

É importante termos em mente que este estudo (i) se at**ém** à produção escrita de apenas dois informantes brasileiros (nascidos e/ou residentes no Rio de Janeiro e em Minas Gerais), que produziram suas cartas pessoais entre fins do século XIX e o início do século XX e (ii) a quantidade de cartas pessoais não est**á** equanimemente distribuída por seus subtipos (amor, amizade e familiar). Esse aspecto representa uma contingência peculiar aos trabalhos no âmbito da sociolinguística histórica, conduzidos por pesquisadores que levantam, editam e analisam amostras a que tiveram acesso, representativas de realidades linguísticas pretéritas. Feitos esses esclarecimentos iniciais, passemos à sistematização das principais generalizações acerca da dinâmica *teu/seu* na amostra em foco.

O possessivo *teu* apresenta maior frequência que a variante *seu*, ainda que este último já ocorra em certos contextos, de preferência, naqueles em que o sujeito corresponde a formas nominais de tratamento. De todo modo, nesta amostra, há um equilíbrio entre as duas formas – *teu/seu* – na amostra de cartas mineiras (Tabela 3).

Essa variação de uso dos possessivos *teu/seu* parece estar formalmente alinhada à ativação da concordância com estratégias (verbais e pronominais) de 3ª pessoa do sin-

[6] No nível 1 de análise, ainda não é levada em conta a interação com outros grupos de fatores. Esta variável que, no primeiro momento, parece ser relevante, não é depois selecionada, talvez pela distribuição irregular por tipo de carta.

gular, ainda que, semanticamente, de referência à 2SG. Em síntese, seria possível supor que a perda gradual da distinção entre possessivos de segunda e terceira pessoas não seja isolada, mas esteja encaixada, representando uma das repercussões gramaticais da reorganização do sistema pronominal do PB, em virtude da inserção do *você* no quadro de pronomes (Lopes 2007; Lopes e Cavalcante, 2012).

Este estudo constituiria, assim, mais uma comprovação no sentido de que é necessário analisar outras amostras, levar em conta outros tipos de textos e considerar uma gama de fatores históricos, linguísticos, extralinguísticos e discursivos, na busca de explicações para a variação e mudança de fenômenos morfossintáticos, em sincronias passadas e recentes. Como afirmam Weinreich *et al.* (1968, p.188), ao final de seu texto, ao tratar dos princípios gerais e da natureza da mudança linguística,

> A generalização da mudança por toda a estrutura linguística não é uniforme nem instantânea; envolve a covariação de mudanças associadas por períodos substanciais de tempo, e se reflete na difusão de isoglossas por áreas do espaço geográfico.[7]

Referências

ARDUIN, J. *A variação dos pronomes possessivos de segunda pessoa do singular teu/seu na região sul do Brasil*. 2005. 123 f. Dissertação (Mestrado em Linguística) - Centro de Comunicação e Expressão. Universidade Federal de Santa Catarina, Florianópolis, 2005.

ARGOTE, J. C. *Regras da lingua portugueza, espelho da lingua latina*. Lisboa: Officina da Musica, 1725.

BARBOSA, G. M. O. *O uso dos pronomes possessivos teu e seu em cartas pessoais de sertanejos baianos do século XX*. 2018. 242 f. Dissertação (Mestrado em Estudos Linguísticos) - Faculdade de Letras, Universidade Estadual de Feira de Santana, Feira de Santana, 2018.

BARROS, J. *Grammatices rudimenta*. 32 f., enc. : perg., il. color. 1538.

BARROS, J. G*rammatica da lingua portuguesa*. Olyssipone: *apud* Lodouicum Rotorigiu[m], Typographum, - 60 f. 1540.

BERGS, A. The Uniformitarian Principle and the Risk of Anachronisms in Language and Social History. *In:* HERNÁNDEZ-CAMPOY, J. M.; CONDE SILVESTRE, J. C. *The Handbook of Historical Sociolinguistics*. Oxford: Wiley-Blackwell, 2012. p. 80-98.

[7] "4. The generalization of linguistic change throughout linguistic structure is neither uniform nor instantaneous; it envolves the covariation of associated changes over substantial periods of time, and is reflected in the diffusion of isoglosses over areas of geographical space".

BERGS, A. *Social Network Analysis and Historical Sociolinguistics*. Berl*in:* Mouton de Gruyter, 2005.

CONDE SILVESTRE, J. C.; HERNÁNDEZ-CAMPOY, J. M. A sociolinguistic approach to the diffusion of Chancery written practices in late fifteenth century private correspondence. *Neuphilologische Mitteilungen*, n. 105, v. 2, p. 135-52, 2004.

CONDE SILVESTRE, J. C.; HERNÁNDEZ-CAMPOY, J. M. Sociolinguistic andngeolinguistic approaches to the historical diffusion of linguistic innovations: incipient standardization. *International Journal of English Studies*, Murcia, v. 5, n. 1, p. 101-34, 2005.

CONDE SILVESTRE, J. C. Sociolinguística histórica. Madrid: Gredos. 2007.

ESPINDOLA, A. F. *Gonzaga Duque – Vida na Arte: uma concepção artístico-filosófica*. 2009. 242 f. (Dissertação) Mestrado - Universidade do Sul de Santa Catarina, Palhoça, 2009.

FONSECA, B. O. *Gonzaga Duque e Revoluções Brasileiras: um olhar para a História do Brasil*. 2015. 160 f. (Dissertação de Mestrado) - Faculdade de Educação, Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

FERREIRA FILHO, J. *Redescobrindo 'Gonzaga Duque' por meio de sua produção escrita: edição de cartas pessoais e descrição de traços paleográficos*. 2023. 204 f. (Dissertação) Mestrado em Estudos Linguísticos - Faculdade de Letras da Universidade Federal de Minas Gerais, Belo horizonte, 2023.

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*. Blackwell Publishing Ltd. 2012, p. 63-79.

LABOV, W. *Principles of Linguistic Change: Internal Factors*. Cambridge: Blackwell Publishers, 1994.

LIMA, A.; MARCOTULIO, L. L.; RUMEU, M. C. B. Experiências metodológicas em constituição de corpora: pistas para um pesquisador iniciante. *In:* CASTILHO, A. T. *História do português brasileiro: corpus diacrônico do português brasileiro*. São Paulo: Contexto, 2019, v. 2, p. 68-91.

LOBATO, A. J. R. *Arte da grammatica da lingua portugueza*. Lisboa: Regia Officina Typografica, - XLVIII, 253 p. 1770.

LOPES, C. R. S. Pronomes pessoais. *In:* BRANDÃO, S. F.; VIEIRA, S. R. (Org.). *Ensino de gramática: descrição e uso*. São Paulo; Contexto, 2007, p. 103-114.

LOPES, C. R. S.; CAVALCANTE, S. R. A cronologia do voceamento no português brasileiro: expansão de você-sujeito e retenção do clítico-te. *Lingüistica* (Madrid), v. 25, p. 30-65, 2012.

LUCENA, R. O. P. *Pronomes possessivos de segunda pessoa: a variação teu/seu em uma perspectiva histórica*. 2016. 203 f. Tese (Doutorado em Letras Vernáculas) - Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

LUZ, R. D. O *Tratamento na Produção Epistolar de João Pinheiro da Silva: análise sociopragmática de TU x VOCÊ e respectivas formas gramaticais*. 2015. Vol. I. 88 f. Vol. II. 163 f. Dissertação (Mestrado em Letras: Estudos da Linguagem) - Universidade Federal de Ouro Preto, Mariana, 2015.

MENON, O. P. S. *Seu/de você: variação e mudança no sistema dos possessivos*. *In:* HORA, D. Diversidade linguística no Brasil. João Pessoa, 1997, Ideia, p. 79-92.

NEVES, M. H. M. Possessivos. *In:* CASTILHO, A. T. *Gramática do Português falado. Volume III: As abordagens*. Campinas, 1993, v. 3.

NEVES, M. H. M. *Gramática de usos do português*. São Paulo: Editora UNESP, 2000.

NÚÑEZ CONTRERAS, L. *Manual de paleografía: fundamentos e história de la escritura latina hasta el siglo VIII*. Madrid: Cátedra; 1994.

OLIVEIRA, F. *Grammatica da lingoagem portuguesa*. Lixboa: Casa d' Germão Galharde, 27 Ianeyro, 38 f. 1536.

OLIVEIRA E SILVA, G. M. *Estudo da regularidade na variação dos possessivos no Português do Rio de Janeiro*. 1982. 458 f. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1982.

PALLADINO NETTO, L. *Uma edição de cartas de mercadores portugueses do século XVIII e o uso variável do artigo diante do possessivo*. 2009. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.

PERINI, M. O surgimento do sistema possessivo do português coloquial: uma interpretação funcional. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, São Paulo, v. 1, nºs 1 e 2, p. 1-16, 1985.

ROCHA, F. C. F. *A alternância nos pronomes pessoais e possessivos do português de Belo Horizonte*. 2009. 107 f. Dissertação (Mestrado em Letras) - Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, 2009.

ROMAINE, S. *Socio-Historical Linguistics: It's Status and Methodology*. Cambridge: Cambridge University Press, 2010 [1982].

SILVA, R. S. *"Carlinda me falou que tu não falou a sua mãe a respeito disto": a variação teu/seu no paradigma de segunda pessoa do singular em cartas de amor interioranas do século XX*. 2023. 125 f. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Universidade Federal de Pernambuco, Recife, 2023.

SOARES, A. S. F. *Segunda e Terceira pessoa – O pronome possessivo em questão: Uma análise variacionista*. 1999. 116 f. Dissertação (Mestrado em Linguística) - Universidade Federal do Paraná, Curitiba, 1999.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. Empirical foundations for a theory of language change. *In:* LEHMANN, W.; MALKIEL, Y. *Directions for historical linguistics*. Aust*in:* University of Texas Press, 1968.

http://doi.org/10.5007/1984-8420.2025.e100972

AS FORMAS "NÓS" E "A GENTE" NAS CARTAS PESSOAIS DA AMOSTRA FAMÍLIA ARTHUR REIS

"NÓS" AND "A GENTE" IN THE PERSONAL LETTERS SAMPLED ON THE ARTHUR REIS FAMILY

Flávia Santos Martins | Lattes | flaviasantos@ufam.edu.br
Universidade Federal do Amazonas

Edson Galvão Maia | Lattes | edson.galvao@ifam.edu.br
Instituto Federal de Educação, Ciência e Tecnologia do Amazonas

José Fabrício Affonso Ferreira Júnior | Lattes | affonsofabricio@gmail.com
PPGL-Universidade Federal de Santa Catarina

Romário Neves Coelho | Lattes | romarioneves16@hotmail.com
Universidade Federal do Amazonas

Resumo: Este artigo tem como objetivo descrever as formas *nós* e *a gente* na função de sujeito em cartas pessoais da família do amazonense Arthur Reis, escritas entre 1940 e 1980. Como embasamento teórico-metodológico, utilizamos a Sociolinguística Histórica (Conde Silvestre, 2007) a qual se ancora na Teoria da Variação e Mudança (WLH, 2006 [1968], Labov, 1994). Para a realização deste estudo, selecionamos 127 cartas que foram transcritas e editadas pela equipe do projeto intitulado PHPB/AM. Dessas cartas, em 83 havia dados da expressão de pronomes P4, expressos e nulos, totalizando 283 ocorrências, sendo 260 (91,9%) da forma *nós* e 23 (8,1%) da forma *a gente*. Essa variação é encontrada, particularmente, nas cartas de duas missivistas, Iria e Alésia, a partir da década de 1940. A nossa principal hipótese, a de que os contextos linguísticos que mais atuam sobre o uso de *a gente* é o *tipo de referente* e o *preenchimento do sujeito*, foi confirmada, pois o uso de *a gente* na escrita dessas duas remetentes é favorecido, linguisticamente, pelo *sujeito expresso*, *pelo traço [-determinado]* e pelo *tempo presente*.

Palavras-chave: Variação e mudança. Cartas pessoais. P4. Arthur Reis. PHPB/AM.

Abstract: The aim of this article is to describe the forms nós and *a gente* in the func-

tion of subject in personal letters from the family of Arthur Reis, from Amazonas, written between 1940 and 1980. As a theoretical-methodological basis, we used Historical Sociolinguistics (Conde Silvestre, 2007), which is anchored in the Theory of Variation and Change (WLH, 2006 [1968], Labov, 1994). For this study, we selected 127 letters which were transcribed and edited by the PHPB/AM project team. From these letters, 83 had data on the expression of P4 pronouns, expressed and null, totaling 283 occurrences, of which 260 (91.9%) were in the form nós and 23 (8.1%) in the form a gente. This variation is found particularly in the letters of two missivists, Iria and Alésia, from the 1940s onwards. Our main hypothesis is that the linguistic contexts, which most affect the use of a gente, are the type of reference and the completion of the subject, it was confirmed, as the use of a gente in the writing of these two senders is favored linguistically by the expressed subject, the [-determined] trait and the present tense.

Keywords: Variation and change. Personal letters. P4. Arthur Reis. PHPB/AM.

1 Introdução

Sob a perspectiva da Sociolinguística Histórica (cf. Romaine, 1982; Conde Silvestre, 2007), o presente artigo objetiva investigar as formas *nós* e *a gente* como variantes da 1ª pessoa do plural (doravante P4), na função de sujeito, no Português Brasileiro (PB) em documentos escritos no estado do Amazonas, no século XX. Essa abordagem está amparada nos pressupostos gerais da Teoria da Variação e Mudança (cf. Weinreich; Labov; Herzog, 2006 [1968]), tais como: i) a variação é inerente ao sistema linguístico; ii) a heterogeneidade é ordenada, ou seja, há condicionadores linguísticos e extralinguísticos atuando sobre a variação e a mudança; iii) toda mudança é resultado de variação linguística; iv) a mudança não é abrupta; v) a variação e a mudança são encaixadas; vi) os métodos quantitativos auxiliam na explicação da variação e da mudança linguísticas.

Pesquisas realizadas a partir de dados de fala e escrita sobre o referido fenômeno têm registrado a implementação da forma *a gente* como pronome, num processo de gramaticalização[1], de acordo com Naro; Görski; Fernandes (1983); Omena (1996); Albán *et al.* (1986); Fernandes; Gorski (1986); Freitas; Albán (1991); Monteiro (1991); Lopes (1998, 2003, 2004); Omena (2003); Santos (2015); Araújo (2016); Vitório (2015); Scherre; Naro; Yacovenco (2018); Freire (2019); Araújo (2023), dentre outros. A expressão *a gente* tem sua origem no substantivo *gente*. Essa forma, enquanto expressão

[1] Segundo Cunha, Oliveira e Martelotta (2015), "a gramaticalização *stricto sensu* se ocupa da mudança que atinge as formas que migram do léxico para a gramática" (p. 43).

nominal, flexionava-se no feminino, podia ser determinada (*Aquela* gente, *Essa* gente *esperta*) e tem um caráter indeterminador e coletivo, além de concordar com o verbo na 3ª pessoa do singular (P3). À medida que essa forma lexical vai se gramaticalizando, ela assume algumas propriedades gramaticais, no caso de pronome: não ser determinado no sintagma nominal (*A gente* saiu), correlacionar-se a adjetivos femininos e masculinos em estruturas predicativas (*A gente* está bonita/bonito) e se referir semanticamente à P4 (*A gente* foi ao supermercado), porém, ao mesmo tempo, mantendo alguns traços nominais, conforme observaremos na seção 2.

Esses trabalhos têm mostrado, dessa forma, a concorrência entre as formas *nós* e *a gente* como pronomes para se referir à P4, evidenciando, sobretudo, que o aumento do uso de *a gente* no PB se deve, principalmente, ao tipo de referente (+/- determinado). Os resultados mostram que *a gente* é favorecido em contextos de indeterminação (*eu + eles / eu + todos*), um traço que é característico da forma de origem *gente*.

No Amazonas, até o momento, encontramos somente uma pesquisa, realizada no município de Benjamin Constant, sobre o fenômeno em foco, e apenas a partir de dados de fala (Reis, 2021). Portanto, não temos uma descrição ampla sobre a variação *nós* e *a gente* no Português falado na região e, muito menos, sobre o Português escrito, o que evidencia o ineditismo desta pesquisa. Com esse objetivo, tomamos como objeto de análise 127 cartas da amostra Família Arthur Reis, datadas entre os anos 1940 e 1980 e organizadas pelo projeto Para a História do Português do Amazonas (PHPB/AM).

Para essa investigação, buscamos, então, responder às seguintes questões:

i) Há variação no uso de *nós* e *a gente* na escrita registrada no Amazonas?

ii) Quais variáveis linguísticas e extralinguísticas estariam atuando no uso da forma inovadora *a gente* na escrita de missivistas manauaras de diferentes gerações?

Esperamos encontrar *a gente* como pronome nas cartas analisadas, tendo em vista que essa forma já é registrada na escrita do PB, desde o final do século XIX (Lopes, 2003; Souza, 2015; Borges, 2019; Monguilhott *et al.*, 2021), assim como, através da observação do tempo real, esperamos que essa forma seja mais produtiva a partir da década de 1970 (Lopes, 2004). Além disso, nossa principal hipótese é a de que os condicionadores que mais atuam sobre o uso de *a gente* é o *tipo de referente* (-determinado) e o *preenchimento do sujeito*.

Este artigo está organizado, além desta introdução, em mais cinco seções: na seção 2, discutimos a gramaticalização de *a gente* no PB; na seção 3, trazemos os resultados de trabalhos a respeito do uso *nós* e *a gente* na escrita do PB, com o objetivo de contextualizar

o Português escrito do Amazonas no Brasil; na seção 4, apresentamos os procedimentos metodológicos; na seção 5, descrevemos a análise do uso de *nós* e *a gente* em cartas pessoais da Amostra Família Arthur Reis. Por fim, na última seção, fazemos as considerações finais.

2 Gramaticalização de *a gente* no PB

Segundo Lopes (2004), a partir do século XIX já se começa a observar a inserção da forma *a gente* como uma das variantes que concorrem à expressão de P4. Considerando o tempo de longa duração, do século XIII ao XX, a referida pesquisadora discute o processo de gramaticalização dessa forma pronominal desde sua origem nominal (*gente* > *a gente*), conforme observamos a seguir.

Lopes (2004), ao discutir a conceituação de gramaticalização, ancorando-se nas reflexões de Lehmann (1982), Hopper (1991) e Hopper e Traugott (1993), assume, como características desse processo, não apenas a assimilação de traços da categoria nova, mas, também, a manutenção de alguns traços da forma anterior. É o que ocorre com o pronome *a gente*, que, mesmo admitindo características de pronome, conserva propriedades do nome *gente*. Para explicar esse processo, a autora adota os traços morfossemânticos de *gênero*, *número* e *pessoa*, a partir da abordagem proposta por Rooryck (1994).

Quanto ao *gênero*, formalmente, o nome *gente* tinha o traço [+fem] e, semanticamente, não havia valor de gênero (tinha um significado neutro). Ao passar para pronome, há uma perda de traço formal (não há gênero formal como os outros pronomes legítimos) e há um ganho de subespecificação semântico-discursiva (passa a admitir uma referência tanto a homens quanto a mulheres, a depender do adjetivo com o qual concorda). No Quadro 1, ilustramos essa alteração nos traços de *gênero*:

Quadro 1 – Exemplos da alteração dos traços de gênero no processo de gramaticalização de gente > a gente

Categoria	**Exemplos de Lopes (2004)**	**Traço formal**	**Traço semântico**
Nome	Todas ***as gentes*** da aldeia ficaram arrasadas.	[+fem]	Neutro
Pronome	Toda ***a gente*** da aldeia ficou *arrasada/arrasado*.	Neutro	[+/-Fem]

Fonte: autoria própria com base em Lopes (2004)

No que diz respeito ao *número*, formalmente, o nome *gente* permitia a marcação positiva ou negativa e, semanticamente, se referia ao plural. Como pronome *a gente*, houve

uma perda da subespecificação formal, passando a ser usado somente no singular [-pl] e mantendo o traço semântico de pluralidade [+pl]. No Quadro 2, exemplificamos essa alteração.

Quadro 2 – Exemplos da alteração dos traços de número no processo de gramaticalização de gente > a gente

Categoria	Exemplos de Lopes (2004)	Traço formal	Traço semântico
Nome	Quen viu o mundo qual o eu já vi, e viu ***as gentes*** que eran enton…" (Séc. XIII, CA, Vasconcelos 1990)	[+/-pl]	[+pl]
Pronome	"e aque ***a gente*** vem a doo de Rachel" (Séc. XIII, CSM, Mettmann, 1972)	[-pl]	[+pl]

Fonte: autoria própria com base em Lopes (2004)

No que tange ao traço de *pessoa*, o nome *gente* tanto no aspecto formal quanto no semântico tem um valor neutro (P3/não-pessoa). Ao mudar para a categoria pronome, *a gente* mantém o valor formal neutro (P3/não-pessoa), porém ganha, semanticamente, o traço de primeira pessoa [+eu] (o eu ampliado). No Quadro 3, mostramos esse processo de alteração no referido traço.

Quadro 3 – Exemplos da alteração dos traços de pessoa no processo de gramaticalização de "gente" > "a gente"

Categoria	Exemplos de Lopes (2004)	Traço formal	Traço semântico
Nome	a **gente** da terra perdem **suas** casas. (Séc. XV, Livro da Cartuxa, Dias 1982).	Neutro (P3/não-pessoa)	Neutro (P3/não pessoa)
Pronome	**a gente** vai mudar as **nossas** coisas para o terreno. (Séc. XX, Mendes 1981:118).	Neutro (P3/não--pessoa)	[+eu] (P4/eu ampliado)

Fonte: autoria própria com base em Lopes (2004)

Vale ressaltar que, para identificar a referência à *pessoa*, Lopes (2004) adota critérios formais, a saber o processo de co-indexação pronominal (na fala culta), além da concordância com a P4 (na fala não-padrão). Para o nome *gente* na fala culta, a co-indexação com as formas *seu/dela* indicam a P3 e para o pronome *a gente* a co-indexação com *nosso(s)/nossa(s)/da gente* para identificar a P4.

Lopes (2004), ao analisar diferentes estágios da língua, observa que *gente* aparece atuando somente como nome entre os séculos XIII e XVIII. Ela identifica que é a partir

do século XIX que há o início da pronominalização dessa forma. Nesse processo de implementação de *a gente* como pronome, como observamos, há manutenção de traços da categoria *nome* (no *número*, formalmente, a concordância com o verbo na P3 e, semanticamente, o caráter indeterminador e coletivo) e há a adoção de propriedades de pronomes legítimos (no *gênero*, passa a se correlacionar a adjetivos no masculino ou feminino em estruturas predicativas e perda da característica de determinação).

Na seção, a seguir, observamos, com mais detalhes, a implementação da forma *a gente* como pronome em trabalhos de dados escritos do PB, sincrônicos e diacrônicos.

3 Estudos sobre as formas *nós* e *a gente* no PB

No Brasil, inúmeros trabalhos têm sido realizados, principalmente em relação à fala, a fim de compreender quais variáveis condicionam o uso das formas de P4 na função de sujeito. Nesta seção, apresentamos algumas pesquisas sobre o referido fenômeno na escrita do PB, separadas por regiões. No Quadro 4, observamos as pesquisas sobre *nós* e *a gente* no Nordeste.

Quadro 4 – Variação nós e a gente em textos escritos na região Nordeste do Brasil

Região Nordeste				
Pesquisas	**Vitório (2015)**	**Santos (2015)**	**Freire (2019)**	**Araújo (2023)**
Localidade	Maceió - Alagoas	Zona rural de Itaibó, distrito de Jequié - Bahia	Montanhas e Pedro Velho - Rio Grande do Norte	Fortaleza - Ceará
Período	Década de 2000 Ano de 2007 (século XXI)	Década de 2000 Ano de 2014 (século XXI)	Década de 2000 Ano de 2017 (século XXI)	Década de 2000 Ano de 2022 (século XXI)
Corpus	120 textos do gênero relato pessoal escritos por alunos do EF[2] e EM[3] em uma escola da rede pública.	28 textos do gênero relato pessoal escritos por alunos do EF II (6º, 7º, 8º e 9º anos) de uma escola pública na zona rural.	155 textos do gênero relato pessoal escritos por alunos do EF II de duas escolas públicas (estadual e municipal) localizadas na zona urbana.	360 redações escritas por alunos do EM de uma escola pública de ensino integral.

[2] Para EF, leia-se: Ensino Fundamental.
[3] Para EM, leia-se: Ensino Médio.

Resultados Gerais	Das 242 realizações, **208 dados (86%) são da variante *nós*** e 34 (14%) são da variante *a gente* na função sujeito.	De 106 ocorrências, **78 referem-se à variante *nós* (73,58%)** e 28 à variante *a gente* (26,42) na função de sujeito.	De 682 ocorrências, **591 referem-se à variante *nós* (86,80%)** e à variante *a gente* (13,30%) na função sujeito.	De 1265 ocorrências, **1131 referem-se à variante *nós* (98,4%)** e 134 à variante *a gente* (10,6%) na função sujeito.
Variáveis Linguísticas	**Paralelismo formal**: *A gente* é favorecido nas seguintes situações: uso isolado, primeiro da série e quando antecedido por ele mesmo. **Marca morfêmica do verbo:** *A gente* ocorreu com a desinência {-mos} e com verbos no singular. **Preenchimento do sujeito:** *A gente* é mais frequente quando há a realização do sujeito pronominal.	**Preenchimento do sujeito:** *A gente* teve poucas ocorrências com sujeito preenchido. **Grafia do pronome** *a gente* (25%) a*gente* (71,4%) *ajente* (3,6)	**Concordância Verbal:** *A gente* é mais frequente com verbos no singular.	**Grau de referência do pronome:** *A gente* é favorecido quando seu uso é genérico. **Paralelismo Discursivo:** *A gente* é favorecido quando antecedido por ele mesmo. **Preenchimento do Sujeito:** *A gente* é favorecido quando o sujeito é lexicalizado.
Variáveis Externas	***Escolaridade:*** O EF tende a usar mais a forma *a gente.* **Sexo/gênero:** As meninas utilizam mais a variante *a gente* **Tema da produção textual:** *A gente* foi mais frequente em gêneros pessoais narrativos.	**Escolaridade:** *A gente* é mais utilizado entre os alunos dos 6º, 7º e 8º anos do EF, diminuindo à proporção que a escolarização aumenta.	---	---

Fonte: autoria própria

Como podemos observar no Quadro 4, os estudos da região Nordeste já registram o uso da variante *a gente* em textos escritos por alunos tanto do EF quanto do EM (em sua maioria, no gênero relato pessoal), embora com uma frequência não muito alta (em torno de 26%, no máximo). Ainda, nos trabalhos de Vitório (2015) e Araújo (2023), verificamos que um dos condicionadores linguísticos que atuam sobre referida variante é o *preenchimento do sujeito*: *a gente* é favorecido quando o sujeito é preenchido. As pesquisas de Vitório (2015) e Freire (2019) atestam o favorecimento considerável da forma inovadora *a gente* com verbos no singular, embora exista a ocorrência da variante com a desinência {*-mos*}. Em relação ao *paralelismo*, Vitório (2015) e Araújo (2023) mostram que, quando a variante *a gente* é usada na primeira posição, tende a se repetir nas sentenças posteriores. Quanto ao *tipo de referência*, apenas Araújo (2023) registrou o favorecimento de *a gente* com uso genérico. Por fim, os resultados de Santos (2015) revelaram também maior uso da forma inovadora quando o sujeito é lexicalizado e expresso pelas grafias *a gente/agente/ajente.*

Finalmente, quanto aos condicionadores extralinguísticos, apenas Vitório (2015) e Santos (2015) fizeram esse controle. Os resultados desses trabalhos mostraram que a variante *a gente* é favorecida pelo sexo/gênero feminino e pelos menos escolarizados. No Quadro 5, observamos os resultados de pesquisas na região Centro-Oeste.

Quadro 5 – Variação nós e a gente em textos escritos na região Centro-Oeste do Brasil

Região Centro-Oeste	
Autor e Ano	Ribeiro e Vieira (2019)
Localidade	Cidade de Goiás - Goiás
Período	Décadas de 1830/37/78/80/85/86/90 (século XIX) e Décadas de 1907/27/32/34/40/55 (século XX)
Corpus	Arquivos da Fundação Cultural Frei Simão Dorvi. Museu dos Bandeirantes, jornais e periódicos antigos, arquivos da cidade de Goiás e Hemeroteca Digital.
Resultados Gerais	O uso da variante *nós* é frequente nos gêneros mais formais, já a variante *a gente* como forma pronominal alternativa de *nós* deu-se pelos anos 1800, a partir do surgimento dos jornais na cidade em textos poéticos, crônicas relatos e cartas de leitores.
Variáveis Linguísticas	**Posição de sujeito** A variante *a gente* foi favorecida na posição de sujeito nos textos antigos desde o início da imprensa em 1830. **Referência genérica** A variante *a gente* foi favorecida com referente genérico com quatro ocorrências nos *corpora.*
Variáveis Externas	Não controlados.

Fonte: autoria própria

Na região Centro-Oeste, os resultados qualitativos do estudo de Ribeiro e Vieira (2019) indicam que a variante *nós* é mais comum em contextos formais, como era de se esperar; já a variante inovadora *a gente* apareceu discretamente em textos poéticos, crônicas e cartas de leitores a partir dos anos 1830. Das poucas ocorrências encontradas, destaca-se favorecimento na *posição de sujeito* e com *referência genérica*, com registro de quatro ocorrências. No Quadro 6, elucidamos alguns resultados de pesquisas na região Sudeste.

Quadro 6 – Variação nós e a gente em textos escritos na região Sudeste do Brasil

Região Sudeste		
Autor e Ano	**Vianna (2006)**	**Souza (2015)**
Localidade	Rio de Janeiro - Rio de Janeiro	Rio de Janeiro - Rio de Janeiro
Período	Década de 1980 e década de 2000 (Séculos XX e XXI).	Década de 1930 - (Século XX).
Corpus	**Oralidade:** 21 inquéritos das amostras do projeto Censo/Peul-RJ, década de 1980. **Escrita:** 39 textos de estudantes de 1º e 2º graus, nos anos 2000, escritos por homens e mulheres com faixa etária entre 15-25, 26-49 e + de 50 anos. **Amostra complementar:** 104 testes de avaliação subjetiva, aplicados no ano de 2005 em duas escolas estaduais a homens e mulheres com faixa etária entre 15-25, 26-49, mais de 50 anos no EF e no EM.	97 cartas pessoais.
Resultados Gerais	A variante *a gente* apareceu nas três décadas analisadas, a saber de 1980 a 2000.	Obteve-se 159 dados de P4. Houve um predomínio da variante *nós* (152 dados, 95,5%). A variante *a gente*, apesar de preterida, já se encontra presente na escrita do PB na década de 1930 (século XX).
Variáveis Linguísticas	**Concordância Verbal:** (*a gente*+P3): 79% (PR. 0,92) (*a gente*+P4): 9% (PR. 0,22) **Concordância de gênero e número:** Singular: 54% (PR. 0,71) Plural: 12% (PR.0,30)	**Função sintática:** Só houve realização de *a gente* no caso nominativo; **Traço [+ determinado] do referente:** Em relação à utilização de *a gente*, não houve caso de referência indeterminada ou sujeito ampliado. Tanto *a gente*
Variáveis Linguísticas (continuação)	**Tempo verbal:** A variante *a gente* foi favorecida no presente do indicativo (38%) (PR. 0.54)	quanto *nós* só apareceram com o traço [+determinado]. **Sujeito preenchido e nulo:** A forma *a gente* sempre se apresenta explicitamente.

Variáveis Externas	**Escolaridade:** A variante *a gente* foi favorecida por alunos do EF (41%) (PR.0.34) **Sexo/Gênero:** As pessoas do sexo/gênero feminino favorecem o uso da variante *a gente.*	**Sexo/ Gênero e Escolaridade:** Utilização categórica da forma *nós* pelos informantes homens (mais escolarizados); Utilização variável entre as formas *nós* e *a gente* pelas informantes mulheres (menos escolarizadas).

Conforme o Quadro 6, em relação à região Sudeste, os resultados dos estudos de Vianna (2006) e Souza (2015) mostram que, apesar de a variante nós ser predominante nas duas amostras, *a gente* já estava presente na escrita do PB desde a década de 1930. Quanto aos condicionadores linguísticos, a variante em questão, no RJ, é favorecida em *concordância verbal* com a P3 (*a gente* + P3, Vianna, 2006), no *presente do indicativo* (Vianna, 2006) e quando o *sujeito é preenchido* (Souza, 2015). No que tange às variáveis extralinguísticas, a forma inovadora *a gente* é mais utilizada pelas mulheres e por pessoas com menor escolaridade. No Quadro 7, observamos os resultados de pesquisas na região Sul.

Quadro 7 – Variação *nós* e *a gente* em textos escritos na região Sul do Brasil

Região Sul				
Pesquisas	**Brustolin (2009)**	**Agostinho (2013)**	**Borges (2019)**	**Monguilhott *et al.* (2021)**
Localidade	Florianópolis - Santa Catarina	Itajaí - Santa Catarina	Rio Grande do Sul	Santa Catarina
Período	Década de 2000 - De maio a outubro de 2008 (século XXI)	Década de 2010 - De março a abril de 2011 (século XXI)	Décadas de 1890 e 1900 (século XIX); Décadas de 1910/20/30/40/ 50/60/70/80/90 (século XX)	Décadas de 1880 e 1890 (século XIX); Décadas de 1960, 1970, 1980 e 1990 (século XX)
Corpus	393 textos narrativos escritos por alunos do EF (5º, 6º, 7º e 8º) em quatro escolas da rede pública	334 Redações (narrativas) elaboradas por alunos dos anos finais do EF (5ª, 6ª, 7ª e 8ª séries) de duas escolas públicas.	11 peças de teatro	Amostras do PHPB-SC: 21 cartas pessoais do século XIX e 136 cartas do século XX
Resultados Gerais	Obtiveram-se 1.284 dados de P4. Houve uma realização de **86% da variante *nós*** e de 14% de *a gente*	Obtiveram-se 2.199 dados de P4, **sendo 1.204 de *nós* (55%)**, 787 com SN+eu (36%) e 208 com *a gente* (9%)	Obtiveram-se 246 dados de P4 expressos formalmente e em função de sujeito, assim distribuídos: 122 de *nós* (49,6%) e **124 de *a gente* (50,4%)** A variante *a gente* já é encontrada na década de 1890 (séc. XIX), com um percentual de 33%	Obtiveram-se 142 dados de P4. No séc. XIX, **todos os casos (13 dados) foram da variante *nós*.** No século XX, houve **129 dados de *nós*** e oito de *a gente* A variante *a gente* é encontrada somente no século XX

Variáveis linguísticas	**Marca morfêmica do verbo:** Há maior produtividade de *a gente* com morfema zero (P3) **Sujeito preenchido e nulo:** Há predominância de *a gente* no que se refere ao sujeito preenchido e não ao sujeito nulo **Saliência fônica:** O pronome *a gente* é empregado com mais frequência nos níveis de menor saliência fônica (34%) do que nos níveis de maior saliência fônica (24%) **Paralelismo formal:** quando o pronome *a gente* inicia uma série, uma forte tendência é que essa escolha influencie no uso de *a gente* + *a gente* nas formas subsequentes	A variável investigada é a variação da concordância verbal de P4	---	----
Variáveis Externas	**Sexo/Gênero:** As mulheres utilizam mais o pronome *a gente* do que os homens (29% e 20%, respectivamente) **Série escolar:** Os alunos que mais utilizam o pronome *a gente* são das séries: 5ª e 6ª (séries iniciais do EF)	A variável investigada é a variação da concordância verbal de P4	**Sexo/Gênero:** A forma inovadora *a gente* foi favorecida pelas personagens femininas, tanto em percentual (56%) como em peso relativo (0,56) **Faixa etária:** As personagens da faixa etária intermediária (de 26-49 anos) foram os que mais utilizaram *a gente* (53% / 0,54); **Classe social:** O uso de *a gente* é favorecido pelos personagens da classe baixa, com percentual de 54% (0,56)	----

Fonte: autoria própria

Como pode ser verificado no Quadro 7, na região Sul, há pesquisas com *corpus* bastante diversificado, formado por textos narrativos (redações), peças de teatro e cartas pessoais, representando a escrita catarinense desde o século XIX até o século XXI. Como resultados gerais, há o predomínio da variante *nós* em quase todas as pesquisas, com exceção de Borges (2019), em que a variante *a gente*, ainda que com uma diferença não muito grande, se sobressai nas 11 peças de teatro escritas nos séculos XIX e XX. Em relação aos condicionadores linguísticos, a pesquisa de Brustolin (2009) mostra a importância da *marca morfêmica do verbo*, do *preenchimento do sujeito*, da *saliência fônica* e do *paralelismo formal* para o uso da variante *a gente*. Sociolinguisticamente, como é constatado em outras pesquisas, as mulheres e os menos escolarizados favorecem o uso da variante inovadora. Por fim, apesar da predominância da forma *nós*, a variante *a gente* já é constatada na escrita brasileira na década de 1890.

Conforme observado nesta seção, as pesquisas descritas levaram em conta, principalmente, os seguintes condicionadores linguísticos: *sujeito preenchido e nulo*, *tipo de referência*, *marca morfêmica do verbo*, *paralelismo formal*, *concordância verbal* e *tempo verbal*. Os resultados mostram que a forma *a gente*, geralmente, é produtiva em contextos em que o sujeito é preenchido, concordando com a P3 e quando é o primeiro de uma série. Além disso, os principais condicionadores extralinguísticos considerados foram: *escolaridade*, *sexo/gênero* e *faixa etária*. De maneira geral, podemos observar que a variante inovadora se faz presente na maioria dos gêneros textuais escritos, com implementação a partir do século XIX, especialmente na escrita de mulheres e de pessoas menos escolarizadas. Esse panorama indica um avanço da variante, mesmo com a pressão que a escola exerce sobre a escrita.

Salientamos que até o momento de elaboração deste artigo, não foram encontrados estudos que investiguem a variação de P4 na escrita na região Norte do país. Essa ausência de pesquisas destaca a importância de realizarmos estudos sobre esse fenômeno na modalidade escrita no Amazonas.

A seguir, detalhamos a amostra da escrita amazonense aqui utilizada.

4 Procedimentos metodológicos

Neste artigo, foram analisados os dados extraídos das cartas pessoais da família de Arthur Reis, datadas das décadas de 1940 a 1980. Esse acervo pode ser encontrado em sua forma física na Biblioteca Arthur Reis, localizada no Centro Cultural Povos da Amazônia,

em Manaus (AM) e compõe o *corpus* mínimo comum[4] do projeto *Para a História do Português do Amazonas* – PHPB/AM. As cartas desse acervo foram digitalizadas e transcritas pela referida equipe e estão disponíveis no *drive*[5] do projeto.

Dessa amostra, foram selecionadas 127 missivas que já haviam sido transcritas pela equipe do PHPB/AM para o estudo sobre as formas de tratamento, realizado por Maia, Martins e Coelho (2024). O referido estudo adotou, como principal critério para seleção, missivas trocadas entre membros da família e amigos[6]. Em relação aos destinatários, a maioria dessas cartas são enviadas à Graziela Reis (esposa de Arthur Reis) por seus filhos, irmãos, sobrinhos, netos e amigos, além de poucas cartas destinadas ao casal ou aos filhos de Arthur e Graziela, conforme se observa no Quadro 8.

Quadro 8 – Tipos de relações na Amostra da Família Arthur Reis

Tipos de relação			**Destinatário**		
Relações assimétricas descendentes	Pai		Graziela Reis	4	
		Emília	Graziela Reis	8	
Relações assimétricas ascendentes		José Augusto	Graziela Reis	9	
		Iria	Graziela Reis		
		Vanja	Graziela Reis	5	
		Miriam	Graziela Reis	2	
		Ruy	Graziela Reis	1	
		Mimar/ Mimá	Graziela Reis	1	
		Márcia Reis	Arthur e Graziela Reis	1	
		Maria da Graça	Graziela Reis	4	

[4] O *Corpus* Mínimo Comum é composto por textos impressos, como cartas de leitores, editoriais/cartas de redatores e anúncios, e por textos manuscritos, incluindo cartas particulares e cartas de leitores (Castilho, 2021, n.p.).

[5] Em breve, esse material, juntamente com outros pertencentes ao projeto, estarão disponíveis ao público também em um *site.*

[6] No conjunto de cartas referentes à família Arthur Reis, há correspondências de diferentes remetentes, algumas das quais não permitem identificar o tipo de relação. Além disso, o material inclui cartas sem data. Esses casos foram excluídos do estudo.

Relações Simétricas		Graziela Reis	Filhos de Arthur e Graziela Reis	2	
			Graziela Reis	1	
		Alésia	Graziela Reis	1	
		Heloísa	Graziela Reis	8	
		Edgar	Graziela Reis	7	
		Isaura	Graziela Reis	5	
		Judith	Graziela Reis	1	
		Fernando	Graziela Reis	1	
Relações Simétricas		Esther	Graziela Reis	1	
		Carlota	Arthur e Graziela Reis	1	
		Quésia	Graziela Reis	1	
TOTAL	21	3	83	--	

Das 127 cartas analisadas, salientamos que em 44 não havia dados de P4, o que resultou em 83 cartas analisadas de 21 missivistas. Vale ressaltar que os dados de P4 em outras funções sintáticas (genitivo, acusativo e oblíquo) não foram considerados para a análise quantitativa desta pesquisa.

4.1 Biografia de Arthur Reis

Arthur Cezar Ferreira Reis foi uma figura de grande relevância, tanto na política quanto na literatura da região amazônica. A amostra aqui analisada integra seu acervo bibliográfico, motivo pelo qual destacamos sua biografia, dada sua significativa influência política no estado e a vasta informação disponível. Em contraponto, não encontramos muitas referências sobre outros membros da família, além do que está nas cartas sobre suas relações (cf. Quadro 8). Sabemos que a família deixou Manaus há muito tempo. Além disso, recentemente, a equipe do PHPB/AM conseguiu contato com uma das bisnetas de Arthur Reis e Graziela Reis, cuja avó é a Iria (a filha mais velha). Ela informou que uma das filhas mais novas do casal, a Miriam, ainda está viva. Por termos informações mais vastas de Arthur Reis, apresentaremos, a seguir, uma breve descrição de sua vida, incluindo informações acadêmicas, atuação profissional, entre outras.

De acordo com o *site* do Instituto Histórico e Geográfico Brasileiro (IHGB) e Sousa (2009), Arthur Reis nasceu em Manaus (AM) em 08 de janeiro de 1906. Ele é filho de Vicente Torres Reis e Emília Ferreira Reis. É importante destacar que seu pai era jornalista[7] e escritor teatral.

Em relação à sua formação inicial, Arthur Reis fez o curso primário e secundário

[7] Vicente Torres Reis foi proprietário de um dos jornais mais antigos e que ainda está em funcionamento no Amazonas, o Jornal do Commercio.

nos mais prestigiados grupos escolares de Manaus, a saber, escola Saldanha Marinho, Marechal Hermes e o Ginásio Amazonense D. Pedro II. Graduou-se pela Faculdade de Direito da Universidade do Rio de Janeiro (hoje UFRJ), em 1927.

Acerca de sua atuação profissional, elencamos suas diversas atribuições, entre elas, seu início no magistério em 1928, em Manaus (AM), como professor de História do Brasil no colégio Dom Bosco, História da Civilização e História do Brasil na escola Solon de Lucena e Economia Política e Ciências das Finanças na Faculdade de Direito do Amazonas.

Mais tarde, Arthur Reis foi vice-diretor da Associação dos Empregados do Comércio do Amazonas, em 1929, e chefe de gabinete da Junta Governativa Revolucionária do Amazonas, em 1930. Ele também atuou como diretor da instrução pública do Estado do Amazonas, Fiscal Federal de Seguros Marítimos e Fiscalização no mesmo estado. Além disso, foi diretor do Departamento de Educação e Cultura do Amazonas e representante da Diretoria do Patrimônio Histórico e Artístico Nacional no Pará, em 1940. Entre 1946 e 1953, chefiou a Divisão do Ministério do Trabalho.

Entre 1953 e 1955, Arthur Reis foi superintendente do Plano de Valorização Econômica da Amazônia (SPVEA) e, de 1950 a 1958, diretor do Instituto Nacional de Pesquisas da Amazônia (INPA). Em 1961, tornou-se diretor do Departamento de História e Documentação do estado da Guanabara e, entre 1961 e 1964, diretor do Departamento Nacional de Indústria do Ministério da Indústria e do Comércio. De 1964 a 1967, ocupou o cargo de governador do Estado do Amazonas. Além disso, foi presidente e membro do Conselho Federal de Cultura e membro da Comissão de Textos de História do Brasil do Ministério das Relações Exteriores.

Arthur Reis também foi sócio grande-benemérito do IHGB, jornalista colaborador em várias revistas e redator-chefe do Jornal do Comércio. Como escritor, publicou mais de 2000 artigos, inúmeros prefácios e livros[8]. Arthur Reis, apesar de pouca idade, era uma promessa na implementação de novas ideias. Portador de um conhecimento de altíssimo nível, revelou-se um homem de virtudes e negócios, destacando-se na sociedade manauense da época. Suas atribuições revelam forte atuação nos locais por onde residiu, desde Manaus, Belém e Rio de Janeiro. Na amostra de cartas aqui analisada, observa-se a troca de missivas entre familiares e amigos que residiam, por vezes, separados em uma dessas cidades. Arthur Reis faleceu em 7 de fevereiro de 1993, na cidade do Rio de Janeiro.

[8] Em sua maioria, sobre História e Geografia da região amazônica.

4.2 O envelope de variação

Para a realização deste estudo, a respeito da variação de *nós* e a *gente* na função de sujeito, na amostra de cartas pessoais da Família Arthur Reis, primeiramente, selecionamos 127 cartas que foram transcritas e editadas pelos membros do projeto PHPB/AM. Logo após essa etapa, separamos todas as sentenças com os pronomes, expressos e nulos, *nós* e *a gente*. A seguir, exemplificamos esses usos com dados destacados das referidas cartas.

(1) ***Nós**, aqui de longe [rasura], | **ficaremos** cheios de saudades, na | revivescencia dos dias felizes de quan - | do estavamos juntos.* [Trecho da carta de Raymundo a sua filha Graziela Reis, em 07/05/1940]

(2) *Na excursão, Paris moderno, **fomos** ao Bosque de Bolanha.* [Trecho da carta de Graziela Reis a seus filhos, em 14/10/1953]

(3) *Disse-me que quando **a gente tem** um quer um, quando **tem** 5 quer [os 5]* [Trecho da carta de Iria a sua mãe Graziela Reis, em 13/12/1954].

Para análise da nossa variável dependente "a expressão de P4 na função de sujeito", como pode se constatar no Quadro 9, controlamos os condicionadores extralinguísticos e linguísticos que se mostraram relevantes nas pesquisas elucidadas na seção 3.

Quadro 9 – Envelope de Variação

Variáveis Linguísticas		
	Fatores	**Hipóteses**
Preenchimento do sujeito pronominal	Nulo	Espera-se que o fator sujeito expresso favoreça o uso de *a gente*, conforme foi encontrado nas pesquisas de Brustolin (2009), Vitório (2015) e Araújo (2023).
	Expresso	
Paralelismo entre as formas de sujeito pronominal	Formas paralelas (formas de P4 iguais)	Espera-se que o fator formas paralelas (formas iguais) favoreça o uso de *a gente*, conforme foi encontrado nas pesquisas de Brustolin (2009), Souza (2015), Vitório (2015) e Araújo (2023).
	Formas não paralelas (formas de P4 diferentes)	

Conjugação Verbal	1ª Conjugação	Os trabalhos mencionados na seção 3 não apresentam discussão para essa variável, mas o trabalho de Araújo (2016), sobre dados de fala, mostra que os verbos de 2ª conjugação (ter, ser) desfavorecem o uso de *a gente*.
	2ª Conjugação	
	3ª Conjugação	
Tempo Verbal	Presente	Espera-se que o fator presente favoreça o uso de *a gente*, conforme foi encontrado na pesquisa de Vianna (2006).
	Pretérito	
	Futuro	
Saliência Fônica (considerando o radical do verbo)	[+marcado]: irregularidade do radical	Espera-se que o fator [-marcado] favoreça o uso de *a gente*, conforme foi encontrado na pesquisa de Brustolin (2009).
	[-marcado]: regularidade do radical	
Tipo de referente (eu ampliado)	eu + você [+determinado]	Espera-se que o fator eu + todos [-determinado] favoreça o uso de *a gente*, conforme foi encontrado nas pesquisas de Ribeiro e Vieira (2019) e Araújo (2023).
	eu + ele [+determinado]	
	eu + eles [+determinado]	
	eu + todos [-determinado]	
Variáveis Extralinguísticas		
Missivista	21 missivistas (conforme Quadro 8)	Espera-se que as missivistas do sexo/gênero feminino favoreçam o uso de *a gente*, conforme foi encontrado nas pesquisas de Vianna (2006), Brustolin, (2009), Vitório (2015), Souza (2015) e Borges (2019).
Década	1940	Espera-se que os fatores 1960 e 1980 favoreçam o uso de *a gente*, conforme foi encontrado na pesquisa de Lopes (2004) e Borges (2019).
	1950	
	1960	
	1980	

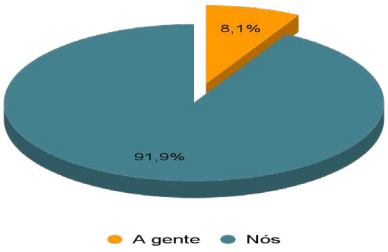
As ocorrências de *nós* e *a gente* na função de sujeito foram categorizadas em planilhas do Excel, conforme as variáveis linguísticas e extralinguísticas controladas e, em seguida, submetidas ao Programa estatístico GoldVarbX[9]. De maneira geral, nas 83 cartas analisadas, encontramos 283 ocorrências das formas *nós* e *a gente* na função de sujeito. Os resultados dessa análise são apresentados na próxima seção.

[9] GoldVarbX (David Sankoff, Sali A. Tagliamonte e Eric Smith, 2005) é uma versão para Windows do pacote estatístico Varbrul. Foi elaborado especialmente para análise multivariada de dados de variação sociolinguística (Guy; Zilles, 2007).

5 Os Resultados

A análise estatística realizada, das 283 ocorrências de *nós* e a *gente* na função de sujeito encontradas nas cartas pessoais da família Arthur Reis, mostrou que 23 foram da forma *a gente* (8,1%) e 260 da forma *nós* (91,9%), conforme podemos visualizar no Gráfico 1.

Gráfico 1 – As formas nós e a gente na Amostra Família Arthur Reis, 1940 a 1980



Fonte: autoria própria

Como observamos no Gráfico 1, a forma *nós*, de maneira geral, predomina na amostra Arthur Reis, que abrange cartas pessoais trocadas entre familiares nas décadas de 1940 a 1980. Esse predomínio também é observado em outras amostras sincrônicas e diacrônicas mencionadas na seção 3, como Araújo (2023), Ribeiro e Vieira (2019), Vianna (2006), Souza (2015), Brustolin (2009), Agostinho (2013) e Monguilhott *et al.* (2021). Por outro lado, a forma *a gente* na amostra de Borges (2019), referente aos séculos XIX e XX, na região sul do Brasil, em 11 peças de teatro, apresenta um resultado contrário às pesquisas citadas anteriormente. Das 246 ocorrências de sujeitos expressos, 124 (50,4%) correspondem à variante *a gente*, enquanto 122 (49,6%) à *nós*.

A fim de compreendermos o uso da forma inovadora *a gente*, considerada como aplicação da regra, na escrita amazonense aqui analisada, controlamos os grupos de fatores linguísticos e extralinguísticos que estão elencados na seção 4.2. Em virtude de *Knockouts* e sobreposições, excluímos algumas variáveis (*paralelismo*[10] e *missivistas*[11]) e fatores (*tempo futuro*[12]), assim como realizamos amálgamas (*tempo verbal*: presente e passado e *década*: 1940, 1950/1960 e 1980). Sendo assim, para a análise de Peso Relativo, controlamos os seguintes grupos de fatores: *preenchimento do sujeito*, *conjugação verbal*,

[10] A forma *a gente,* em sua maioria, aparece na função de sujeito. Há apenas dois casos em que aparece como possessivo.
[11] Essa variável se sobrepôs à *década*.
[12] Não apareceu nenhum *a gente* com verbo no tempo no futuro.

tempo verbal, saliência fônica, tipo de referência e *década*. Desses, o programa estatístico selecionou os seguintes, em ordem de seleção: *preenchimento do sujeito, tipo de referência, década* e *tempo verbal*.

A seguir, na Tabela 1, visualizamos os resultados encontrados referentes ao condicionador *preenchimento do sujeito*:

Tabela 1 – Frequência e probabilidade da variante "a gente", segundo a variável 'preenchimento do sujeito'

Fatores	**Aplicação/ Total**	**%**	**P.R**
Expresso	19/50	38%	0,96
Nulo	4/233	1,7%	0,33
Significância: 0,012			
Input: 0,014			

Fonte: autoria própria

Os resultados mostram que a forma *a gente* é favorecida quando o sujeito é expresso (0,96) e desfavorecida quando o sujeito é nulo (0,33), corroborando com os resultados encontrados por Vitório (2015) e Araújo (2023), no nordeste do Brasil, assim como em Souza (2015), no Sudeste, e Brustolin (2009), no Sul.

A forma *a gente* quando passa à categoria de pronome no Português, conforme discutido na seção 2, mantém o traço formal de P3, embora, semanticamente, seja interpretada como P4. Como as formas verbais de P3 não apresentam marca morfêmica, a ausência do pronome na frase poderia gerar uma ambiguidade em relação ao sujeito. Por esse motivo, provavelmente, a forma *a gente* é preenchida.

Os quatro casos de *a gente* nulos que apareceram, na amostra aqui analisada, correspondem a situações de paralelismo entre sujeitos na mesma sentença, conforme os exemplos 1 ao 3, a seguir:

(1) *Até nos armários de | roupa, quando* ***a gente*** *abria, Ø via pulga passean- | do.* [Trecho da carta de Iria para sua mãe Graziela Reis, 29/09/1954].

(2) ***A gente*** *| faz do boneco o que Ø quer.* [Trecho da carta de Iria para sua mãe Graziela Reis, 19/08/1957).

(3) *Resolvi escrever porque os telefonemas são | tão rápidos que* ***a gente*** *quasi* não pode *con- | versar, dizer tudo o que Ø queria dizer.* [Trecho da carta de Iria para sua mãe Graziela Reis, 20/02/1989)].

No que se refere à variável *paralelismo entre formas de sujeito,* na Tabela 2, ilustramos os resultados:

Tabela 2 – As formas nós e a gente na função de sujeito, segundo a variável linguística 'paralelismo'

Paralelismo de formas do sujeito pronominal	**A expressão de P4**	
	A gente	***Nós***
Formas paralelas (sujeitos iguais)	0/165 = 0%	165/165 = 100%
Formas não paralelas (sujeitos diferentes)	23/93 = 24,7%	70/93 = 75,3%
TOTAL	**23/258 = 8,1%**	**235/258 = 91,1%**

Fonte: autoria própria

Os resultados expostos na Tabela 2 indicam que, nas cartas de um mesmo missivista em que havia dados de *a gente*, a forma *nós* também era utilizada, mostrando a variação existente entre esses dois pronomes para expressão de P4. No exemplo 4, apresentamos esse caso de *não paralelismo entre sujeitos* em uma mesma carta:

(4) *Realmente, Ø* ***temos*** *que | auxiliar o próximo, mas, o que não é pos- | sivel é a vida toda* ***a gente*** *estar sempre |* á *disposição dos outros.* [Trecho da carta de Iria para sua mãe Graziela Reis, 19/04/1955].

Por sua vez, conforme mostram os resultados da Tabela 2, o *paralelismo entre sujeitos* sempre ocorre com a forma *nós*, como no exemplo 5 que segue:

(5) [...] não estou certa | de que ***nós vamos****, mas papai falou que talvez* ***nós | possamos*** *ir aí.* [Trecho da carta de Márcia Reis para seus avós Arthur e Graziela Reis, 19/06/1966]

Na Tabela 3, apresentamos os resultados da segunda variável selecionada, o *tipo de referência:*

Tabela 3 – Frequência e probabilidade da variante "a gente", segundo a variável 'tipo de referência'

Grau de indeterminação do sujeito	**Fatores**	**Aplicação/ Total**	**%**	**P.R**
-determinado (grau 3)	Eu + todos	15/69	21,7%	0,83
-determinado (grau 2)	Eu+ ele	2/85	2,4%	0,37
-determinado (grau 1)	Eu + eles	6/121	5,0%	0,36
+ determinado	Eu + você	0/8	0%	—-
Significância: 0,012 *Input*: 0,014				

Fonte: autoria própria

De acordo com a Tabela 3, considerando o seguinte grau de indeterminação (graus 1, 2 e 3), observamos que o fator *eu + todos* (0,83), que é o de maior grau de indeterminação do sujeito, favorece o uso da forma *a gente*, enquanto os demais fatores, *eu + ele* (0,37) e *eu + eles* (0,36), a desfavorecem. Vale ressaltar que no fator *eu + você*, que é o mais determinado, não apareceu nenhum uso de *a gente*, podendo indicar, assim, que essa forma, ao ser gramaticalizada como pronome no Português, ainda mantém o traço de indeterminação da categoria nome, conforme apontado por Lopes (2004). Os trabalhos de Araújo (2023), no Nordeste, e Ribeiro e Viana (2019), no Centro-Oeste, também mostram essa tendência. A pesquisa de Souza (2015), no Sudeste, por outro lado, não encontrou nenhum dado de *a gente* e *nós* com grau de indeterminação.

Nos exemplos 6 a 8, ilustramos *a gente* nesse sentido mais genérico (o eu ampliado):

(6) (2) *Disse-me que quando* ***a gente*** *tem um Ø quer um, quando Ø tem 5 Ø quer [os 5]* [Trecho da carta de Iria para sua mãe Graziela Reis, 13/12/1954].

(7) *Já se foi o | tempo que* ***a gente*** *ia á vontade para Copacabana.* [Trecho da carta de Iria para sua mãe Graziela, 06/10/1954].

(8) *No | final, quando todo o conjunto estava dansan- | do, com uma harmonia de movimentos im- | pressionante, as luzes todas acesas, [rasurado] a rou- | pas com coloridos os mais diferentes,* ***a gente*** *ti- | nha ímpeto de ficar de pé e aplaudir.* [Trecho da carta de Iria para sua mãe Graziela, 13/12/1954].

Na Tabela 4, apresentamos os resultados a respeito da variável *década,* selecionada em terceiro lugar:

Tabela 4 – Frequência e probabilidade da variante "a gente", segundo a variável 'década'

Fatores	Aplicação/ Total	%	P.R
Década de 1980	3/7	42,9%	0,92
Década de 1950 e 1960	19/203	9,4%	0,67
Década de 1940	1/73	1,4%	0,09

Significância: 0,012

Input: 0,014

Fonte: autoria própria

Conforme a Tabela 4, a forma *a gente* é favorecida pelas décadas de 1980 (0,92) e 1950/1960 (0,67), ao passo que é desfavorecida pela década de 1940 (0,09).

Considerando as particularidades da amostra analisada, na qual só registramos dados de *a gente* em cartas de duas remetentes (Iria e Alésia), esse resultado diacrônico, na verdade, é enviesado pela variável *missivista*. Na Tabela 5, mostramos a quantidade de *a gente* utilizada por cada uma e por década:

Tabela 5 – As formas *nós* e *a gente* na função de sujeito por 'missivista' e 'década', segundo as relações sociais entre os interlocutores

Missivista	**Década**	**Número de cartas**	**A EXPRESSÃO DE P4**	
			A GENTE	**NÓS**
Iria (filha de Graziela e Arthur Reis)	1950	17	19	82
	1960	1	0	3
	1980	1	3	4
Alésia (Irmã de Graziela Reis)	1940	1	1	3
TOTAL		20	23	92

Fonte: autoria própria

Na Tabela 5, observamos que não há regularidade em relação à distribuição do número de cartas por década/missivista. Por exemplo, na década de 1980, há somente uma carta de Iria; nas décadas de 1950/1960, há 18 de Iria e na década de 1940, somente uma de Alésia. Desse modo, não é possível discutirmos a implementação da forma *a gente* na escrita amazonense na perspectiva da mudança em tempo real, pois seriam necessários dados de outros missivistas em cada década.

Por outro lado, a forma *a gente*, como mencionado anteriormente, foi utilizada na nossa amostra por duas mulheres, o que pode indicar que a variável *sexo/gênero* influencia na implementação dessa forma no PB. Vale ressaltar que alguns trabalhos discutidos na seção 3 (Vitório, 2015; Viana, 2006; Brustolin, 2009; Souza, 2015 e Borges, 2019) mostram as mulheres liderando o uso da forma *a gente,* corroborando com os resultados desta pesquisa. Os primeiros estudos labovianos já mostravam essa tendência de que as mulheres ocidentais podem liderar a mudança linguística, mas quando a forma inovadora tem algum prestígio (Labov, 2008 [1972]; Paiva (2015). Ressalta-se que, nas pesquisas de Vitório (2015), Santos (2015), Vianna (2006) e Souza (2015), a forma *a gente* é mais produtiva na escrita dos menos escolarizados, por esse motivo não podemos ainda correlacionar essa forma com prestígio.

No que se refere à nossa amostra, não podemos ter certeza do nível de escolaridade dos missivistas, porém, o que se sabe sobre Iria, por exemplo, conforme informações ex-

postas nas cartas, é que se mudou para o Rio de Janeiro, na década de 1950, casando-se e, posteriormente, cursando nível superior. Quanto à Alésia, não há nenhuma informação sobre escolaridade.

Na Tabela 6, apresentamos os resultados da variável *tempo verbal*, quarta e última variável selecionada.

Tabela 6 – Frequência e probabilidade da variante "a gente", segundo a variável 'tempo verbal'

Fatores	**Aplicação/ Total**	**%**	**P.R**
Presente	18/103	17,5%	0,62
Passado	5/115	4,3%	0,38
Significância: 0,012			
Input: 0,014			

Fonte: autoria própria

Conforme os resultados expostos na Tabela 6, verificamos que a forma *a gente* é favorecida pelo tempo presente (0,62), ao passo que é desfavorecida pelo tempo passado (0,38). Nos exemplos de 9 a 12, observamos o uso de *a gente* no tempo presente, assim como *nós* com formas verbais no passado:

(9) *[...] é horrivel* ***a gente estar*** | *separada dos seus.* [Trecho da carta de Alésia para sua irmã Graziela, 21/05/1941].

(10) As vezes | *sente dor,* ***a gente vê*** *pela expressão dele,* | *pelo modo de gritar, mas, a maioria das* | *vezes é só medo* [Trecho da carta de Iria para sua mãe Graziela Reis, 13/12/1954].

(11) *o Eduardo Cézar esteve adoentado sábado, com* | *forte gripe, e* ***nós fomos*** *visitá-lo, mas domingo,* êle | *já estava bem, e pouco [inint.] de casa.* [Trecho da carta de José Augusto para sua mãe Graziela, 20/08/1957].

(12) ***Nós tivemos*** *muito [p]razer com* | *isso porque para nós são pessoas de muita distin-* | ção. [Trecho da carta de Isaura para sua cunhada Graziela, 02/01/1940].

Esses dados corroboram os resultados encontrados por Vianna (2006), no Rio de Janeiro. Segundo a autora, alguns estudos (Omena, 1996; Lopes, 1993) têm mostrado que os tempos menos marcados tendem a condicionar o uso de *a gente*. Ainda, podemos levantar como hipótese que a forma *a gente* está sendo implementada a fim de desfazer a neutralização existente entre as formas verbais na P4 do presente e do passado em verbos regulares (Câmara Jr., 2011[1970]).

6 Considerações Finais

Neste artigo, tivemos como objetivo investigar as formas *nós* e *a gente* na função de sujeito nas cartas pessoais da amostra Família Arthur Reis, no período de 1940 a 1980, a fim de observarmos se havia variação entre essas duas formas na escrita amazonense, assim como verificar quais condicionadores linguísticos e extralinguísticos estariam atuando sobre a forma inovadora *a gente*. Considerando as características peculiares dessa amostra, conseguimos alcançar o objetivo proposto.

Os resultados mostraram que há variação entre *nós* e *a gente* para expressar a P4 na função de sujeito. Essa variação é encontrada, particularmente, nas cartas de duas missivistas, Iria e Alésia, a partir da década de 1940. Além disso, nossa principal hipótese de que os condicionadores que mais atuam sobre o uso de *a gente* são o *tipo de referente* e o *preenchimento do sujeito* foi comprovada, tendo em vista que uso de *a gente* na escrita dessas duas remetentes é favorecido, linguisticamente, pelo *sujeito expresso*, *pelo traço [-determinado]* e pelo *tempo presente*. Esse resultado é semelhante ao encontrado em outras pesquisas realizadas no PB, tanto em dados sincrônicos (Brustolin, 2009; Agostinho, 2013; Vitório, 2015; Santos, 2015; Freire, 2019 e Araújo, 2023) quanto diacrônicos (Vianna, 2006; Souza, 2015; Ribeiro e Vieira, 2019; Borges, 2019; Monguilhott *et al.*, 2021).

Vale destacar, ainda, que observamos uma tendência já evidenciada nas pesquisas descritas na seção 3 (Vianna, 2006; Brustolin, 2009; Vitório, 2015; Souza, 2015; Borges, 2019) de que as mulheres usam com mais frequência a forma inovadora para expressão de P4.

Espera-se que, com a ampliação da amostra, já em andamento pela equipe do PHPB/AM, possamos ter uma distribuição de dados equitativos por década/missivista e, assim, compreender de forma mais concreta a atuação dos condicionadores no uso da forma inovadora *a gente* na escrita manauara ao longo do tempo.

Referências

AGOSTINHO, S. R. N. *A variação na concordância verbal de primeira pessoa do plural na escrita de alunos do Ensino Fundamental.* 2013. 318 f. Dissertação (Mestrado em Linguística) – Programa de Pós-graduação em Linguística. Universidade Federal de Santa Catarina, Florianópolis, 2013.

ALBÁN, M. R; RAPP, C.; PASSOS, F.; OLIVEIRA, I.; CRUZ, R. Nós e a Gente: uma sondagem na norma culta brasileira. *Atas do I Simpósio sobre a Diversidade Linguística no Brasil.* Salvador, Instituto de Letras da UFBA, p. 147-156, 1986.

ARAÚJO, M. A. A. A primeira pessoa do plural em redações escolares de alunos do ensino médio: uma análise descritiva à luz da teoria variacionista. *Revista Contemporânea*, [*s.l.*], v. 3, n. 12, p. 30005-30028, 18 dez. 2023. South Florida Publishing LLC. Disponível em: https://ojs.revistacontemporanea.com/ojs/index.php/home/article/view/2678. Acesso em: 15 maio 2024.

ARAÚJO, M. A. A. *Será que a gente usa mais o nós?* Uma fotografia sociolinguística do falar popular de Fortaleza. Dissertação (mestrado acadêmico) – Universidade Estadual do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Linguística Aplicada, Fortaleza, 2016.

BORGES, Paulo Ricardo Silveira. As dimensões sociais da mudança em peças de teatro de autores gaúchos: inserção e propagação do pronome a gente no português brasileiro. *Cadernos do Il,* [*s.l.*], v. 1, n. 59, p. 71-88, 29 out. 2019. Universidade Federal do Rio Grande do Sul. http://dx.doi.org/10.22456/2236-6385.92310. Disponível em: https://seer.ufrgs.br/cadernosdoil/article/view/92310. Acesso em: 01 jun. 2024.

BRUSTOLIN, A. K. B. da S. *Itinerário do uso e variação de nós e a gente em textos escritos e orais de alunos do Ensino Fundamental da Rede Pública de Florianópolis.* 2009. 245 f. Dissertação (Mestrado em Linguística) – Programa de Pós-graduação em Linguística. Universidade Federal de Santa Catarina, Florianópolis, 2009.

CÂMARA JR., J. M. *Estruturas da língua portuguesa.* 44. ed. Petrópolis: Vozes, 2011 [1970].

CASTILHO, A. T. de. Produção científica do *Projeto Para a História do Português Brasileiro*, de 1988 a fevereiro de 2019. Bancos de dados do CEDOCH. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, 2021. Disponível em: https://cedoch.fflch.usp.br/producao-cientifica-do-projeto-para-historia-do-portugues-brasileiro. Acesso em: 03 out. 2024.

CONDE SILVESTRE, J. C. C. *Sociolinguística Histórica.* Madrid: Editorial Gredos, 2007.

CUNHA, M. A. F. de; OLIVEIRA, M. R. de; MARTELOTTA, M. E. *Linguística Funcional*: teoria e prática. São Paulo: Parábola Editorial, 2015.

FERNANDES, E; GOSKY, E. A concordância verbal com os sujeitos Nós e a Gente: um mecanismo do discurso em mudança. *Atas do I Simpósio sobre a Diversidade Linguística no Brasil.* Salvador, Instituto de Letras da UFBA, p. 175-83, 1986.

FREIRE, J. B. Concordância da 1ª pessoa do plural: o que dizem os textos escolares? Primeira Escrita, Aquidauana, n. 6, p. 157-168, 06 nov. 2019. *Dossiê Dialetologia e Geolinguística - princípios, abordagens e resultados.* Disponível em: https://periodicos.ufms.br/index.php/revpres/article/view/8523. Acesso em: 15 maio 2024.

FREITAS, J; ALBÁN. M.R.*"Nós ou a gente?" Estudos lingüísticos e literários,* Salvador, UFBA, 11:75–90, 1991.

HOPPER, P. J.; TRAUGOTT, E. C. *Grammaticalization*. Cambridge: Cambridge University Press, 1993.

HOPPER, P. On some principles of grammaticization. In: TRAUGOTT, E. C.; HEINE, B. (eds.): *Approaches to grammaticalization*, v. 1, Amsterdam/Philadelphia: John Benjamins Company, 1991.

INSTITUTO HISTÓRICO E GEOGRÁFICO BRASILEIRO. *Arthur César Ferreira Reis*: sócios falecidos brasileiros. Disponível em: https://ihgb.org.br/perfil/userprofile/acfreis.html. Acesso em: 15 maio 2024.

LABOV, W. *Padrões sociolinguísticos.* Trad. de Marcos Bagno; Maria Marta Scherre; Caroline Cardoso. São Paulo: Editorial Parábola, 2008 [1972].

LABOV, W. *Principles of linguistic change*: Internal factors. Cambridge: B. Blackwell, 1994.

LEHMANN, C. Grammaticalization: Synchronic Variation and Diachronic Change. *Lingua e Stile*, XX, 3, p. 303-318, 1982.

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, p. 47-80, jul. 2004. Disponível em: https://dialnet.unirioja.es/servlet/articulo?codigo=6137781. Acesso em: 15 maio 2024.

LOPES, C. R. dos S. A inserção de 'a gente' no quadro pronominal do português. Linguística Iberoamericana, Madrid/Frankfurt: *Iberoamericana/Vervuert*, v. 18, p. 1-167, 2003. Disponível em: https://laborhistorico.letras.ufrj.br/producao/Lopestese.pdf. Acesso em: 15 maio 2024.

LOPES, C. R. dos S. Nós e a gente no português falado culto do Brasil. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada,* [*s.l.*], v. 14, n. 2, 1998. Disponível em: https://revistas.pucsp.br/index.php/delta/article/view/44300. Acesso em: 29 jun. 2024.

LOPES, C. R. S. *Nós e a gente no português falado culto do Brasil,* Dissertação de Mestrado em Língua Portuguesa, Rio de Janeiro, Faculdade de Letras/UFRJ, 1993.

MAIA, E. G.; MARTINS, F. S.; COELHO, I. L. As formas de tratamento nas cartas pessoais da amostra família Arthur Reis. *Revista LaborHistórico*, v.10, n.2, e61210, 2024.

MILROY, J. *Linguistic Variation and Change*: On the Historical Sociolinguistics of English. Oxford: Blackwell, 1992.

MONGUILHOTT, I. O. S; CHAVES, R. G.; BRUSTOLIN, A. K. B. S.; CHAGAS, J. F. Variação nós e a gente em Santa Catarina: do presente para o passado. In: COELHO, I. L; MONGUILHOTT, I.O.S.; MARTINS, M.A.R.; GÖRSKI, E. M. (Orgs.). *Aspectos sócio-históricos e linguísticos do português escrito em Santa Catarina nos séculos XIX e XX*. Florianópolis: Editora da Ufsc, 2021. p. 155-170. Disponível em: https://editora.ufsc.br/estante-aberta. Acesso em: 01 jul. 2024.

MONTEIRO, J. L. *Os pronomes pessoais no português do Brasil.* Tese (doutorado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 1991.

NARO, A. J.; GÖRSKI, E; FERNANDES, E. Uma mudança linguística em curso: a concordância com o sujeito nós / a gente. *Trabalho apresentado durante o Seminário sobre Variação em Sintaxe.* Universidade Federal do Rio de Janeiro, 1983.

OMENA, N. P. A referência à primeira pessoa do plural: variação ou mudança? In: PAIVA, M. da C.; DUARTE, M. E. L. (Org.). *Mudança linguística em tempo real.* Rio de Janeiro: Contra Capa Livraria, 2003.

OMENA, N. P. A referência variável da primeira pessoa do discurso no plural. In: SILVA, G. M. de O. e; SCHERRE, M. M. P. (Org.). *Padrões Sociolinguísticos – Análises de fenômenos variáveis do português falado no Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro, p. 183-215, 1996.

PAIVA, M. C. A variável gênero/sexo. In: MOLLICA, M. C.; BRAGA, M. L. (org.). *Introdução à Sociolinguística*: o tratamento da variação. 4. ed. São Paulo: Contexto, 2013. p. 33-42.

RIBEIRO, L. C.; VIEIRA, M. S. Uso variável de nós e a gente em jornais publicados na Cidade de Goiás. *Entrepalavras*, [*s.l.*], v. 9, n. 3, p. 96, 2 dez. 2019. Disponível em: http://www.entrepalavras.ufc.br/revista/index.php/Revista/article/view/1536. Acesso em: 15 maio 2024.

ROMAINE, S. Socio-historical Linguistics: Its Status and Methodology. Cambridge: *Cambridge University Press*, 2009 [1982].

ROORYCK, J. On two types of underspecification: Towards a feature theory shared by syntax and phonology. *Probus,* 6, Berlin/New York, 1994, p. 207-233.

SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. *Goldvarb X*: A variable rule application for Macintosh and Windows. Department of Linguistics, University of Toronto, 2005. Disponível em: http://individual.utoronto.ca/tagliamonte/goldvarb.html. Acesso em: 20 maio 2024.

SANTOS, A. A. S. *O uso dos pronomes nós e a gente na escrita de alunos do ensino fundamental II da zona rural de Itaibó.* 2015. 116 f. Dissertação (Mestrado Profissional em Letras) –, Universidade Estadual de Santa Cruz, Ilhéus, 2015. Disponível em: http://www.biblioteca.uesc.br/biblioteca. Acesso em: 15 maio 2024.

SCHERRE, M. M. P.; NARO, A. J.; YACOVENCO, L. C. Nós e a gente em quatro amostras do Português Brasileiro: revisitando a escala da saliência fônica. *Revista Diadorim,* [*s.l.*], v. 20, p. 428-457, 30 dez. 2018. Programa de Pós-Graduação em Letras Vernáculas - PPGLEV. Disponível em: https://revistas.ufrj.br/index.php/diadorim/article/view/23285. Acesso em: 26 jun. 2024.

SOUSA, L. C. *Arthur Reis e a História do Amazonas*: um início em grande estilo. 2009. 146 f. Dissertação (Mestrado em História) - Universidade Federal do Amazonas, Manaus, 2009. Disponível em: https://tede.ufam.edu.br/handle/tede/3739. Acesso em: 15 jun. 2024.

SOUZA, K. P. F. A representação da 1ª pessoa do plural em cartas de autores não ilustres da sociedade carioca. In: BAALBAKI, A.; CARDOSO, J.; ARANTES, P.; BERNARDO, S. (Org.). *Linguagem*: teoria, análise e aplicações. Rio de Janeiro: UERJ / Programa de Pós-Graduação em Letras, 2015. p. 596-613. Disponível em: http://www.pgletras.uerj.br/linguistica/linguagem08.html. Acesso em: 29 jun. 2024.

VIANNA, J. B. S. *A concordância de nós e a gente em estruturas predicativas na fala e na escrita carioca.* 2006. 133 f. Dissertação (Mestrado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

VITÓRIO, E. S. L. A. Variação nós e a gente na posição de sujeito na escrita escolar. *Letras & Letras,* [*s.l.*], p. 128-143, 29 dez. 2015. EDUFU – Editora da Universidade Federal de Uberlândia. Disponível em: https://seer.ufu.br/index.php/letraseletras/article/view/31429. Acesso em: 15 maio 2024.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. *Fundamentos empíricos para uma teoria da mudança linguística.* Tradução de Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].

http://doi.org/10.5007/1984-8420.2025.e101203

MANUTENÇÃO LINGUÍSTICA E A VARIÁVEL 'GRAU DE RURALIDADE' EM FOCO: UM INSTRUMENTO METODOLÓGICO PARA ANÁLISE DE FENÔMENOS LINGUÍSTICOS RURAIS

LINGUISTIC MAINTENANCE AND THE VARIABLE 'DEGREE OF RURALITY' IN FOCUS: A METHODOLOGICAL INSTRUMENT FOR THE ANALYSIS OF RURAL LINGUISTIC PHENOMENA

Ivelã Pereira | Lattes | ivela.pereira@ifsc.edu.br
Instituto Federal de Santa Catarina

Loremi Loregian-Penkal | Lattes | loremi.loregian@gmail.com
Universidade Estadual do Centro-Oeste do Paraná

Resumo: Esta investigação tem como foco a *manutenção linguística* a partir da descrição quantitativa e qualitativa do fenômeno variável da alternância vocálica precedente ao sufixo de PN /mos/, nos verbos regulares de 1ª e 2ª conjugação, em formas canônicas, como *falamos* e *aprendemos,* e não canônicas, como *falemo* e *aprendimo*, bem como sua relação com a ruralidade e o conservadorismo linguístico. Trata-se de uma pesquisa no âmbito da Teoria da Variação e da Mudança (cf, Labov, 2008 [1972], 1982, 1994; Weinreich, Labov, Herzog, 2006 [1968]; Guy, 1981, 2000), Dialetologia (Vasconcelos, 1901; Boléo, 1943; Melo, 1946) e Linguística Histórica (Castilho, 1968, 1992, 2016). Tomamos por base a defesa de Naro e Scherre (2003) de que áreas rurais constituem-se em "ilhas de conservadorismo linguístico" ao manterem usos de um português menos recente (dos séculos XIX e XX, considerado pelos dialetólogos como "português moderno"). Nossa amostra é composta por 168 entrevistas sociolinguísticas do banco-base VARLINFE (Variação Linguística de Fala Eslava) – vinculado ao NEES (Núcleo de Estudos Eslavos), da UNICENTRO (câmpus Irati), de característica majoritariamente rural. A análise quantitativa considerou algumas variáveis linguísticas e extralinguísticas, sendo selecionada, pelo programa de análise multivariada GOLDVARB-X, a variável linguística complexa 'grau de ruralidade', construída especificamente para este estudo com base em estudos dialetológicos que mapearam fenômenos rurais e conservadores da língua. Assim, foram eleitas algumas particularidades linguísticas consideradas rurais e, a cada informante entrevistado, fez-se uma escala de uso, considerando três graus de ruralidade: *baixo*, *médio* e *alto*. Os resultados mostraram que o uso das variantes não-canônicas (/e/ em "falemo" e /i/ em "aprendimo") estava significativamente inter-relacionado ao alto grau de ruralidade dos informantes, um indício de que seu uso seja mais proeminente em contextos de conservadorismo linguístico e, consequentemente, de manutenção linguística de formas conservadoras. Acreditamos que o uso da variável independente interna

'grau de ruralidade' possa auxiliar pesquisadores a trabalharem com fenômenos rurais variáveis e também com o tópico ainda pouco explorado (cf. Ribeiro e Lacerda, 2013) da *manutenção linguística*.

Palavras-chave: Alternância vocálica. Manutenção linguística. Variação morfofonológica. Ruralidade.

Abstract: This investigation focuses on linguistic maintenance based on the quantitative and qualitative description of the variable phenomenon of vowel alternation preceding the PN suffix /mos/, in regular verbs of the 1st and 2nd conjugation, in canonical forms, as we speak and learn, and not canonical, how we speak and learn, as well as its relationship with rurality and linguistic conservatism. This is research within the scope of the Theory of Variation and Change (cf, Labov, 2008 [1972], 1982, 1994; Weinreich, Labov, Herzog, 2006 [1968]; Guy, 1981, 2000), Dialectology (Vasconcelos, 1901; Boléo, 1943; Melo, 1946) and Historical Linguistics (Castilho, 1968, 1992, 2016). We take as a basis the defense of Naro and Scherre (2003) that rural areas constitute "islands of linguistic conservatism" by maintaining uses of a less recent Portuguese (from the 19th and 20th centuries, considered by dialectologists as "modern Portuguese"). Our sample is made up of 168 sociolinguistic interviews from the base database VARLINFE (Linguistic Variation of Slavic Speech) – linked to NEES (Center for Slavic Studies), at UNICENTRO (Irati campus), with a mostly rural characteristic. linguistic and extralinguistic, using the GOLDVARB-X multivariate analysis program to select the complex linguistic variable 'degree of rurality', constructed specifically for this study based on dialectological studies that mapped rural and conservative language phenomena. linguistic particularities considered rural and, for each interviewed informant, a usage scale was made, considering three degrees of rurality: low, medium and high. The results showed that the use of non-canonical variants (/e/ in "falemo" and /i/ in "aprendimo") was significantly interrelated to the high degree of rurality of the informants, an indication that their use is more prominent in contexts of linguistic conservatism and, consequently, linguistic maintenance of conservative forms. We believe that the use of the internal independent variable 'degree of rurality' can help researchers work with variable rural phenomena and also with the topic that is still little explored (cf. Ribeiro and Lacerda, 2013) of linguistic maintenance.

Keywords: Vowel alternation. Linguistic maintenance. Morphophonological variation. Rurality.

1 Introdução

O foco desta investigação é a *manutenção linguística*, *conservadorismo linguístico* e sua relação com a variável complexa 'grau de ruralidade' (Pereira[1], 2021), a partir do estudo da alternância vocálica na primeira pessoa do plural em verbos de 1ª e 2ª conjugação no presente e pretérito perfeito do indicativo (como *falamos* e *aprendemos*, que são realizados de forma não canônica como *falemo* e *aprendimo*).

Este fenômeno já havia sido mencionado por Amaral (1920) e por Castilho (1992), que o descrevera afirmando haver uma "elevação da vogal temática *a* para *e* e *e* para *i* no pretérito perfeito do indicativo, para distingui-lo do presente do indicativo: *fiquemo* (por *ficamos*), *bebimo* (por *bebemos*)" (Castilho, 1992, p. 250). Também em outra obra, o pesquisador, ao tratar sobre a morfologização dos sufixos de modo e tempo do latim vulgar até o português, especifica que a distinção antes ocorrida entre os tempos verbais no latim vulgar seria restabelecida no português popular, "elevando a vogal temática no pretérito de C1 e C2 (cf. *amamos ~ amemos*, *bebemos ~ bebimos*)" (Castilho, 2016, p. 152). Isto é, o português popular, tendo uma neutralização[2] dessas formas de presente e pretérito perfeito do indicativo em primeira pessoa do plural, acabam por trazer uma estratégia de diferenciação entre os tempos verbais, alternando-se a vogal tônica precedente ao sufixo de pessoa e número /mos/.

Assim, fica uma reflexão: a vogal em alternância (/e/ em 1ª conjugação e /i/ em 2ª conjugação) consistiria numa variante inovadora (por divergir da forma padrão instaurada) ou numa variante conservadora, por tentar restabelecer uma diferenciação de tempo-modo-aspecto que ocorria no latim vulgar? Ainda não temos respostas convincentes para essa questão, mas os dados coletados na contemporaneidade têm mostrado a manutenção/conservação linguística das formas em /emo/ e /imo/, de maneira que é possível considerá-las conservadoras no sentido de que atravessam séculos na língua sendo utilizadas por falantes de comunidades rurais. Neste artigo, portanto, defendemos que o uso das variantes não padrão /e/ e /i/ consista numa estratégia mais conservadora da língua.

Com vistas a uma contextualização, trazemos, a seguir, alguns exemplos do fenômeno em variação retirados de nosso córpus:

---

[1] Parte das ideias deste artigo foi publicada na tese de Doutorado de Pereira (2021).

[2] Lembrando que essa distinção entre passado e presente ocorre na norma padrão do português europeu. De acordo com Castilho (1992, p. 246), a respeito de "características fonológicas" do português do Brasil: "O PB não opõe timbres abertos a timbres fechados da vogal a seguida de nasal: cf. PB presente e pretérito *cantamos*, PP presente *cantamos* / pretérito **cantámos**".

1ª conjugação

(1) Entrevistadora: E, assim, hoje em dia, você fala ucraniano com teu marido? Entrevistada: **Fal[ɐ̃]mos**. **(PRU[3], fem[4], fund. II, mais jovem)**

(2) Entrevistadora: E falando em mudá... que festas que vocês assim já... Natal, Páscoa, assim, vocês comemoravam sempre ou não? Entrevistada: Não, era só o Natal da baba, né? Sempre comera- **comemor[ɐ̃]mo** até o dia de hoje o Natal aqui em casa é dia sete de janeiro. [...] Em dezembro, nós sempre dizia assim: Natal dos brasileiro. [sobre o que fazem hoje em dia] **(IRA, fem, fund. I, mais jovem)**

(3) Entrevistadora: Mas tá OK, já deu nosso tempo, [nome feminino ocultado], brigada pela... Entrevistada: Hum... não **fal[e]mo** nem metade, não **convers[e]mo** nem metade... Entrevistadora: [risos] Quanta história que cê tem pra contá, mulher? Entrevistada: Teim, teim... e não te mostrei minhas boneca... **(IRA, fem, fund. I, mais velha)**

(4) Nós **mud[e]mos** umas treis veiz lá pra Guarapuava, sabe? Até mais pra lá de Pinhão, sabe? Fomo, **volt[e]mo**, **volt[e]mo** embora, vol- se **mud[e]mo** de novo pra Guarapuava, porque aí ele num queriam que ficasse sozinho. **(IRA, fem, fund. II, mais velho)**

2ª conjugação

(5) Eu, até outro dia, nós fomo na casa do tio do meu marido, porque ele mor- mora em Concórdia, nossa! Aí nós **com[e]mos** no velório, almoçamos [velarizado] i [hes] cantum e rezum assim... **(PRU, fem, ens. médio, mais velho)**

(6) Na verdade, na Quaresma ali, a gente, como agora pode teR carne, o nosso costume é ansim, nas quarta e sexta-fera, nós num **com[e]mo** carne. **(REB, mas, fund. I, mais velho)**

(7) [sobre como era antigamente] Então, era sufrido. Nóis **sufr[i]mo**. **(REB, mas, fund. I, mais jovem)**

(8) Os professores, na época que eu estudava, professores rigorosos, mas, pelo menos, nós **aprend[e]mos** coisas boas e valorizamos hoje tudo isso que nós **aprend[i]mos**. **(CRU, mas, fund. II, mais velho)**

---

[3] Os municípios onde foram feitas as entrevistas estão localizados na mesorregião sudeste do Paraná e são os seguintes, com suas respectivas abreviaturas: Irati (IRA), Prudentópolis (PRU), Ivaí (IVA), Cruz Machado (CRU), Rebouças (REB), Rio Azul (RIO) e Mallet (MAL).

[4] Todos os dados apresentados neste artigo trarão estas informações, conforme a estratificação do banco-base do VARLINFE: município, gênero (feminino e masculino), grau de escolaridade (fundamental I, fundamental II e ensino médio) e faixa etária (mais jovem e mais velho).

É possível observar que a variação entre as vogais ocorre, inclusive, em proximidade numa mesma frase de um mesmo informante, como se pode ver em (8), quando o sujeito produz "aprendemos" e, logo na sequência, "aprendimos". Além disso, o fato de ambas as ocorrências serem produzidas com o sufixo *mos* – sem a supressão de */s/* – é um indício de que o uso não padrão possa ocorrer até mesmo em contextos de fala monitorada, pois o uso de /s/ no sufixo número-pessoal é pouco usual até mesmo em variedades cultas, aparecendo em falas mais monitoradas.

Parece-nos que e o uso de /e/ em 1ª conjugação e de /i/ em 2ª conjugação consista numa característica intrínseca a comunidades de fala rurais. Mas essa constatação já havia sido feita por Amaral (1920) ao descrever tal fenômeno como típico do "dialeto caipira", e também aparece em descrições de outros pesquisadores que o reconheceram como característico a variedades rurais no Brasil (Frosi, Mioranza, 1983; Costa, 1990; Bortoni-Ricardo, 2005, 2011; Naro, Scherre, 2007; Pereira, 2016; Pereira e Margotti, 2018) e em Portugal, numa variedade menos recente do português moderno – séculos XIX e XX (Vasconcelos, 1970 [1901]).

Assim, em consideração a esse ponto crucial da ruralidade, embora a variabilidade das formas linguísticas seja um aspecto essencial à nossa investigação – especialmente por se tratar de uma pesquisa no âmbito da TVM (Teoria da Variação e da Mudança) – priorizamos aqui um ponto que ainda está sendo pouco explorado na área: a *manutenção linguística* e o *conservadorismo linguístico*. Como afirmam Ribeiro e Lacerda (2013, p. 91), a mudança linguística tem sido foco dos estudos sociolinguísticos, mas "o tratamento da *não mudança* não se deu com a mesma intensidade. Assim, na maior parte dos estudos brasileiros, a não mudança (conservadorismo linguístico ou manutenção linguística) ficou em segundo plano".

Nesse contexto, ao considerarmos que estudos com essa temática se fazem essenciais, procuramos trazer uma proposta de análise que considere a variável independente 'grau de ruralidade' (construída especificamente para a pesquisa quantitativa sobre /emo/ e /imo/) e é medida por critérios puramente linguísticos, que podem ser captados mediante entrevistas orais gravadas. Ou seja, a variável linguística construída leva em conta outras características linguísticas que são denominadas como "rurais" por linguistas portugueses e brasileiros.

O objetivo é que o pesquisador possa utilizá-la como um instrumental metodológico, bem como estabelecer relações entre seu fenômeno linguístico, características rurais e o conservadorismo linguístico. Tal variável independente é aplicada num estudo

quantitativo no âmbito da Teoria da Variação e da Mudança (cf, Labov, 2008 [1972], 1982, 1994; Weinreich, Labov, Herzog, 2006 [1968]; Guy, 1981, 2000) sobre a alternância vocálica não canônica na primeira pessoa do plural no presente e pretérito perfeito do indicativo, procurando testar se o uso não canônico apresentaria mais significância em informantes com nível maior de ruralidade.

Dessa forma, este estudo se justifica por se direcionar a contribuir para pesquisas no âmbito da Sociolinguística, Dialetologia e Linguística Histórica voltados a variedades rurais e à questão (ainda em coadjuvância) da *manutenção linguística*. Esperamos também que nossas discussões relativas à alternância vocálica precedente a /mos/ em associação às "ilhas de conservadorismo linguístico" sejam fecundas no campo dos estudos históricos da língua, em busca do fortalecimento da tese de Naro e Scherre (2003) sobre "ilhas de conservadorismo linguístico" que mantêm, por períodos de tempo consideráveis, usos de fenômenos da língua fora do padrão, os quais são encontrados hodiernamente, mas também em séculos anteriores (como XIX e XX, por exemplo).

Isso posto, aqui estão as questões primordiais que guiam esta pesquisa:

(i) É possível relacionar a variável 'grau de ruralidade' com "ilhas de conservadorismo linguístico", isto é, a ruralidade linguística e o conservadorismo na língua estão intrinsecamente ligados?

(ii) Se um fenômeno linguístico, embora divirja da norma padrão da língua portuguesa (e, portanto, não é ensinado na escola, nem propagado pela mídia e outros meios), aparece em dados do português dos séculos XIX e XX, mas também em dados atuais do português brasileiro, *como* e *por que* tem atravessado séculos e continua sendo encontrado em variedades brasileiras? Poderia, então, ser considerado como algo característico a "ilhas de conservadorismo linguístico" no Brasil?

(iii) De que maneira uma descrição linguística mais ampla num estudo quantitativo pode ajudar o pesquisador a perscrutar melhor o fenômeno linguístico que é foco de análise e classificá-lo como rural (ou não)?

A fim de responder a esses e outros questionamentos, este artigo está assim organizado: após a primeira seção (da Introdução), na qual apresentamos o objeto de pesquisa, o objetivo geral, as questões norteadoras e a justificativa, está a segunda seção, "Ilhas de conservadorismo linguístico: características gerais", a qual busca definir o conceito de *ilhas de conservadorismo linguístico* e sua relação com a *manutenção linguística*.

Na terceira seção, com o título "'Grau de ruralidade': uma variável linguística complexa para o estudo de fenômenos linguísticos rurais", é apresentada a proposta de uma variável construída para estudos quantitativos focados em variedades rurais. Testamos o 'grau de ruralidade' a partir de um estudo sociolinguístico do fenômeno variável da alternância vocálica feito em localidades rurais do sudeste do Paraná, estabelecendo-se os devidos cruzamentos de variáveis e discussões de resultados, tanto da primeira quanto da segunda conjugação verbal.

Por fim, finalizamos com a seção das "Considerações finais", retomando questões norteadoras já elencadas na seção introdutória.

2 Ilhas de conservadorismo linguístico: características gerais

Objetivamos tratar sobre variedades rurais no português brasileiro e sua relação com "ilhas de conservadorismo linguístico", a fim de, após tais explanações, construir uma variável complexa linguística para o estudo de fenômenos rurais. Nesse sentido, trazemos à vista algumas especificações a respeito da fala rural.

Em relação a isso, Ribeiro (2013, p. 42, grifos nossos) elucida que:

> [...] **a fala rural é conservadora e desprestigiada, enquanto a fala urbana é inovadora e dotada de prestígio social**. Não podemos perder de vista que o estabelecimento da variedade urbana do português brasileiro é fruto de um processo bastante divergente daquele que constituiu a variedade rural. Enquanto nesta o aprendizado foi assistemático, de "oitiva", naquela houve, inicialmente, grande influência da escola, com seu caráter normatizador e, posteriormente, de um forte papel unificador desempenhado pela mídia. Essa diferença foi fundamental na formação do português brasileiro e ainda o é pelo que se pode perceber na diversidade linguística em nosso país.

Também Amaral (1920) já havia descrito minuciosamente a variedade rural do português: "o nosso falar caipira – bastante característico para ser notado pelos mais desprevenidos como um sistema distinto e inconfundível – dominava em absoluto a grande maioria da população e estendia a sua influência à própria minoria culta." (Amaral, 1920, p. 1).

Porém, o pesquisador parece ter ser enganado quando afirmou que "este [dialeto caipira] acha-se condenado a desaparecer em prazo mais ou menos breve. Legará, sem dúvida, alguma bagagem ao seu substituto, mas o processo novo se guiará por outras determinantes e por outras leis particulares." (Amaral, 1920, p. 2). Passados mais de cem

anos desde que Amaral (1920) fez tal afirmação, os fenômenos linguísticos descritos por ele em 1920 continuam a aparecer em comunidades rurais do Brasil, como é o caso da alternância vocálica de /amo/ ~ /emo/ e /emo/ ~/imo/, por exemplo.

Ademais, alguns deles aparecem igualmente na *variedade rurbana*, que está no *intermezzo* entre rural e urbano, sendo descrita por Bortoni-Ricardo (2011) como um *continuum* que perpassa os dois espectros (de urbanidade e ruralidade). De modo similar, o fenômeno da alternância vocálica também foi encontrado e descrito por Bortoni-Ricardo (2005; 2011) em uma comunidade rurbana de Brasília.

Verificamos hoje, nos anos 2020, a permanência e manutenção da variedade rural em localidades diversas no Brasil e também a *manutenção linguística* de fenômenos que pareciam estar fadados ao desaparecimento (cf. Amaral, 1920), mas sobreviveram firmes e fortes na contemporaneidade. Nesse contexto, concordamos com Naro e Scherre (2007) ao afirmarem que a manutenção linguística ocorre porque as áreas rurais apresentam um grau de isolamento maior, em oposição aos grandes centros urbanos, fazendo delas *ilhas de conservadorismo* em que formas arcaicas do português europeu teriam se mantido (e fidelizado).

Enquanto as comunidades urbanas têm uma característica mais cosmopolita e, por consequência, parecem estar mais abertas a inovações linguísticas, as comunidades rurais brasileiras apresentam certo isolamento geográfico, o qual resulta num isolamento também linguístico, que lhes permite manter formas arcaicas da língua, inclusive as trazidas de um português europeu. Nessa direção, Melo (1946, p. 91) já asseverava: "a nossa língua popular, falando-se de um modo geral, é substancialmente o português arcaico deformado [sic], ou se quiserem, transformado em certo aspecto da morfologia e alguns da fonética pela atuação dos índios e dos negros."

Também Elia (1975, p. 209) acreditava que a "característica dos nossos falares é [...] a *arcaicidade*. De fato, o português do Brasil se aproxima bastante da maneira de falar do português quinhentista". De modo similar, a pesquisadora portuguesa Cardeira (2006), menciona a existência de inúmeros *arcaísmos* existentes no Brasil, os quais estão presentes em âmbitos fonético-fonológico, morfológico, prosódico, sintático e lexical. Em decorrência de uma *tendência arcaizante* do português brasileiros, muitas variantes já suplantadas no português europeu ainda ocorrem em regiões do interior brasileiro.

Precisamos enfatizar também que, além da tendência de manutenção de fenômenos linguísticos mais antigos no português, as ilhas de conservadorismo mantêm, outrossim, outros elementos em conservação, como os costumes, tradições, aspectos culturais

e, até mesmo, outras línguas que foram trazidas por imigrantes de outros países (além de Portugal) em certas regiões do Brasil.

Um exemplo de conservação de sua(s) língua(s) de imigração é o da comunidade pesquisada por Vandresen (2009), localidade em que o Pomerano manteve-se por décadas e passou por um apagamento linguístico ínfimo, se comparado, por exemplo, a comunidades urbanas que falavam essa língua, mas deixaram de falá-la. Como afirma Vandresen (2009):

> Esta atitude ou **lealdade linguística** ao Pomerano tem garantido sua manutenção, graças à localização da comunidade dentro de uma **ilha linguística Pomerana**, em que praticamente todas as pessoas são bilíngues e, por exercerem atividades agrícolas, suas redes de comunicação são fechadas e densas, favorecendo a resistência a forças inovadoras, como a troca pelo português. Como vimos nos dados, mesmo na faixa de 25 aos 50 anos, prevalece o uso do dialeto pomerano no lar, apesar de esta geração e a mais velha (mais de 51 anos) terem frequentado escolas em língua portuguesa. Desta forma, o **isolamento geográfico** explica, em parte, os padrões de escolha das línguas, em função dos interlocutores. (Vandresen, 2009, p. 11, grifos nossos).

Outra comunidade bilíngue que apresenta essa mesma característica de isolamento é a comunidade ucraniana de Dorizon, que foi estudada por Wouk (1981). Segundo a pesquisa feita, "a comunidade continua vivendo relativo isolamento geográfico e cultural, o qual se vê reforçado pela organização familiar, através dos casamentos quase sempre endogâmicos". (Wouk, 1981, p. 64). Assim, o que possibilita a conservação/manutenção linguística, tanto do português (língua da localidade) quanto do ucraniano (língua de imigração), é o seu isolamento.

E, quando se fala em *isolamento*, é preciso ressaltar que ele está relacionado não somente no âmbito geográfico, mas também no relacionamento entre os sujeitos da comunidade, além do seu índice de localismo e de mobilidade. Por isso, ao se tratar sobre isolamento de comunidades, os estudos sobre redes sociais (Cf. Milroy, 2002) são cruciais. Justamente nessa direção. Bortoni-Ricardo (2011, p. 135, grifos nossos) assegura que:

> **As redes isoladas tendem a favorecer a manutenção da cultura rural** e, portanto, a focalização do vernáculo. Exibem um alto grau de densidade consensual ou moral que funciona como um mecanismo de resistência à mudança. Tal resistência não opera necessariamente ao nível da consciência, *i.e.*, a resistência pode não ser uma atitude consciente motivada por um tipo de oposição intergrupal aguda e conflitiva, frequentemente encontrada nos países industriais ocidentais Pelo

> contrário, é consequência do próprio estado de isolamento. A função de reforço mútuo tende a ser menos influente em redes integradas.

Desse modo, a densidade das redes sociais estabelecidas numa comunidade está igualmente relacionada ao isolamento dessa comunidade. De acordo com Ribeiro (2013, p. 28): "logo, associando as definições de Milroy (1980) às de Bortoni-Ricardo (2011), as redes isoladas tendem a ser densas e multiplex [multiplexas], enquanto as redes integradas são frouxas e uniplex [uniplexas]".

Assim como Ribeiro (2013), consideramos que *manutenção linguística* e *conservadorismo linguístico* sejam sinônimas, considerando que existe uma prevalência da utilização de variantes conservadoras nas comunidades rurais, que se constituem em "ilhas de conservadorismo linguístico". Nessa lógica, as comunidades com maior isolamento linguístico teriam uma tendência maior para o uso de variantes conservadoras/arcaicas.

> [...] evidencia-se que a divisão geográfico-social da história de nosso país está atrelada à história da nossa língua. Nesse sentido, o fato de o Brasil ter sido "um país essencialmente rural" no período colonial e hoje se configurar como "um país eminentemente urbano" se reflete na polarização linguística rural/urbana que se evidencia no português brasileiro atual, em função de um "conjunto de duas épocas estratificadas". Da mesma forma – e como consequência –, os espaços geográfico-sociais, polarizados em função da história do nosso país, tendem a refletir o perfil linguístico conservador/inovador, embora, para essa caracterização, seja preciso levar em consideração, também, a influência da mídia e da escola. (Ribeiro, 2013, p. 37).

As comunidades rurais do Brasil, portanto, apresentariam características não apenas geográficas de isolamento, mas também históricas, linguísticas, sociais e culturais. Destarte, fenômenos linguísticos com tendências ao desaparecimento em comunidades urbanas mais suscetíveis às mudanças linguísticas podem traçar um caminho bastante inverso em comunidades rurais: no viés do conservadorismo/manutenção linguística. Acreditamos que seja esse exatamente o caso da alternância vocálica precedente a /mos/ no português brasileiro.

Como já mencionado, o uso de /emo/ e /imo/ está relacionado ao contexto rural, e isso já foi demonstrado em diversas outras pesquisas[5] (cf. Amaral, 1920; Bortoni-Ricardo, 2005, 2011; Frosi e Mioranza, 1983; Costa, 1990; Naro, Scherre, 2007, Pereira e Margotti, 2018). Mas outro ponto que não pode ser desconsiderado é que dados de uso de /emo/ e /imo/ foram registrados não apenas em Amaral (1920), mas também na tese

[5] Por uma questão de concisão e foco, não tratamos de todas essas pesquisas neste artigo. Mas o leitor pode ter acesso a elas conforme as nossas referências.

do dialetólogo Vasconcelos (1901) ao descrever o português de comunidades rurais no final do século XIX em Portugal. Naro e Scherre (2007) já sinalizaram que traços linguísticos do português europeu não-padrão podem ser vistos no português brasileiro, e citam o uso de -êmos ao invés de –*amos*, com base em portugueses como Braga (1971) e Cruz (1991). Os linguistas brasileiros defendem haver uma confluência de formas não-padrão que ocorrem no português europeu e no português brasileiro rural.

Além disso, também Foeger (2014), baseada em Braga (1971) e Naro e Scherre (2007), relacionou a hipótese de Naro e Scherre (2007) com a existência de dados não canônicos obtidos em seu córpus (*estudemo*, *aprendimo* e *fumo*) cujo tema principal era a alternância dos pronomes *nós/a gente* e a aplicação da concordância verbal de primeira pessoa do plural:

> É interessante observarmos que esse é um traço verificado na variedade popular do PE, como notado por Naro e Scherre (2007). Os autores localizam em terras lusitanas "a origem de estruturas linguísticas portuguesas não-padrão, que em função de uma confluência de motivações, se ampliaram e se tornaram visíveis em terras hoje brasileiras" (Naro; Scherre, 2007, p. 23). **Partindo desse pressuposto, podemos pensar na elevação da vogal temática em verbos de primeira conjugação na 1PP como uma herança do português popular europeu que ainda se conserva na área rural do Brasil.** (Foeger, 2014, p. 140, grifos nossos).

Não podemos nos deter nesses registros e sua relação com o português europeu neste artigo, mas supomos que, conforme Melo (1946, p. 117), "a 'língua brasileira' é muito 'portuguesa' demais", no sentido de que muitos fenômenos linguísticos não padrão localizados em Portugal encontram-se "conservados" em variedades rurais do Brasil. Nesse sentido, em ancoragem à defesa de Naro e Scherre (2007), bem como no estudo de Svobodová (2017) e nas explanações de dialetólogos portugueses (Vasconcelos, 1901; Delgado, 1951; Moura 1960; Santos, 1967; Oliveira, 1966; Baptista, 1967; Cruz, 1969; Carrancho, 1969; Braga, 1977; Maia, 1977; Garcia, 1979; Gonçalves, 1988; Alves, 1993; Faria, 1999; Florencio, 2001; Simão, 2011), supomos, neste artigo, que a alternância vocálica precedente ao SNP /mos/ no presente e pretérito perfeito possa ter suas raízes no português europeu. Mas como podemos estabelecer essa relação se ainda não pudemos captar (a partir de estudos precedentes) como emerge este uso na língua?

Por falta (ainda) de uma recuperação histórica pontual da emergência desse fenômeno na língua, o que não nos permite traçar uma linha temporal, diacrônica, que revele detalhadamente o surgimento desse fenômeno e sua manutenção/conservação ao longo dos séculos, esta pesquisa se detém em como verificar o conservadorismo linguístico com base em evidências apenas sincrônicas.

Expliquemos melhor: as similaridades entre os registros de dados em Portugal e no Brasil é o aspecto da ruralidade, isto é, se as pesquisas mostraram que o uso de /emo/ e /imo/ está intrinsecamente relacionado a variedades rurais, as quais, por sua vez, costumam ser conservadoras, estabelecer concatenações entre fenômenos linguísticos conservadores e a ruralidade se faz primordial.

Com esse objetivo e buscando aprimorar as pesquisas linguísticas sincrônicas que tratam de fenômenos já registrados em séculos passados, a próxima seção tem como foco a ruralidade e uma proposta de instrumental metodológico de viés quantitativo e qualitativo para estudos sociolinguísticos voltados a fenômenos rurais/conservadores e à manutenção linguística.

3 'Grau de ruralidade': uma variável linguística complexa para o estudo de fenômenos linguísticos rurais

Defendemos haver uma relação intrínseca entre fenômenos linguísticos conservadores e a ruralidade, mas como podemos atestar essa relação numa pesquisa de âmbito sincrônico, que não dispõe de dados em distintos recortes temporais para se fazer uma comparação? Ou seja, como um pesquisador sociolinguista consegue verificar se o fenômeno linguístico em análise é de fato conservador na língua, se ele não possuir a seu dispor dados que lhe permitam mostrar *mudança* (ou *manutenção*) *linguística* em tempo real ou mesmo em tempo aparente[6]?

Respondemos essa questão argumentando, neste artigo, que é possível verificar indícios de conservadorismo linguístico analisando-se, de modo mais qualitativo, a fala dos informantes por meio de um mapeamento de outros fenômenos da língua que são considerados rurais.

Assim, primeiramente trazemos à vista um arrolamento de usos linguísticos que estão associados à ruralidade, para, na sequência, sistematizar um quadro de análise que, posteriormente, é usado na oitiva de entrevistas de informantes, a fim de fazer um levantamento geral de uso dos fenômenos elencados, classificando o informante em um nível de ruralidade.

3.1 Definições metodológicas: a construção de níveis de ruralidade

Existem fenômenos nitidamente idiossincráticos a variedades rurais conforme Amaral (1920), Rodrigues (1975), Penha (1974), Elia (1975), Head (1978), Bortoni-Ricardo (2011) e Ribeiro (2013). De acordo com Bortoni-Ricardo (2011, p. 59), "as variedades caipiras exibem um alto nível de uniformidade em suas características fono-

[6] Nossa amostra não permite fazer estudos de mudança em tempo aparente, porque o banco VARLINFE está estratificado em duas faixas etárias, segundo o mesmo padrão do primeiro banco de dados do VARSUL na década de 1990.

lógicas e morfonêmicas, o que é comprovado pela extensa comparação de descrições de falares rurais fornecida por Elia (1975)."

Com base nisso, elencamos um conjunto de características linguísticas nitidamente rurais conforme a descrição de linguistas e também levando em consideração que seriam facilmente reconhecíveis como rurais por falantes do português.

Ribeiro (2013), descrevendo a concordância verbal e nominal em uma comunidade linguística do interior de Minas Gerais, faz um pequeno compilado deste teor, o qual tomamos por base num primeiro momento:

> a) No nível fonético: nasalização do /i/ (igual>ingual, igreja>ingreja); perda
> de nasalização da vogal átona final (virgem>virge); síncope (cócega>cosca);
> inversão do /w/ (tábua>tauba); redução dos ditongos (baixo>baxo,
> autoridade>otoridade); vocalização da palatal /ʎ/ (filha>fia) ou hipercorreção
> (alfaiate>arfalhate); permuta de /l/>/r/ e /v/>/b/ (problema>probrema;
> verruga>berruga); apagamento de diferentes segmentos sonoros (pode>pó,
> perto>per, como é>cumé, com a >ca, dentro da>denda, para>pa/pra, pra
> você>procê); etc. b) No nível morfológico: flexão de plural apenas em um
> elemento do sintagma nominal (os livros>os livro); ausência ou pouco uso do
> subjuntivo (ame>amá); prevalência da desinência de terceira pessoa do
> singular no uso dos verbos, exceto na primeira pessoa do singular (eu amo,
> tu/você/ ocê,/cê ama, ele/ei ama, nóis ama, voceis/ôceis/ceis ama, eles/eis
> ama), o que resulta na ausência de concordância verbal; alteração fonética das
> desinências do pretérito perfeito do indicativo (foram> foru;
> andaram>andaru); etc. c) No nível sintático: emprego dos pronomes retos
> como acusativo (ele viu nóis na casa; não vi tu lá); etc. (Ribeiro, 2013, p. 39-40).

Dentre todas essas particularidades mencionadas, as de nível fonético aparentam ser as mais salientes às variedades rurais, sendo uma delas mencionada por Bortoni-Ricardo (2011, p. 77): "a vocalização do /ʎ/ tornou-se um claro estereótipo caipira no sentido usado por Labov (1972, p. 180). No processo de urbanização, os migrantes rurais parecem tornar-se logo conscientes do estigma associado com a variável e muitas vezes recorrem a hipercorreções [...]". Ademais, a vocalização da lateral alveopalatal (*despalatalização* ou *yeísmo*), é caracterizada como rural também por Amaral (1920), Elia (1975), Rodrigues (1975) e Penha (1974).

Seria esse, então, um traço importante para considerarmos em nosso rol de fenômenos rurais. Os outros pontos foram igualmente pensados com base nas descrições dos dialetólogos já mencionados. Com base em todos esses estudos, sistematizamos o quadro a seguir, que servirá como base para a definição do 'grau de ruralidade' de cada informante:

Quadro 1: Fenômenos linguísticos característicos de variedades rurais

CARACTERÍSTICAS DA VARIEDADE RURAL		
Fenômenos linguísticos	**Exemplos**	**Pesquisadores que o citam**
Vocalização da lateral palatal [ʎ]	Palhaço – *paiaço* Falha – *faia* Filha – *fia*	Amaral (1920), Rodrigues (1975), Penha (1974), Elia (1975) e Bortoni-Ricardo (2011)
Nasalização de vogal que não antecede segmentos nasais	Igreja – *ingreja* Igual – *ingual* Eleição – *inleição*	Amaral (1920), Rodrigues (1975), Penha (1974), Elia (1975) e Bortoni-Ricardo (2011)
Rotacismo (permuta de [l] para [r])	Planta – *pranta* Bloco – *broco* Volta – *vorta*	Amaral (1920), Rodrigues (1975), Penha (1974) e Bortoni-Ricardo (2011)
Permuta de fricativa labiodental sonora [v] em oclusiva bilabial sonora [b] e vice-versa	Travesseiro – *trabissero* Bravo – *brabo*	Amaral (1920), Penha (1974) Ribeiro (2013)
Prótese do -a- em vocábulos que iniciam por consoante ou aférese (queda) do -a-	Divertir – *adivertir* Voar – *avoar*	Amaral (1920), Penha (1974), Elia (1975) e Bortoni-Ricardo (2011)
Rótico pertencente ao rol dos retroflexos em coda e onset silábicos – tepe retroflexo [ɽ] e aproximente retroflexo [ɻ]	Cor – [ˈkoɽ]	Amaral (1920), Rodrigues (1975), Head (1978)

Fonte: Pereira (2021, p. 192)

A criação desse quadro é a base do instrumental metodológico para aplicação em pesquisas acerca de fenômenos linguísticos rurais. O objetivo é utilizá-lo para demarcar o grau de ruralidade em que cada informante da pesquisa estaria, conforme o mapeamento de uso dessas características. A partir do parâmetro desse quadro, são arquitetados três graus de ruralidade linguística: *baixo* (com três ou menos características mapeadas); *médio* (de três a cinco); e *alto* (de cinco a seis).

Assim, na oitiva das entrevistas, havendo a presença de um vocábulo que contemple cada fenômeno linguístico ali elencado, é preciso anotar no quadro a sua ocorrência. Ao fim da entrevista, o pesquisador contabiliza quantos fenômenos rurais foram mapeados e classifica o informante conforme o 'grau de ruralidade' alcançado.

Depois dessa classificação, esse detalhamento é incorporado à codificação das variáveis independentes linguísticas da análise sociolinguística, a qual, por sua vez, passa pelo programa GOLDVARB numa análise multivariada, de acordo com o que se tem feito no Brasil nas pesquisas pautadas na Teoria da Variação e da Mudança.

Com isso, o objetivo é que o 'grau de ruralidade' atue como uma variável independente de caráter mais qualitativo, pois é preciso ouvir e descrever minuciosamente as entrevistas sociolinguísticas para captar o fenômeno que constitui a variável dependente (neste caso, alternância vocálica precedente a /mos/) que é foco de pesquisa, mas também ter ouvidos atentos às seis peculiaridades linguísticas complementares à pesquisa, componentes do quadro.

Passemos, pois, aos resultados obtidos com o uso deste instrumental metodológico do 'grau de ruralidade' em aplicação à variação vocálica precedente a /mos/.

3.2 Mapeando resultados: a aplicação da variável 'grau de ruralidade' no estudo da alternância vocálica precedente a /mos/ em verbos de primeira e segunda conjugação

Antes da aplicação da variável em questão, cabe explicitar as características básicas do banco VARLINFE (Variação na fala eslava), pertencente à UNICENTRO (câmpus Irati), que constituiu a composição de nossa amostra. A característica principal do banco é a sua particularidade eslava, isto é, os informantes conheciam, falavam e/ou escreviam em algum grau, as línguas polonesa e/ou ucraniana.

Para além disso, trata-se de um banco de característica rural composto por 168 entrevistas de aproximadamente 1 hora. Os entrevistadores utilizaram um roteiro de perguntas e gravador, tendo se deslocado para comunidades rurais de difícil acesso, bem como relativamente distante de áreas urbanas (cf. Costa; Loregian-Penkal, 2015), o que revela, já neste aspecto, o nível de *isolamento* das comunidades entrevistadas. O banco apresenta 7 municípios e 24 entrevistas para cada um deles, sendo estratificado confor-

me os padrões na área da sociolinguística no Brasil, de modo similar ao banco-base do VARSUL.

Das 168 entrevistas, foram coletados 1.858 dados, do fenômeno linguístico da alternância vocálica sendo que somente 227 ocorrências foram de 2ª conjugação, ao passo que 1.631 foram de 1ª conjugação. A discrepância considerável em relação ao número de ocorrências (88% do córpus em 1ª conjugação, e 12% em segunda) se dá pela produtividade mais abundante dos verbos de 1ª conjugação na língua portuguesa

A fim de detalharmos melhor os resultados de pesquisa e também considerando como foram feitas as rodadas estatísticas no programa GOLDVARB-X, apresentamos primeiramente os resultados de primeira conjugação (com as variantes /a/ ~ /e/) e, na sequência, os resultados de segunda conjugação (com as variantes /e/ ~ /i/).

Em ambas as conjugações, a variável 'grau de ruralidade' foi selecionada como significativa no que se refere ao fenômeno da alternância vocálica, revelando sua importância no estudo do fenômeno variável rural citado. Portanto, acreditamos que tal variável possa ser bastante útil no estudo de outros fenômenos linguísticos rurais/conservadores.

3.2.1 Análise de 1ª conjugação: as variantes canônica /a/ e não canônica /e/ em foco

Dos 1.631 dados de 1ª conjugação coletados, houve 900 ocorrências da variante canônica /a/ (55,2%) e 731 da variante não canônica /e/ (44,8%), mostrando que a variante canônica ainda prepondera na comunidade linguística, embora o uso não canônico seja bastante considerável.

Sobre a variável independente 'grau de ruralidade', o programa GOLDVARB-X selecionou-a na segunda posição, perdendo apenas para a variável 'TMA' (tempo-modo-aspecto), o que demonstrou sua altíssima significância.

Tabela 1: Efeito da variável 'grau de ruralidade' sobre a aplicação do fenômeno morfofonológico da alternância vocálica de /a/ para /e/

VARIÁVEL	**Categorias**	**Apl./tot.**	**%**	**PR**
Grau de ruralidade	Baixo	27/155	17.4%	0.155
	Médio.	225/570	39.5%	0.379
	Alto	479/906	52.9%	0.646
	TOTAL	731/1631	44,8%	

Input: 0.250
Log likelihood = -446.771
Significance = 0.011

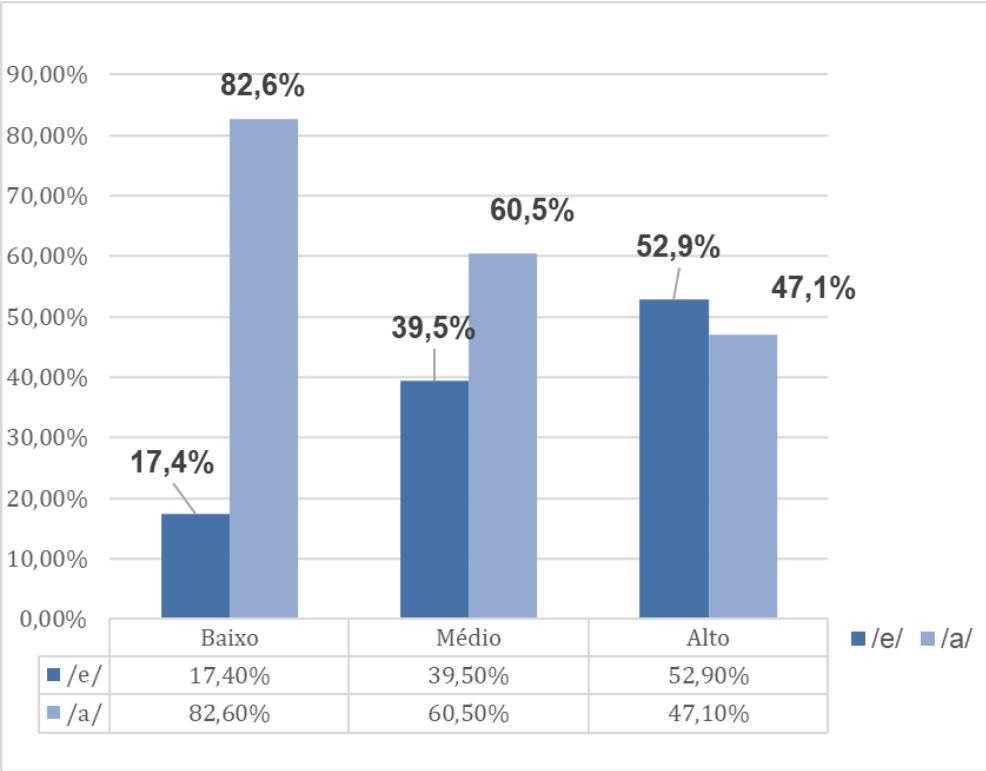
Fonte: Pereira (2021, p. 343)

As 479 ocorrências da variante não padrão /e/ em relação ao grau alto de ruralidade, comparando-se ao número baixíssimo de apenas 27 ocorrências de tal variante nos casos de grau baixo de ruralidade já revelam que existe uma forte tendência dos informantes com grau alto utilizarem formas verbais não padrão como *cantemo*, *falemo*, entre outros. Porém, conforme o grau de ruralidade diminui, o uso dessas formas também se torna ínfimo, mostrando que o 'grau de ruralidade' tem forte influência para o uso não canônico da alternância vocálica.

Além disso, o peso de 0.646 associado ao grau alto de ruralidade deixa ainda mais evidente a significância para o uso de /emo/ na primeira conjugação, da mesma forma como os pesos relativos de 0.155 e 0.379 estarem relacionados, respectivamente, aos graus baixo e médio evidenciam que os falantes com grau mais baixo ainda optam majoritariamente pelos usos mais padrão, prestigiados e menos rurais.

O gráfico a seguir mostra mais claramente a curva ascendente do uso de *emo* conforme cresce o grau de ruralidade do informante:

Gráfico 1: Porcentagens das variantes /a/ e /e/ conforme 'grau de ruralidade' em 1ª conjugação



Fonte: Pereira (2021, p. 344)

Tendo sido expostos os resultados quantitativos a que chegamos, fazemos uma análise qualitativa a partir do 'grau de ruralidade'. Um perscrutamento mais fino dos dados mostrou o seguinte: os informantes que mais produziram dados não canônicos, os quais são considerados rurais, foram justamente os que apresentaram maior grau de ruralidade. Isso pode ser visto no quadro a seguir:

Quadro 2: Dados de características da variedade rural com base em um informante com 'grau de ruralidade' alto

Fenômenos linguísticos	Exemplo	Utilizou?
Vocalização da lateral palatal [ʎ]	Palhaço – *paiaço*	*Famia, oiada, baruio, trabaioso, veia, joeio, mio, assoaio, muié, coeio*
Nasalização da vogal que não antecede segmentos nasais	Igreja – *ingreja*	*Inducado* *Indaí (e daí), ingreja, inguar (igual)*
Rotacismo (permuta de /l/ para /r/)	Planta – *pranta*	*Argum, arguma, prantava, cumpricado, facurdade*
Permuta de fricativa labiodental sonora [v] em oclusiva bilabial sonora [b] e vice-versa	Travesseiro – *trabissero*	*Venzê, venzido*
Prótese do -a- em vocábulos que iniciam por consoante ou aférese (queda) do -a-	Divertir – *adivertir*	*adirubá* (*derrubar*, sem africada)
Rótico pertencente ao rol dos retroflexos em coda e onset silábicos – tepe retroflexo [ɽ] e aproximante retroflexo [ɻ]	Cor – ['koɽ]	*Forte, moderna, governo*

Fonte: Pereira (2021, p. 346)

O informante em questão era da cidade de Irati, de gênero masculino, com ensino médio completo e mais jovem, tendo apresentado alto grau de ruralidade justamente por contemplar, com variados exemplos, todas as 6 características rurais que constam no quadro de ruralidade. Já no que diz respeito ao fenômeno da alternância vocálica, o sujeito entrevistado produziu 20 dados, sendo apenas 2 deles com a variante canônica, ou seja, apenas 10% de uso da variante padrão.

Isso pode ser confirmado em trechos de sua fala:

(9) [sobre uma história de infância quando colocaram uma lata com fogo dentro da sala de aula para se esquentar] Entrevistador: Mais e daí? Fizerum o que ca- ca lata? **Entrevistado:** Puis i daí fomo [inint] **taqu[e]mo** água **apagu[e]mo** as brasa i **jogu[e]mo** pra fora, né? Vai fazê o que co aquilo lá, né? **(IRA, mas, ens. médio, mais jovem)**

(10) Daí peguei comprei uns negócio de boracharia e fiquei daí. [est] Daí até **pens[e]mo** de- de- de- de arumá outro serviço e saí, ma daí, ah, já me estacionei por aqui e... **(IRA, mas, ens. médio, mais jovem)**

(11) Daí já o piá-, o piá já começô a trabaiá em Irati, né? [est] Aí num **procur[e]mo** outra coisa. **(IRA, mas, ens. médio, mais jovem)**

Além do uso considerável da variante não canônica e dos fenômenos linguísticos presentes no quadro que propusemos, há outros elementos a se ressaltar, como o uso de tepe em "borracharia", o alçamento vocálico de /o/ em "pois", além do item lexical "piá", que é bastante comum à realidade linguística do sul paranaense. Outro ponto a se destacar e que não pode ser deixado de lado é que tal informante apresenta grau de escolaridade III (Ensino Médio), e isso revela que, ao tratarmos de comunidades rurais, considerar apenas a variável 'escolaridade' para averiguar o uso de formas canônicas e não canônicas deve ser questionado.

Algumas informações extras também se fazem necessárias, como: outro informante desta mesma cidade com grau bastante alto produziu 19 dados, sendo, dentre eles, 17 da variante não canônica /e/ e apenas 2 da variante canônica /a/. Com base neste levantamento, podemos perceber que o uso do 'grau de ruralidade' possibilita mostrar caminhos cada vez mais nítidos para a identificação de variantes rurais e conservadoras. Um resultado similar pôde ser verificado nos resultados de 2ª conjugação, conforme descrito na subseção a seguir.

3.2.2 Análise de 2ª conjugação: as variantes canônica /e/ e não canônica /i/ em foco

Dos 227 dados coletados na 2ª conjugação, houve 110 ocorrências da variante canônica /e/ (48,5%) e 117 da variante não canônica /i/ (51,5%), ou seja, o uso de /imo/ foi preponderante nesta comunidade linguística, revelando a forte tendência deste uso de cunho rural.

O conjunto de dados a seguir mostra um pouco dessa realidade linguística das localidades investigadas:

(12) [sobre não terem tido aula naquele ano] **Perd[i]mo** otro ano daí. **(RIO, mas, fund. I, mais jovem).**

(13) Entrevistadora: E como é que a senhora ia então? Entrevistada: Di a pé! A pezinho... Nóis levantava seis horas e tava de a pezito lá na coloi. Não fartava aula e... uma veiz nós **perd[i]mo** essa vez foi... hoje em dia não... cinco minuto nóis atrasemo, nóis vortemo de Gonçalves Júnio pra num entrá que tinha vergonha de entrá, que nóis tava atrasado cinco minuto na iscola. **(IRA, fem, fund. II, mais jovem)**

(14) Entrevistadora: E a senhora não sabe lê? Entrevistada: Em ucraniano? Sei. Entrevistadora: Sabe lê também? Entrevistada: Sei. Entrevistadora: Ahm...E a senhora aprendeu em casa também? Entrevistada: É... **Aprend[i]mo** na escola. Entrevistadora: Sim. Lá com as irmãs, né? Entrevistada: Lê e escrevê, não assim tão bem, que nem ainda, até esses dia nóis tava conversando ali cas mulher da igreja. Pra lê assim alto, eu me apuro, eu gaguejo, sabe? Eu leio pra mim. Não pronunciando as palavras... eu leio tudo, é, um jornal eu leio, mas, se for pra mim pronunciá as palavras, daí a gente já se perde um poco. **(IRA, fem, fund. I, mais velha)**

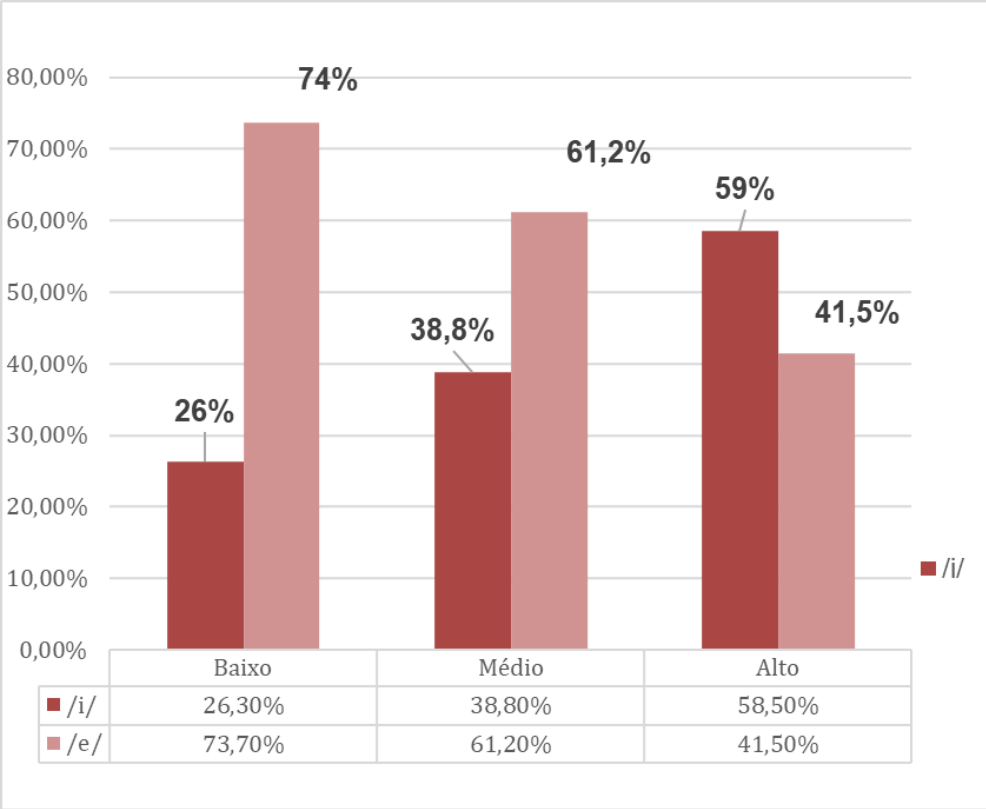
(15) Antes de aprendê o português, nóis sabia rezá em ucraniano. Depois, **aprend[i]mo** a rezá em português, nunca mais [hes] rezamo em ucraniano. **(CRU, fem, ensino médio, mais velha)**

(16) Entrevistadora: E, assim, e de comida? O que que a sua vó, a sua mãe fazia que é, assim, típico ucraniano? Entrevistada: Nóis, que nóis **com[i]mo,** que é o borsht, que é ucraniano, né? Entrevistadora: É uma sopa? Entrevistada: É uma sopa. **(IRA, fem, fund. II, mais jovem)**

Os trechos de fala não apenas trazem à vista os usos não canônicos, mas também evidenciam a relação de conservação linguística das línguas eslavas, mostrando, mais uma vez, a faceta conservadora da comunidade em questão. Ou seja, assim como há uma manutenção da língua eslava entre os informantes, também se mantém uma variedade da língua portuguesa rural.

A respeito da variável independente 'grau de ruralidade', o programa GOLDVARB a selecionou na quarta posição, o que mostra seu grau de significância. O gráfico 2 mostra os resultados gerais obtidos:

Gráfico 2: Porcentagens das variantes /e/ e /i/ conforme 'grau de ruralidade' em 2ª conjugação



Fonte: Pereira (2021, p. 381)

A mesma curva de ascendência ocorrida em relação à primeira conjugação ocorre agora na 2ª conjugação, quando o uso de /i/, que é a vogal não padrão, cresce de acordo com o crescimento do grau de ruralidade. Do mesmo modo, há uma curva descendente ao mostrar o uso da variante padrão /e/ quando grau de ruralidade cresce. No 'grau de ruralidade' baixo, há o uso considerável de 73,7% (14 ocorrências) da variante canônica /e/, com apenas 26,3% (5 ocorrências) da não canônica /i/. Já o número de uso não canônico de /i/ aumenta no grau médio, quando há um número de 38,8% (33 ocorrências) de uso, em oposição a 61,2% (52 ocorrências) da variante canônica.

A categoria de destaque no que diz respeito à nossa regra variável é o grau alto de ruralidade, em que se mostra uma porcentagem de 58,5% (72 ocorrências) de formas com a variante /i/, em contraposição a 41,5% (51 dados) da variante canônica /e/.

Com isso, assim como o que foi revelado na alternância vocálica de /a/~/e/, o grau de ruralidade alto aparece como o condicionador de uso da forma não canônica. Essa revelação se confirma com a geração dos pesos relativos na variação de /e/~/i/.

Tabela 2: Efeito da variável 'grau de ruralidade' sobre a aplicação do fenômeno morfofonológico da alternância vocálica de /e/ para /i/

VARIÁVEL	**Categorias**	**Apl./ tot.**	**%**	**PR**
Grau de ruralidade	Baixo	5/14	26.3%	0.438
	Médio.	33/85	38.8%	0.350
	Alto	72/123	58.5%	0.615
	TOTAL	110	48,5%	

Input: 0.428
Log likelihood = -88.509
Significance = 0.035

Fonte: Pereira (2021, p. 382)

O peso relativo de 0.615 em relação ao grau alto exibe o favorecimento dessa categoria relativamente ao uso de formas como *aprendimo*, *comimo*, *perdimo*, *vencimo*, *corrimo*, *sofrimo*, *conhecimo*, entre outras que apareceram em nosso córpus. Por outro lado, os outros pesos mais baixos evidenciam o desfavorecimento do uso não canônico nos graus de ruralidade baixo e médio.

Enfim, esses resultados deixam evidente que os usos não canônicos de /e/ em primeira conjugação e de /i/ em segunda conjugação estão altamente relacionados ao alto grau de ruralidade, constituindo-se em uma possível prova de que o uso de /emo/ e /imo/ se trate de uma característica rural.

Mas, para além disso, a relação dessa alternância com outros fenômenos linguísticos considerados rurais e descritos como surgidos num português menos recente podem dar indícios de que essas comunidades linguísticas caminham para um processo de *manutenção linguística* ou *conservadorismo linguístico*, considerando que muitos desses usos já têm sidos suplantados em comunidades urbanas, nas quais algumas mudanças linguísticas se efetuaram, mas continuam "firmes e fortes" (além de muito produtivos) em comunidades rurais no Brasil, como se pôde verificar nessas comunidades linguísticas do banco VARLINFE.

5 Considerações finais

Os dados sobre alternância vocálica obtidos nesta pesquisa revelaram suas inter-relações entre conservadorismo linguístico e 'grau de ruralidade', uma vez que a produção de formas não canônicas (de /emo/ e /imo/) ocorreram preponderantemente entre in-

formantes cujo grau de ruralidade era mais alto. Além disso, a parte quantitativa da pesquisa mostrou uma curva ascendente de uso das variantes não padrão de acordo com o aumento do grau de ruralidade, tanto em 1ª quanto em 2ª conjugação.

No tangente ao questionamento **(i)** – acerca da possibilidade de relacionar a variável 'grau de ruralidade' a "ilhas de conservadorismo linguístico" – alicerçamo-nos em pesquisadores que já estabeleceram conexões entre o âmbito rural e a conservação de fenômenos linguísticos menos recentes, como Naro e Scherre (2003), Ribeiro (2013) e Ribeiro e Lacerda (2013), para asseverar que há uma concatenação entre esses dois aspectos, ao passo que, por outro lado, as variedades cultas, *grosso modo*, estariam mais ligadas à escolaridade mais alta e, portanto, usos mais padrão, além de estarem mais abertos a inovações linguísticas.

A respeito da questão **(ii)** – sobre *como* e *por que* o fenômeno da alternância vocálica em formas verbais que contêm /emo/ e /imo/ tem atravessado séculos e continua sendo encontrado em variedades brasileiras –, acreditamos que isso ocorra devido à manutenção linguística que ocorre em "ilhas de conservadorismo linguístico" existentes em localidades rurais no Brasil. Nesses lugares, as variações e fenômenos linguísticos ocorridas num português menos recente permanecem entre os falantes, de modo a manter um falar menos moderno e mais conservador.

Contudo, em relação aos porquês da manutenção linguística de certas variantes não padrão, esta pesquisa ainda não consegue trazer respostas, não obstante possamos supor que a questão da identidade desses sujeitos possa ser um aspecto importante a se considerar para explicar tal ponto.

Por fim, no que se refere à questão **(iii)** – a respeito de uma descrição linguística mais ampla num estudo quantitativo que possa auxiliar o pesquisador a classificar um fenômeno linguístico como rural –, propusemos o uso da variável linguística complexa 'grau de ruralidade', a partir da qual é possível analisar a fala de informantes com mais detalhamento, constatando se são falantes mais (ou menos) adeptos de uma variedade rural.

Partindo da premissa de que o conservadorismo linguístico está no contrafluxo da mudança linguística, queremos fechar este texto trazendo algumas reflexões para pesquisas futuras:

(a) Se a ruralidade está intrinsecamente ligada ao conservadorismo linguístico, o qual, por sua vez, está no contrafluxo da mudança linguística, é possível afirmar que comunidades rurais são mais resistentes a mudanças linguísticas?

(b) Se mudanças linguísticas ocorrem mesmo em realidades em que há resistência para

o uso de novas variantes e para a mudança linguística, quais são os aspectos/ as variáveis que influenciariam na aceitação dos falantes de localidades rurais em relação a novas variantes linguísticas e a efetivação da mudança?

Almejamos, ao findar deste artigo, que a variável 'grau de ruralidade' possa surgir como um instrumento metodológico relevante para pesquisas na área de sociolinguística quantitativa respectivas a fenômenos linguísticos conservadores, levando-se em conta também seu aspecto qualitativo e seus possíveis direcionamentos de análise.

Referências

AMARAL, A. *O dialeto caipira*. 3.ed. São Paulo: Hucitec, 1976 [1920].

BORTONI-RICARDO, S. M. *Nós cheguemu na escola e agora?* Sociolinguística e educação. São Paulo: Parábola Editorial, 2005.

BORTONI-RICARDO, S. M. *Do campo para a cidade:* estudo sociolinguístico de migração e redes sociais. São Paulo: Parábola Editorial, 2011.

BORTONI-RICARDO, Stella Maris. A concordância verbal em português: um estudo de sua significação social. In: VOTRE, Sebastião; RONCARATI, Cláudia (org.). *Anthony Julius Maro e a Linguística no Brasil*: um abordagem acadêmica. Rio de Janeiro: 7Letras, 2008, p. 362-380.

BORTONI-RICARDO, S. M; VELLASCO, A. M. de M. S.; FREITAS, V. A. de L. *O falar candango:* análise sociolinguística dos processos de difusão e focalização dialetais*:* estudo sociolinguístico de migração e redes sociais. Brasília: Editora UnB, 2010.

CARDEIRA, Esperança. *O essencial sobre a história do português.* Lisboa: Caminho, 2006

CASTILHO, Ataliba T. de. *Introdução ao estudo do aspecto verbal na língua portuguesa*. Marília - SP, Faculdade de Filosofia, Ciências e Letras. Tese de Doutorado, 1968.

CASTILHO, A. T. de. O Português do Brasil. In: ILARI, Rodolfo. *Linguística Românica.* São Paulo, Ática, p. 237-285, 1992.

CASTILHO, Ataliba T. de. *Nova gramática do português brasileiro*. São Paulo: Contexto, 2016.

COSTA, I. B. *O verbo na fala de camponeses*: um estudo de variação. Tese de Doutorado. Campinas: Universidade de Campinas, 1990.

ELIA, Sílvio. *Ensaios de Filologia Linguística*. Rio de Janeiro: Grifo, 1975.

FOEGER, Camila Candeias; YACOVENCO, Lilian Coutinho; SCHERRE, Maria Marta Pereira. A primeira pessoa do plural em Santa Leopoldina/ES: correlação entre alternância e concordância *Letrônica*. Porto Alegre, v. 10, n. 1, p. 5-17, janeiro-junho, 2017.

FROSI, V.; MIORANZA, C. *Dialetos italianos*. Caxias do Sul: EDUCS, 1983.

LABOV, W. *Padrões sociolinguísticos*. Tradução de Marcos Bagno, Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].

LABOV, W. *Language in the Inner City*: The Black English Vernacular. Philadelphia: University of Pennsylvania Press. 255-92, 1972.

LABOV, W. Building on Empirical Foundations. In: LEHMANN, W.P.; MALKIEL, Y. *Perspectives on historical Linguistics* (eds.). Amsterdam/Philadelphia: Johns Benjamin Publishing Company, 1982.

LABOV, W. *Principles of linguistic change*: Internal factors. Cambridge: B. Blackwell, 1994.

LABOV, W. *Principles of linguistic change* – Social factors. Cambridge: B. Blackwell, 2001.

LABOV, W. *The social stratification of English in New York City*. 2. ed. New York: Cambridge University Press, 2006a.

MILROY, L. (1992). *Linguistic variation and change*. On the historical sociolinguistics of

English. GB: Brasil Blackwell.

NARO, A. J.; GÖRSKI, E.; FERNANDES, E. Change without change. *Language Variation and Change*, v. 11, n. 2, p. 197-211, 1999.

NARO, Anthony Julius; SCHERRE, Maria Marta Pereira. *Origens do português brasileiro*. São Paulo: Parábola Editorial, 2007.

NARO, A. J.; SCHERRE, M. M. P. Variação e mudança linguística: fluxos e contrafluxos na comunidade de fala. *Cadernos de Estudos Lingüísticos*, Campinas, SP, v. 20, p. 9-16, 2012. Disponível em: https://periodicos.sbu.unicamp.br/ojs/index.php/cel/article/view/8636853. Acesso em: 10 out. 2020.

PEREIRA, Ivelã. Cuidamo(s) e cuidemo(s): a variação morfêmica na p4 em verbos regulares de 1ª conjugação. *Working Papers em Linguística*, v. 2, n. 14, p. 49-71, Florianópolis, ago/dez. 2014.

PEREIRA, Ivelã. A voz de polono-brasileiros: um contexto histórico sul-paranaense. *Working Papers em Linguística*. Florianópolis, 18(1): 23-45, jan./jul., 2017.

PEREIRA, Ivelã. O caso de –a-mo(s) versus –e-mo(s) e –e-mo(s) versus –i-mo(s): variação morfêmica ou especialização temporal? *Caderno Seminal Digital Especial*, nº 1, v. 1, jan-dez, 2018.

PEREIRA, Ivelã. *"O pirogue, nós aprendimo da mãe" e "agora nós mudemo o borscht"*: variação morfofonológica em comunidades rurais eslavo-brasileiras no sudeste do Paraná. Tese (doutorado). Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão, Programa de Pós-Graduação em Linguística, Florianópolis, 2021.

PEREIRA, Ivelã; LEHMKUHL-COELHO, Izete; LOREGIAN-PENKAL, Loremi. Variação na concordância Verbal de nós no presente e pretérito perfeito em Verbos regulares de 1ª e 2ª conjugação: produtiva no sudeste paranaense? *Signótica*. Goiânia, v. 28, n. 2, p. 481-508, jul./dez. 2016.

PEREIRA, Ivelã; MARGOTTI, Felício Wessling. Sobre onde nós fiquemo: mapeamento diatópico de um traço rural brasileiro. *Web-Revista SOCIODIALETO* – NUPESDD / LALIMU, v. 8, nº 24, mar, 2018.

VANDRESEN, Paulino. *Fonologia do vestfaliano de Rio Fortuna*. Dissertação de Mestrado. (Programa de Pós-Graduação em Lingüística). Universidade Federal do Rio de Janeiro, Rio de Janeiro: UFRJ, 1968.

VANDRESEN, Paulino. Línguas em contato: um panorama da pesquisa no Brasil. In: SAVEDRA BARRETO, M.M.G.; SALGADO, A.C.P. (Orgs.) *Sociolinguística no Brasil*: uma contribuição dos estudos sobre línguas em/de contato. Homenagem ao prof. Jürgen Heye. Rio de Janeiro: 7Letras, 2009, p. 6-16.

VASCONCELOS, J. L. de. *Esquisse d'une dialectologie portugaise*. Lisboa: Centro de Estudos Filológicos, 1970 [1901].

WEINER, J.; LABOV, W. Constraints on the agentless passive. *Journal of Linguistics* 19, 1983.

WEINREICH, Uriel. *Languages in contact*: findings and problems. New York: De Gruiter, 1953.

WEINREICH, U; LABOV, W; HERZOG, M. Empirical foundations for a theory of language change. In: W. LEHMANN; Y. MALKIEL (eds.). *Directions for Historical Linguistics*. Austin: University of Texas Press, 1968.

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

ZILLES, A. M. S.; MAYA, L.; SILVA, K. A concordância verbal com a primeira pessoa do plural em Panambi e Porto Alegre, RS. *Organon,* Porto Alegre, v.14, n.28/29, p.195-219, 2000.

ZILLES, A. M.; BATISTA, H. H. A concordância verbal na primeira pessoa do plural na fala culta de Porto Alegre. In: VANDRESEN, Paulino (org.). *Variação, mudança e contato linguístico no português da região sul*. Pelotas: EDUCAT, p. 100-124, 2006.

RIBEIRO, P. R. O. (2013). *O perfil sociolinguístico do município de Oliveira Fortes-MG: a concordância nominal e verbal*. Dissertação de Mestrado. Juiz de Fora: Universidade Federal de Juiz de Fora.

RIBEIRO, Patrícia Rafaela Otoni; LACERDA, Patrícia Fabiane Amaral da Cunha. Variação, Mudança e não mudança linguística: ressignificando o conservadorismo linguístico no português do Brasil. *Revista LinguíStica* / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro. Volume 9, número 2, dezembro de 2013. ISSN 1808-835X 1. http://www. letras.ufrj.br/poslinguistica/revistalinguistica

http://doi.org/10.5007/1984-8420.2025.e101203

A (RE)EXISTÊNCIA DOS POVOS INDÍGENAS NAS ENTRANHAS DO SERTÕES BAIANOS: UMA BREVE HISTÓRIA SOCIAL LINGUÍSTICA

THE (RE)EXISTENCE OF INDIGENOUS PEOPLES IN THE DEPTHS OF BAHIA'S BACKLANDS: A BRIEF SOCIAL LINGUISTIC HISTORY

Paloma Reis Soares | Lattes | paloma.aleoncio@gmail.com
Universidade Estadual de Feira de Santana

Edson Tosta Matarezio Filho | Lattes | etmfilho@uefs.br
Universidade Estadual de Feira de Santana

Resumo: Este artigo tem como objetivo investigar as dinâmicas sócio-históricas e linguísticas dos povos indígenas nos sertões-norte da Bahia. A pesquisa segue uma abordagem bibliográfica e descritiva, fundamentada nos princípios da História Social Linguística, ramo da Linguística Histórica. A metodologia utilizada envolve a análise de fontes bibliográficas, com o propósito de explorar o processo histórico e de compreender a formação dos sertões do semiárido baiano, considerando os atravessamentos históricos desde o período colonial até suas configurações atuais. O impacto violento do processo de colonização resultou na supressão e no apagamento dos elementos históricos e socioculturais dos povos originários, evidenciando a diminuição da diversidade linguística e os danos causados às línguas indígenas. Este estudo almeja compreender a resistência, a organização e o legado dos povos indígenas nos semiáridos baianos, considerando a relevância das línguas pertencentes à família kariri e a imperatividade de salvaguardar a diversidade linguística e cultural que a caracteriza. A resistência dos povos indígenas é reconhecida como um componente fundamental na configuração étnica e linguística da Bahia e, por extensão, do Brasil. Entretanto, enfatiza-se que ainda há muito a ser investigado e protegido, incumbindo-nos a tarefa contínua de conhecer e de respeitar a herança linguística, histórica e cultural dos povos indígenas.

Palavras-chave: História Social linguística. Povos indígenas. Colonização. Catequização. Sertões baianos.

Abstract: This article aims to investigate the socio-historical and linguistic dynamics of Indigenous peoples in the north backlands of Bahia. The research adopts a bibliographic and descriptive approach, grounded in the principles of Social Linguistic History, a branch of Historical Linguistics. The methodology employed involves the analysis of bibliographic sources to explore the historical process and understand the formation of the backlands of semi-arid Bahia, taking into account the historical intersections from the colonial period to their current configurations. The violent impact of the colonization process led to the suppression and erasure of the historical and sociocultural elements of Indigenous peoples, highlighting the decline in linguistic diversity and the damage inflicted on Indigenous languages. This study aims to understand the resistance, organization, and legacy of Indigenous peoples in the Bahian semi-arid region, considering the significance of the languages belonging to the Kariri family and the imperative to safeguard the linguistic and cultural diversity that characterizes them. However, it is important to emphasize that there is still much to be investigated and protected, and it is our ongoing responsibility to explore and respect Indigenous peoples' linguistic, historical, and cultural heritage.

Keywords: Social linguistic history. Indigenous peoples. Colonization. Catechization. Bahia's backlands.

1 INTRODUÇÃO

Desbravar os sertões é seguir os caminhos que nos levam à história soterrada dos povos originários. Ao despertar a memória do ambiente que guarda a raiz violenta do processo de colonização, responsável por dizimar, por doutrinar e por apagar os aspectos históricos e socioculturais dos povos indígenas, confrontamos a manifestação linguística de um país político e ideologicamente fragmentado. No Brasil, um país que vive um paradoxo linguístico originado pelos mecanismos coloniais, nota-se, de um polo, a hegemonia da cultura e da Língua Portuguesa, falada por 98% da sua população e, do outro, uma das maiores diversidades de línguas, em sua maioria, oriundas das mais diversas etnias indígenas, somando-se mais ou menos 180 línguas (Simas e Lucchesi, 2020). Esse complexo cenário nos revela o intenso massacre dos povos originários visto que, segundo Franchetto e Balykova (2020) "no decorrer do tempo colonial, foram apagadas pelo menos 80% das línguas dos povos nativos", essa significativa depreciação gerou danos inestimáveis para as línguas indígenas, em especial, para aquelas que primordialmente atuavam, desde o século XVI, nos sertões do estado baiano.

Diante desse caminho, o referente artigo debruça-se quanto aos aspectos da História Social e linguística dos povos indígenas dos sertões baianos, precisamente da família Kariri, atribuída ao tronco macro-jê, a fim de compreender de forma breve a resistência, a organização e o legado dos povos originários no semiárido baiano, disposto a visualizar como se comporta na atualidade e, através deste, traçar perspectivas para a preservação das suas narrativas socioculturais e manifestação das suas línguas maternas.

Por intermédio de uma pesquisa bibliográfica, busca-se explanar o processo de colonização e de catequização dos povos indígenas no sertão baiano, durante o século XVII e XVIII, e retratar as modificações que esse processo ocasionou nos aspectos linguísticos da família Kariri, fornecendo uma atenção especial quanto às prerrogativas históricas com interesse de entender como se estabelece essa família no tempo atual. Para a realização desta pesquisa, considera-se o campo da Linguística Histórica, a qual se atenta para a mudança linguística consoante ao tempo histórico, com ênfase nos fatores estruturais e sócio-históricos, vertente *stricto sensu* (Mattos e Silva, 2008).

É preciso retomar a história da organização das primeiras comunidades indígenas, desse modo, entenderemos os seus conflitos e a sua configuração atual, assim como a sua participação histórica e linguística para a formação social da Bahia. No interior do estado, ocorreram os primeiros aldeamentos indígenas com a intenção de converter e de "civilizar" esses povos (Santos, 2012, p. 44). Assim, nos interessa compreender os desafios e a atuação das populações indígenas provenientes da família linguística Kariri, que povoaram e, ainda povoam, os semiáridos baianos. É notório que esse contato secular com o estado resguarda a participação dos povos originários nos aspectos culturais e, precisamente, linguísticos da região e, por esse viés, sobressalta-se neste trabalho o protagonismo dos referidos povos quanto à sua formação.

Com a intencionalidade de abarcar as discussões aqui citadas, a abordagem será dividida em quatro aspectos. Em *Do Grão-Pará à Bahia: um breve panorama histórico das políticas linguísticas coloniais,* será retratado um breve panorama histórico de como se sucedeu o processo colonial e a política de catequização jesuíta referente aos povos indígenas aldeados, e como isso afetou os povos do sertão baiano[1]. Após apresentar essa parte histórica, será necessário compreender os atravessamentos políticos, sociais e ideológicos que o sistema colonial depositou sobre a existência das línguas indígenas do tronco macro-jê, assim como sobre os seus falantes, a fim de vislumbrar a sua resistência, dessa forma com-

[1] Para compreender com mais detalhamento esse processo histórico recomendo a leitura em: SANTOS, Fabricio Lyrio. **Da catequese à civilização:** colonização e povos indígenas na Bahia (1750-1800) – Salvador/BA: Universidade Federal da Bahia / Faculdade de Filosofia e Ciências Humanas, 2012.

põe-se a seguinte seção, *Os "tapuias" do sertão baiano e as línguas do tronco macro-jê*. Isto posto, para debruçar-se sobre a análise de como se comportam na atualidade as línguas da família Kariri e quais as discussões pertinentes realizadas por pesquisadores nos últimos anos, em *A (re)existência do macro-jê: tronco, famílias e línguas no interior da Bahia*, poder-se-á visualizar caminhos que trilhem a revitalização, a retomada, a constituição histórica e a política da identidade étnica desses povos, além de estimular os holofotes da pesquisa quanto a esse tema. Por fim, nas considerações finais, serão retomados alguns pontos já discutidos, destacando-se a importância de lançar luz sobre a história, a língua e os aspectos socioculturais do tronco macro-jê. Isso se faz especialmente relevante devido ao grave processo de redução da multiculturalidade e do multilinguismo desse tronco, promovido pelo sistema colonialista, que resultou na escassez de documentos sobre ele.

2 DO GRÃO-PARÁ À BAHIA: UM BREVE PANORAMA HISTÓRICO DAS POLÍTICAS LINGUÍSTICAS COLONIAIS

A Companhia de Jesus teve o seu primeiro contato com o solo brasileiro em 1549, por meio da comitiva do primeiro governador geral Tomé de Sousa. Desde a sua chegada, a intenção dos jesuítas era converter os nativos ao cristianismo e, para isso, foi necessário estabelecer uma política de catequização de acordo com a realidade cultural dessa nova nação. Em vista disso, a base da conversão dos povos originários foi por meio da realização do aldeamento, ou seja, agrupamentos de povos indígenas em um determinado território, os quais se constituíam a partir de três aspectos: construção de uma igreja em uma comunidade compreendida como indígena ou em uma "légua em quadra" fornecida pelo governo da época, a partir da qual se estabelecia a aldeia; catequização, instrução da doutrina cristã seguida de batismo; e por último, o deslocamento de grupos indígenas, em geral das brenhas do sertão para regiões com povoações estruturadas, onde a transmissão dos dogmas cristãos seria melhor efetivada, esse processo é chamando de "descimento". (Santos, 2012, p. 44).

> Os religiosos da Companhia de Jesus foram os primeiros que adotaram a prática de aldear ou reunir os índios com o objetivo de torná-los cristãos, sendo responsáveis pela criação das primeiras aldeias ou aldeamentos situados nas proximidades da cidade de Salvador, no século XVI, a partir do empenho conjunto do padre Manoel da Nóbrega e do terceiro governador-geral da colônia, Mem de Sá (Santos, 2012, p. 43).

Aldear os indígenas em um determinado espaço permitia que os missionários

propagassem o dogma cristão e se introduzissem no cotidiano indígena e, por esse viés, aprendiam sobre os aspectos socioculturais e linguísticos desses povos e, precisamente, utilizavam essa política de conversão como forma de controle dos seus costumes fundamentais, os quais pareciam hediondos aos olhos do homem europeu, cristão e colonizador. O nomadismo era um dos hábitos indígenas que contrastava com a cultura do aldeamento, já que esse propagava a ideia de se fixar em algum lugar por um longo período, mas, para além desse aspecto, a antropofagia, a poligamia, o ato de guerrear e a devoção aos seus encantados se chocavam diretamente com a noção de civilização advinda dos jesuítas (Santos, 2012).

A organização da vida social dentro das aldeias corroborava com a imposição de um novo conceito de mundo aos povos aldeados, que destituía incisivamente o *modus vivendi* indígena em prol dos aspectos doutrinais do cristianismo, demonstrando, assim, o quão a percepção colonial e cristã via a cultura dos povos originários como imoral, desqualificando-a, marginalizando-a, em relação a cultura da coroa portuguesa, que se alocava numa posição de prestígio social. A percepção dos portugueses sobre esses povos foi profundamente marcada pelas ambivalências da alteridade, o que, segundo Pompa (2003), designou um processo histórico de marginalização.

> O aldeamento em si, a disposição das casas, a posição central assumida pela igreja (normalmente, tendo à frente um cruzeiro e uma enorme praça), a divisão da rotina entre as atividades produtivas e de aprendizado, tudo isso, para além dos conteúdos doutrinais e culturais transmitidos pelos missionários, contribuía para introduzir novos hábitos, novas concepções de tempo, espaço, moradia, família, trabalho, produção e sobrevivência (Santos, 2012. p. 49).

Isto posto, em acordo com as palavras de Santos (2012), foi por meio dessa fragmentação dos costumes indígenas que os jesuítas transformaram "cada aldeia em um 'grande projeto pedagógico total'".[2] Mediante os agrupamentos, os missionários usufruíram da oportunidade para aprender a língua dos indígenas aldeados na intenção de facilitar a conversão, portanto, foi a partir desse raciocínio que surgiram o *Catecismo da Doutrina Christãa na Lingua Brasilica da Nação Kiriri* (1698), *Arte de Grammatica da Lingua Brasilica da Naçam Kariri* (1699), ambos registrados pelo Pe. Mamiani, e o *Katecismo Indico da lingua Kariris* (1709), do Pe. Bernado de Nantes. Dessa forma, a catequização dos indígenas foi realizada em sua língua materna, reforçando a ideia de Nascimento (2019), de

[2] Referência utilizada por Lyrio Santos (2012), na sua tese de doutorado, em acordo com o seguinte trabalho: NEVES, O Combate dos soldados de Cristo, op. cit., p. 162.

que as prerrogativas coloniais se deram, principalmente, pela dominação linguística. No entanto, faz-se necessário considerar que inúmeras prerrogativas linguísticas podem ter ocorrido entre o processo de aquisição da língua pelos missionários em relação à tradução da visão de mundo cristã e colonizadora para base sociocultural e linguística de um povo que possuía uma visão cosmológica completamente diferente do homem europeu.

Por conta dessa inserção, os missionários da Companhia de Jesus obtiveram uma grande autoridade sobre os povos originários e, esse fato, foi malvisto pelos colonos da época, por acreditarem que os padres jesuítas monopolizavam a mão de obra indígena. Com o passar dos tempos, os jesuítas solicitaram ao rei o domínio da legislação temporal e espiritual das aldeias, o qual foi sancionado. Dessa maneira, os religiosos tornaram-se responsáveis pela administração total dos aldeamentos. Porquanto, como menciona Santos (2012, p. 45), "a salvação dos povos nativos era o principal objetivo da colonização", mas, a partir desse momento, a autoridade dos padres jesuítas foi minimizada gradativamente por pressões políticas e socioeconômicas, as quais envolviam domínio das terras ocupadas pelas aldeias e poder ideológico e simbólico dos missionários. Mediante os conflitos pertinentes às aldeias desse período, Pompa (2003, p. 299) elenca os principais problemas evidentes no sertão: a fome de terra, que gerou conflitos com colonos e "guerras justas, o poder administrativo e judiciário dos missionários, o controle dos "índios" fugitivos, os descimentos e a mistura étnica.

A partir desse ponto, a soberania jesuítica torna-se um estorvo para o Estado[3]. Com a intenção de amenizar os conflitos políticos dos colonos, dos civis e de outros grupos sociais contra ao grande poderio das missões, foram adotadas medidas legislativas encabeçadas por Mendonça Furtado, governador geral do Grão-Pará e do Maranhão, sob os conselhos do irmão Sebastião José de Carvalho e Mello, posteriormente conhecido como o Marquês de Pombal. É preciso ressaltar que as referidas medidas foram empregadas na província do Grão-Pará e que, tempos depois, foi instaurada a nível nacional, por meio do Alvará de 8 de maio de 1758. A sua empregabilidade nas outras províncias do Estado do Brasil ocorreu de forma particular para cada lugar. Na Bahia, em concordância com Santos (2012),

> [...] houve um esforço maior para que o Diretório fosse aplicado na capitania de Porto Seguro, por exemplo, do que nos demais espaços pertencentes à jurisdição da capitania da Bahia, a saber, Sergipe, Ilhéus e Espírito Santo. **Nesses espaços, tiveram maior impacto as leis de 1755,**

[3] Como atesta Santos (2012, p. 21) sobre as instruções de governo de Mendonça Furtado, "datadas de 31 de maio de 1751, alertavam-no a respeito do "excessivo poder que têm nesse Estado os Eclesiásticos, principalmente no domínio temporal nas suas aldeias"

> **as quais, embora fossem válidas, inicialmente, para o Estado do Grão-Pará e Maranhão, foram ampliadas para o Estado do Brasil por meio do alvará de 8 de maio de 1758**. (Santos, 2012, p. 27, grifo meu)

A lei de 17 de agosto de 1758 aprovou a aplicação do Diretório dos Índios, que complementava as duas primeiras leis criadas por Mendonça Furtado. Esse documento, assinado pelo Marquês de Pombal em 3 de maio de 1757, estabelecia diretrizes a serem seguidas nas povoações indígenas do Grão-Pará e do Maranhão, com o objetivo principal de impor a língua portuguesa e de promover a "civilização" (Santos, 2019, p. 247). Para entender plenamente o Diretório, é essencial destrinchar as medidas que o compunham, pois cada uma delas representava uma supressão específica que afetava o multilinguismo indígena, restringia a atuação dos jesuítas, subjugava os povos indígenas e atendia às demandas econômicas coloniais.

Uma das medidas, datada de 4 de abril, autorizava o casamento misto entre colonos e indígenas. Essa união não acarretaria infâmia social, e os filhos resultantes não teriam dificuldades em ocupar cargos públicos, ganhando prestígio sociopolítico e direito a terras. No entanto, a união matrimonial entre indígenas e não indígenas diluía os valores culturais dos povos originários, reforçando, em muitos casos, dinâmicas de subordinação, o que também se refletia na supressão de suas línguas. Outra medida, de 7 de junho, instituiu a Companhia de Comércio do Grão-Pará e do Maranhão, responsável pela introdução de africanos escravizados, com o objetivo de evitar que os colonos continuassem a escravizar os indígenas (Santos, 2011, p. 23).

O governador achava impróprio os jesuítas terem tanto domínio sobre os povos indígenas, dessa forma, buscou por problemas no regimento das missões para que, em sequência, os usasse para enfraquecer seu poderio. Inúmeros conflitos emergiram entre colonos e os aldeamentos, tanto pelo interesse nas terras que habitavam quanto pela possibilidade de subjugar a sua população. Em face da intensidade dos eventos recorrentes, o governador desconfiava da Companhia de Jesus e a acusava de dificultar a implementação do Tratado de 1750 com os espanhóis. Isso porque nas terras que seriam cedidas à Espanha, havia aldeamentos indígenas, e a remoção desses povos se tornaria necessária.

Assim, os jesuítas opuseram-se a essa medida e, a partir desse instante, conflitos militares e pressão política dos colonos levaram à ruína o poder político e espiritual que a Companhia de Jesus detinha. A promulgação da Lei de 6 de junho de 1755 referia-se à libertação dos indígenas e restringia o poderio dos jesuítas a respeito do regimento das aldeias. Doravante, os aldeamentos foram transformados em vilas e sua população poderia

ser empregada, a catequização foi mantida, no entanto, era o governador quem indicava os missionários para tal função (Santos, 2012).

Ao analisar o Diretório à luz desse contexto mais amplo, é possível identificar uma intenção subjacente de transformação social e cultural, revelando um projeto civilizatório que visava moldar as populações indígenas de acordo com os interesses coloniais. As medidas e o estabelecimento do Diretório promovem intencionalidades políticas-governamentais a respeito dos povos indígenas, por meio da visão de "civilização" veiculada no documento, o qual reflete um desejo mais profundo de impor valores e normas que coincidam com a agenda sociopolítica do colonizador português. No que diz respeito à regulamentação da língua portuguesa em detrimento das línguas indígenas, essa medida é uma ferramenta destinada a moldar a sociedade colonial conforme os ideais da coroa. Nos dizeres de Jesus (2022), fica explícito a intenção de assimilação total da cultura indígena à portuguesa, quando informa que as medidas do *Diretório* "objetivavam torná-los cristãos e falantes da língua portuguesa, além de fazê-los fabricar as suas casas não mais da maneira como as faziam, com sua própria identidade, mas imitando as casas dos brancos" (Jesus, 2022, p. 106). Em outras palavras, vivendo aos modos do homem branco.

3 LÍNGUA ATRAVÉS DO TEMPO HISTÓRICO: UM OLHAR INVESTIGATIVO

Desde tempos remotos, as terras brasileiras têm sido o palco de uma ampla diversidade de interações linguísticas, impulsionadas pela riqueza étnica do país. Essas interações, que abrangem uma variedade de aspectos demográficos, sociais e culturais, desempenharam um papel crucial na configuração do português brasileiro. No entanto, durante o período colonial, as autoridades coloniais procuraram promover uma certa padronização na língua, apesar de reconhecerem a diversidade multiétnica e linguística. Para uma compreensão abrangente do processo de formação do português brasileiro, é essencial revisitar sua história social desde suas origens até os dias atuais. Somente assim podemos compreender plenamente os fatores sócio-históricos e linguísticos que contribuíram para moldar a língua que é falada majoritariamente hoje no Brasil.

No âmbito da pesquisa em curso, busca-se ampliar a compreensão das línguas indígenas, considerando sua intersecção com a história do português brasileiro, afinal ambas histórias dividem trajetórias que se intercruzam até certo ponto, como demonstra os estudos e as reflexões propostas por Mattos e Silva no livro *Ensaios para uma sócio-histórica do Português Brasileiro (2004)*. Como reforça a própria autora, "se faz necessário conhecer como se passa o processo de contato sócio-histórico e linguístico entre as línguas indí-

genas e língua portuguesa" (Mattos e Silva, 2004, p. 16). A complexa interação entre o português europeu, as línguas dos povos originários e dos africanos escravizados trazidos demanda uma análise cuidadosa das dinâmicas de contato entre essas línguas. Essa análise nos permite compreender como o português brasileiro emergiu como uma língua dinâmica, heterogênea e profundamente influenciada pelo contexto histórico e cultural do país.

No ensaio "A sócio-histórica do Brasil e a heterogeneidade do Português Brasileiro", é apresentada uma análise significativa da trajetória da língua, que enfatiza a importância de compreendê-la dentro de um contexto social multilíngue, que, por sua vez, remodela seus aspectos linguísticos. Nesse sentido, Mattos e Silva (2004) examina as diversas perspectivas sobre a sócio-história do Brasil, as quais se refletem na heterogeneidade desse português. O referido ensaio inicia destacando a significativa contribuição de Serafim Leite e A. Houaiss para a compreensão da história externa do Português Brasileiro. Serafim Leite, ao se posicionar contra as teses "indianófilas e africanófilas" e defender a predominância do português, marcou o início da reconstrução sócio-histórica da língua. Apesar de sua ênfase na "vitória" do português, seu trabalho foi fundamental nesse processo. Por outro lado, o trabalho de Houaiss (1985) contextualiza o Português Brasileiro dentro de um cenário multilíngue e multidialetal. Ele descreve as dinâmicas demográfico-linguísticas em termos de "sucessivas distribuições" dos falantes (Mattos e Silva, 2004, p. 31-33).

Para uma compreensão abrangente do Português Brasileiro em suas manifestações oral e escrita, é essencial seguir as perspectivas delineadas por Houaiss (1985) e investigar cuidadosamente algumas problemáticas. Neste artigo, direcionamos nossa análise e nossa reflexão para a primeira delas, que envolve o "levantamento exaustivo de depoimentos diretos e indiretos sobre todos os processos linguísticos ocorridos desde o início da colonização, englobando os estudos históricos a partir do século XIX até os dias atuais". Ao nos depararmos com essa questão, torna-se clara a relevância fundamental de compreender e de reconstruir a sócio-história do Brasil e da língua portuguesa, seguindo a ideia de Houaiss de que "o português nasce com diversidade" (Houaiss, 1985, *apud*, Mattos e Silva, 2004).

A percepção do *multilinguismo generalizado*, termo consolidado pela grande pesquisadora Mattos e Silva, resguarda em si a história social do povo brasileiro e o seu aspecto linguístico desde os primórdios da colonização. Por meio de estudos demográficos, mostra-se que, do século XVI ao século XVIII, a maioria dos habitantes em território

brasileiro eram os indígenas e os africanos, enquanto os portugueses estavam em menor número. Esse fator demonstra que, nesse período, havia uma grande pluralidade linguística no território brasileiro. No entanto, desde meados do século XVIII até os tempos atuais, esse multilinguismo não é mais generalizado, ele é localizado. Portanto, o contexto sociolinguístico do Brasil é marcado por um paradoxo linguístico, como define Simas e Lucchesi (2020), ao mencionar que 98% dos brasileiros falam a língua portuguesa e 2% falam mais de 180 línguas indígenas e 50 falam línguas de imigração. Ou seja, ainda mantemos essa característica do multilinguismo, a qual, segundo Matos e Silva, também é responsável por caracterizar a língua portuguesa do Brasil, que possui particularidades distintas do português europeu.

Diante desse cenário, nos vem o seguinte questionamento: o que aconteceu na história linguística do Brasil para que o multilinguismo deixasse de ser generalizado e passasse a ser localizado? Além disso, quais foram os fatores sócio-históricos e culturais que marcaram as diferenciações linguísticas do que compreendemos ser hoje o português brasileiro? Para, então, obter respostas a esse respeito, é necessário investigar profundamente os encontros e os desencontros com os maiores detalhes possíveis do contato massivo entre línguas no que concerne à história da formação do Brasil.

Para isso, em conformidade com os princípios da Linguística Histórica (LH), que destaca a ideia de que as línguas não são estáticas, mas estão em constante transformação, buscamos compreender os processos dinâmicos que as afetam. As línguas estão sempre em movimento, moldadas pela capacidade do homem de dar significação a partir da sua construção histórica, nesse sentido, como afirma Costa (1996, p. 51) "A língua não é, como muitos acreditam, uma entidade imutável, homogênea, que paira por sobre os falantes. Pelo contrário, todas as línguas vivas mudam no decorrer do tempo e o processo em si nunca para". Dessa forma, a LH é uma ferramenta imprescindível para notar as mudanças estruturais da língua em decorrência do tempo e, também, para compreender a história social que cada língua possui. Nessa perspectiva, será utilizada para pesquisa a vertente *stricto sensu*[1] da Linguística Histórica que, segundo Mattos e Silva (1999, p. 151), "[...] se concentra na mudança lingüística no tempo, levando em consideração fatores intralingüísticos ou estruturais e fatores extralingüísticos ou sócio-históricos".

Conforme exposto por Conde Silvestre no livro *Sociolinguística Histórica* (2007), a Linguística Histórica investiga o desenvolvimento das línguas ao longo do tempo, concentrando-se nos modos de mudança e nas condições que a propiciam. Para tal, utiliza

a comparação entre manifestações textuais da mesma língua ou de variedades afins pertencentes a períodos diferentes. O objetivo é compreender os fatores que explicam essas mudanças, além de buscar uma representação e explicação de forma universal.

Neste estudo, destaca-se o capítulo "*Problemas y principios*", que enfatiza a dependência da sociolinguística histórica na recuperação de fatos linguísticos do passado a partir dos textos que sobreviveram. Faz-se uso dos princípios e dos métodos propostos por Conde Silvestre, com especial atenção aos problemas relacionados ao material linguístico histórico. O linguista-historiador deve estar atento às dificuldades de lidar com informações fragmentadas e frequentemente desconectadas da produção real dos falantes, situação agravada pela preservação dos dados linguísticos, que distanciam os textos da realidade da língua falada e do contexto em que foram gerados. Nesse sentido, a sociolinguística histórica estende sua metodologia aos dados históricos, estabelecendo uma correlação entre a variação manifestada nos documentos preservados e determinados fatores sociais condicionantes.

De acordo com Conde Silvestre (2007), os princípios metodológicos na sociolinguística histórica envolvem a construção e a utilidade do material linguístico histórico, a reconstrução do contexto social e a investigação do estilo. A Linguística Histórica contribui para a formação de *corpora* linguísticos ao selecionar textos representativos, ao realizar análises diacrônicas e ao desenvolver ferramentas específicas. Além disso, promove o alinhamento temporal, identifica mudanças linguísticas e integra perspectivas sociolinguísticas.

A reconstrução do contexto social é fundamental e divide-se em duas vertentes: a investigação do estilo, que revela as mudanças linguísticas e os usos da linguagem em diferentes contextos sociais, e a contribuição da História Social, que contextualiza essas mudanças em um quadro mais amplo de transformações sociais. Essa interseção entre linguística histórica e história social enriquece a compreensão das complexas relações entre linguagem e sociedade ao longo do tempo, permitindo uma análise mais profunda da evolução da língua (Conde Silvestre, 2007, p. 41-43).

Considerando os aspectos extralinguísticos, baseamo-nos nos princípios da História Social Linguística, conforme proposto por Mattos e Silva (2008). Essa abordagem explora a interseção entre a língua e a sociedade, reconhecendo sua conexão intrínseca e levando em conta suas transformações ao longo do tempo histórico. Ao adotarmos essa perspectiva, torna-se evidente a impossibilidade de dissociá-las. No que se refere à

realidade brasileira, a história social linguística do Brasil destaca a importância de considerar não apenas a língua portuguesa, mas também outras línguas indígenas, africanas e europeias presentes no país.

Esse Português no Brasil é marcado desde o período colonial por cinco fatores importantes para sua interpretação: (1) a demografia histórica do Brasil do século XVI ao XIX; (2) a mobilidade populacional dos africanos no Brasil colonial e pós-colonial; (3) a escolarização ou a sua ausência do século XVI ao XIX; (4) as reconfigurações socioculturais, políticas e linguísticas ao longo do século XIX (Mattos e Silva, 2004, p. 125). Isto posto, é essencial a tentativa de interpretação do passado sócio-histórico-linguístico do Brasil por meio desses fatores, a fim de compreender a natureza heterogênea, variável, plural e polarizada do português brasileiro contemporâneo.

A aproximação ao processo sócio-histórico da origem e da formação do português brasileiro revela uma história marcada por silenciamentos e por esquecimentos. Desde os primeiros momentos da colonização, o Brasil foi caracterizado por um multilinguismo generalizado, com portugueses e com diversas etnias indígenas e africanas convivendo no nordeste do país durante o primeiro século de colonização. Posteriormente, esse multilinguismo tornou-se mais localizado em determinadas regiões do território brasileiro (Santos, 2019). As políticas linguísticas implementadas no Brasil, notadamente as do Marquês de Pombal em 1757, aliadas à ação da escola, contribuíram para estabelecer o português como língua nacional. Estudos demográficos, como demonstra o trabalho de Mussa (1991), mostram a diminuição dos povos indígenas devido ao genocídio e ao aumento da população africana durante o século XVII, devido ao ciclo do açúcar. Esse contexto propiciou o contato linguístico entre as línguas indígenas e o português, assim como entre os africanos e o português como segunda língua.

Por meio dos fatos históricos e linguísticos abordados na seção anterior, pode-se notar que essa concepção foi utilizada como um dos mecanismos coloniais, Nascimento (2019, p. 44-45) diz que "o sistema perverso de colonialidade, que produziu no Ocidente séculos de escravidão negreira e dizimação dos povos originários de cada lugar onde se colonizava, não se deu fora, mas dentro dos sistemas linguísticos". Diante disso, é crucial reconhecer que muitas narrativas ainda aguardam redescoberta, reconstrução e análise aprofundada, incluindo a história do Português Brasileiro. Essa língua, tão multifacetada e diversa, carrega consigo as marcas dessas complexas interações históricas, da mesma forma que a história das línguas indígenas do Brasil ao longo dos séculos.

4 A (RE)EXISTÊNCIA DA FAMÍLIA LINGUÍSTICA KARIRI: TRONCO, POVO E LÍNGUA NO INTERIOR DA BAHIA

A campanha jesuítica seguiu com o projeto de adentrar os sertões para tornar os povos que ali se estabeleciam em fiéis do Cristianismo, devido a isso, quanto mais distâncias as missões percorriam, maior foi o contato com a multiculturalidade indígena. Porquanto, notou-se que os povos originários dos "caminhos de dentro"[4] se diferenciavam culturalmente dos povos do litoral, a começar pela língua, a qual era diferente daquela falada pelos indígenas da costa, a então denominada *Língua Geral*, oriunda do tronco Tupi e utilizada para intermediar a relação com os indígenas em favor do avanço colonial e comercial.

Segundo Rodrigues (2002), visualiza-se nos sertões baianos a presença da língua "*Nhengaíba* (referência de fala ruim, língua travada) e, em oposição, aos povos que falavam o *Nhengatu* (referência a uma boa fala) na Ilha de Marajó", referindo-se às línguas faladas pelos indígenas do interior do estado. É interessante notar que a designação de "fala ruim" foi inicialmente atribuída pelos habitantes da costa litorânea aos povos das entranhas dos sertões. Mais tarde, essa percepção foi utilizada também pelo viés colonizatório. Isto posto, faz-se necessário a contribuição de Viveiros de Castro a respeito do uso dos etnônimos.

> Não é tampouco por acaso que a maioria dos etnônimos ameríndios que passaram à literatura não são autodesignações, mas nomes (freqüentemente pejorativos) conferidos por outros povos: **a objetivação etnonímica incide primordialmente sobre os outros, não sobre quem está em posição de sujeito**. (Viveiros de castro, 1996, p. 125, grifo meu)

Em outras palavras, os etnônimos são frequentemente termos pejorativos devido a terem sido atribuídos por outros povos ou terem sido inventados pelos colonizadores. No caso aqui analisado, o termo destinado à língua contraposta ao Nhengatu possui uma carga desqualificadora. Se essa língua é considerada travada e de "fala ruim", ela não deveria ser considerada nos processos de comunicação. E, assim, o Nhengaíba foi estigmatizado, assim como os seus falantes, como será discutido a seguir. Em consonância com Santana e Mendes (2020) *apud* Cardim (1583[1977]), a língua *Nhengaíba* era considerada corrupta e travada e utilizada pelos povos indígenas que viviam espalhados pelo território dos semiáridos baianos, conhecidos como "tapuias"[5].

[4] Termo encontrado no trabalho de Santana e Mendes (2019), que é utilizado para a caracterização de regiões situadas longe da costa litorânea, o interior do estado. Segundo os autores mencionados, o termo "fora revelado pelo cronista Pero de Magalhães de Gândavo" e está citado em Amado (1995).

[5] Santos (2011, p. 61) faz alusão ao trabalho de antropólogos e de historiadores a respeito do não reconhecimento do termo enquanto designação étnica, mas como "um termo genérico usado, inicialmente, pelos próprios grupos falantes do Tupi-guarani – e adotada pelos portugueses – para se referir aos demais povos indígenas.

Os "tapuias" eram considerados bárbaros, arredios e selvagens, nomeados assim pelos indígenas da costa, como já mencionado, por conta da contraposição das suas particularidades socioculturais e linguísticas. Logo, os povos indígenas dos sertões foram designados com esse termo, o qual pertence ao tronco Tupi e significa "inimigo". O *Dicionário Tupi (antigo) - Português* (1987, p. 278), define o vocábulo como "Bárbaro; Escravo. Tapúia (desprezível para o tupi, como inimigo)".

> A expressão tapuio é trazida por alguns dicionaristas, a exemplo do dicionário Houaiss, como uma denominação dada aos índios, pelos portugueses, de grupos que não falavam as línguas do tronco linguístico Tupi e que habitavam mais para o interior do país, pertencentes, naturalmente, ao tronco linguístico Macro-Jê. (Cunha e Sousa, 2011, p. 46)

Observa-se, então, uma generalização de todos os povos que habitavam o interior do estado e que não falavam a *língua geral* ou as línguas do tronco tupi, e se diferenciavam dos indígenas do litoral, enquanto "tapuias", perpetuando, assim, um ideal de marginalização (Dantas, Sampaio e Carvalho, 1992). E esse é mais um aspecto da iniciativa colonial em tornar a cultura indígena, de meados do século XVII, em algo hegemônico e estigmatizado. Por essa via, Santos (2012) informa que é difícil indicar atualmente quais etnias, especificamente, integravam à nação dos "povos arredios", no entanto, é certo que são predominantemente da família linguística Kariri. Segundo Taunay (1936, p. 11), "ao interior do Nordeste habitavam as tribos tapuias, designadas sob a denominação genérica de Cariris".

Os povos indígenas originários dos sertões baianos, pertencentes à família Kariri, foram amplamente documentados nos aldeamentos administrados pelos jesuítas. Marcaram presença nos aldeamentos sob a administração dos padres jesuítas, certificada pelo *Catecismo da Doutrina Christãa na Lingua Brasilica da Nação Kiriri* (1698) e *Arte de Grammatica da Lingua Brasilica da Naçam Karirí* (1699) do Pe. Mamiani, e pelo Capuchinhos, na *Relação de uma missão no Rio São Francisco* (1702) e *Katecismo Indico da lingua Kariris* (1709) dos irmãos Martin de Nantes e Bernardo de Nantes, respectivamente, assim como cartas ânuas e demais documentações da época. Quanto à integração da família Kariri ao tronco Macro-Jê, pode-se dizer que eles são membros presumidos, conforme sugerido hipoteticamente por Rodrigues (1986). Além disso, Soares e Carvalho (2014, p. 551) afirmam que, em comparação com o Tupi, "o chamado tronco Macro-Jê apresenta um caráter muito mais hipotético, pois postula uma relação de ancestralidade comum para um conjunto de línguas e famílias linguísticas mais diversas e distintas entre si".

> Entre os 'Tapuias, no sertão da Baía', os índios chegavam a 2.500, distribuídos em quatro aldeias: Natuba, Canabrava e Saco dos Morcegos, na capitania da Bahia e Jeru, em Sergipe, cujas populações eram formadas principalmente por índios do grupo Kiriri ou Cariri (Câmara Jr, 1965, p. 99).

Percebe-se, então, que a principal etnia designada pelas missões como povo "tapuia" são os indígenas da família Kariri. Por esse caminho, pode-se afirmar que os indígenas que vivem, hoje em dia, na Aldeia Indígena de Mirandela, na Aldeia Massacará e na Aldeia do Geru, são remanescentes dos "tapuias do sertão" do século XVII e meados do sécúlo XVIII. Demarquet (1991, p. 27) informa que "as missões jesuíticas de Natuba, Canabrava, Saco dos Morcegos e Geru correspondem as localidades atuais de Nova Soure, Ribeira do Pombal, Mirandela e Jeru, respectivamente". Logo, a população que vive nesses municípios (indígenas ou não-indígenas) exprime em seu cotidiano o legado linguístico e cultural dos seus ancestrais, mesmo sofrendo ações empregadas pelo colonialismo.

Devido à pluralidade linguística entre os "tapuias", a conversão só seria possível por meio da língua materna desses povos ou pelo estabelecimento de uma língua geral. Cardim (1583[1977]) afirma que, para tal feito, seria necessário que esses indígenas aprendessem a língua dos *filhos do mar,* tronco Tupi, para mediar o processo de catequização, basicamente como tradutores. Inicialmente, alguns missionários empregaram a língua geral (Nhengatu) como instrumento nos esforços de conversão dos povos considerados "hostis", mas, ao longo do tempo, passaram a utilizar as línguas maternas dessas comunidades para facilitar a comunicação e a integração. Como designa, Martins de Nantes (1979 [1709], p.18) "todos os nossos missionários entre os cariris se empenharam em aprender a sua língua, sem o uso da qual era impossível transformá-los em verdadeiros cristão, pois que é natural do bárbaro a barbaria." Assim, observamos a primeira indicação de contato linguístico entre a língua geral, o português colonial e as línguas dos *tapuia*, sugerindo a possibilidade de processos de transferência linguística e dentre outros devido ao intenso contato.

A referida declaração carrega uma forte carga pejorativa ao associar os povos indígenas a um estado de barbárie intrínseca, reproduzindo o discurso colonial que desqualificava culturas e línguas nativas ao colocá-las em oposição à "civilização" europeia. Dessa

forma, a eleição de uma língua geral, acompanhada de um discurso depreciativo em relação às demais línguas indígenas faladas pelos povos "tapuia", teve como efeito secundário a redução da multiplicidade linguística dos povos originários, que predominavam no interior da Bahia. As diversas línguas gerais brasileiras não eram utilizadas por um grupo homogêneo de falantes que compartilhassem uma base cultural comum. Contudo, essas línguas funcionavam como códigos supraétnicos, adotados por diversas etnias que, posteriormente, foram exterminadas ou forçadas a migrar para o interior (Bortoni-Ricardo, 2021).

De acordo com registros históricos e estudos sobre os povos indígenas do Nordeste, observa-se que uma parte significativa do grupo identificado como *tapuia* estava vinculada à família Kariri, como já mencionado. Nota-se que a população de algumas aldeias no período colonial era registrada como *tapuia* ou *nação Kariri*, o que indica uma considerável mistura de etnias nos sertões. É fundamental considerar as práticas de *entradas* e *descimentos*, expedientes de expansão e domínio territorial que tinham como objetivo a conversão forçada dos indígenas. Essas investidas não apenas promoviam o deslocamento compulsório de diversos grupos, empurrando-os ora para o litoral, ora para o interior, mas também desencadearam profundas transformações nas dinâmicas sociais e culturais das aldeias. O contato forçado, a dispersão territorial e a reorganização imposta resultaram em um intenso processo de miscigenação, impactando diretamente a preservação dos costumes tradicionais de cada povo (Soares e Matarezio Filho, 2025).

As línguas indígenas pertencentes à família linguística *Kariri,* de que se têm notícia, são quatro: *Kipeá, Dzubukua, Kamuru* e *Sapuya*. O Dzubukuá foi registrado pelos capuchinhos, Martin e Bernardo de Nantes, nas aldeias de Aracapa, de Cavalo e de Pambu. Enquanto o Kipeá foi registrado pelo jesuíta Luis Vicêncio Mamiani nas aldeias de Jeru, de Natuba, de Canabrava e de Saco dos Morcegos, as quais foram incluídos nas diligências referentes ao alvará de 8 de maio, em 1758 (Santana e Mendes, 2020).

Novais-Carneiro e Almeida (2011), com base em Dantas, Sampaio e Carvalho (1992), indicam que os falantes do Kipeá se estendiam até o Rio Salitre, os do Dzubukuá no médio São Francisco, os do Sapuyá (ou Sabujá) na região do Rio Paraguaçu, e o Kamuru, também conhecido como Pedra Branca (Bandeira, 1972), na região do Rio de Contas. Rodrigues (1986) localiza, de forma semelhante, o Kipeá (ou Kiriri, como ele próprio o denomina) no nordeste da Bahia e em Sergipe, e o Dzubukuá entre a Bahia e Pernambuco, próximo a Cabrobó.

Em decorrência das estratégias da colonialidade, todas as línguas pertencentes à família Kariri desapareceram ao que se refere ao seu uso vernacular[6]. No entanto, há registros das línguas Kipeá (ou Kiriri)[7] e Dzubukuá, datados dos séculos XVII e XVIII (Rodrigues, 1986). Segundo Bandeira (1972) *apud* Bernard de Nantes (1896), as aldeias de língua Kipeá se diferenciavam das aldeias de língua Dzubukuá, sendo essa distinção principalmente no campo linguístico.Contudo, havia uma certa unidade cultural entre os grupos, como observado por Bandeira (1972) ao analisar os relatos de missionários em documentos da época. As "práticas proibidas", mencionadas nas narrativas, eram as mesmas entre esses dois grupos, o que evidencia essa similaridade cultural. No âmbito linguístico da família Kariri, os diferentes grupos étnicos falavam línguas específicas dessa família (Soares e Matarezio Filho, 2025).

Observa-se então que as quatro línguas que compõem a família linguística Kariri são linguística e culturalmente presentes em grupos étnicos distintos[8]. Como mencionado anteriormente, a língua Kipeá manifesta-se entre o povo Kiriri da Aldeia Mirandela (antiga Saco dos Morcegos) e, adicionalmente, entre o povo Kaimbé, Aldeia Massacará, na região do Rio Itapicuru, no nordeste da Bahia, bem como nas aldeias de Canabrava e Natuba (Azevedo, 1965; Bandeira, 1972; Rodrigues, 1986).

A língua Dzubukuá é atestada entre os povos da bacia do São Francisco (Nantes, 1896), frequentemente referidos como tapuias ou pertencentes à nação Kariri, conforme discutido anteriormente. Segundo Letícia Tumbalala (Santana, 2023, p. 13), "os tapuias' falavam diversas línguas, entre elas o Dzubukuá, que também é reconhecida como pertencente aos Tumbalalá, no Pambu". Além do povo Tumbalalá, outros grupos da região do rio São Francisco também são falantes do Dzubukuá, como os povos Tuxá e Truká. Segundo Segundo Durazzo (2021), embora Tuxá e Truká se reconheçam etnicamente

[6] Embora essas línguas sejam consideradas extintas do ponto de vista linguístico (UNESCO, 2003), este trabalho considera que elas continuam vivas, sendo faladas em contextos socioculturais e ritualísticos específicos, em um estado de adormecimento. Este ponto será abordado nas próximas seções, com base no conceito de "língua encantada" de Durazzo e Bonfim (2023).

[7] Nos estudos de Aryon Rodrigues, o termo "Língua Kiriri" é utilizado para se referir também à língua Kipeá, devido ao fato de ela ser amplamente falada pelo povo Kiriri. Em razão do valor significativo de seu trabalho nas áreas linguística e antropológica, alguns estudos adotam a designação "língua Kiriri" para a língua Kipeá. Contudo, neste trabalho, opta-se por usar especificamente o nome da língua.

[8] Durazzo e Bonfim (2024) abordam o adormecimento e o encantamento das línguas indígenas ancestrais do Nordeste, situando-as na dimensão dos "mais-que-humanos" e destacando sua manifestação em contextos de rituais como línguas encantadas, mantidas vivas pela cosmologia dos povos indígenas. Com base nessas reflexões, esta pesquisa reafirma a vivacidade dessas línguas, tema que será aprofundado em outra seção. Conforme Durazzo (2021, p.431) "Apesar de não possuírem a língua ancestral, como a chamam, ativa no registro cotidiano e linguisticamente fluente, os Tuxá afirmam possuir, sobretudo em contextos rituais, a *língua* da *ciência* (DURAZZO, 2019)."

como parte da nação Proká, indígenas aldeados pelos capuchinhos no sertão de Rodelas, mas essa relação não se estende à esfera linguística.

> É certo que a Ilha da Assunção, onde habitam os Truká, era missão estabelecida já na fronteira entre as aldeias e populações de Prokás e Kariris (NIMUENDAJU, 2017). Talvez por isso o vínculo histórico entre Ilha da Assunção e Rodelas, esta mais aproximada dos Proká e Brancararus, tenha permitido aos pesquisadores indígenas atuais estabelecerem relações de autoafirmação proká, quando étnica, e kariri, quando linguística (Durazzo, 2021, p.434).

Além das informações disponibilizadas por Novais-Almeida (2011) sobre a localização das línguas Sapuyá e Kamuru, Azevedo (1965) registra que a língua Sapuyá foi documentada na região sul da Bahia, enquanto a Kamuru foi identificada na Aldeia de Pedra Branca, também no estado. Ambas foram documentadas por Martins de Nantes em 1867. Além disso, em 1818, ele compilou uma lista de 100 a 150 palavras dessas línguas, cuja localização atual permanece desconhecida (Azevedo, 1965, p. I-III). Para facilitar a compreensão da distribuição dessas línguas entre os respectivos povos, territórios e registros coloniais pesquisados por diversos autores, apresenta-se a Quadro 01 a seguir.

Quadro 01 - Distribuição geográfica, étnica e missionária das línguas Kariri nos sertões da Bahia

Família Linguística Kariri	**Localização**	**Povos**	**Resgistro**	**Aldeias missionárias**
Kipeá	Bacia do Itapicuru até o Rio Salitre	Kiriri Kaimbé	Pe. Mamiani (1698,1699)	Saco dos morcegos; Canabrava; Natuba e Massacará
Dzubukuá	Bacia do Rio São Francisco	Tumbalalá Tuxá Truká	Bernardo de Nantes (1702) Martinho de Nantes (1979, 1709)	Aracapá; Pambú; Rodelas; Cavalo; Irupua; Vargem etc
Kamuru (Pedra Branca)	Região do Rio de Contas; Amargosa – Ba	_	Martins de Nantes (1867)	Aldeia Pedra Branca
Sapuyá (Sabujá)	Região do Rio Paraguaçu; Sul da Bahia	_	Martins de Nantes (1867)	_

Fonte: adaptado de Bandeira (1972); Santana (2023), Durazzo (2021); Novais-Carneiro e Almeida (2011); Dantas, Sampaio e Carvalho (1992).

Dentre os povos e línguas mencionados na tabela acima, apenas as línguas Kipeá e Dzubukuá, localizadas especificamente na região norte dos sertões da Bahia, foram concentradas nas aldeias missionárias dos Kariri, como já mencionado anteriormente. Se houve algum contato linguístico com as demais línguas que compõem a família Kariri, até o momento, isso não foi devidamente explorado em estudos, tampouco há registros que documentem tal interação. Além disso, não há indícios de que tenha ocorrido um contato direto com essas línguas.

No entanto, não se pode descartar a possibilidade de interação entre as línguas originárias dos sertões-norte da Bahia e outras línguas da mesma família ou até de diferentes grupos linguísticos. Considerando a dinâmica já discutida dos processos coloniais, os contatos linguísticos são plausíveis, especialmente devido às intensas migrações e deslocamentos forçados que os povos indígenas enfrentaram nos dois primeiros séculos de colonização. Essas condições criaram cenários propícios para trocas culturais e linguísticas entre diferentes comunidades.

Para compreender com mais cautela a História Social da família linguística Kariri, no capítulo *Os Povos Indígenas no Nordeste Brasileiro: um esboço histórico*, Dantas (1992) contribui com um aparato explicativo sobre a diversidade étnica presente nos sertões, particularidades sócio-históricas, organização social, as revoltas e os conflitos que os povos indígenas sofreram desde o Brasil colônia, precisamente nas áreas sertanejas do nordeste brasileiro, evidenciando, por meio de tal explanação, que a língua é um fator histórico.

4.1 As línguas dos "tapuia" e o tronco macro-jê

A paisagem étnica e linguística brasileira é composta por uma riquíssima diversidade e, dentro desse cenário, há mais ou menos 180 línguas indígenas, as quais pertencem a troncos e famílias distintas e, em alguns casos, se configuram como isoladas. Como já foi discutido até aqui, a composição dessa paisagem possui raiz na história colonial, dessa maneira, a integridade das ramificações linguísticas foi atravessada por mudanças e variações em decorrência das políticas linguísticas que desejavam a homogeneização indígena. Estabeleceu-se nesse período a categorização "índio" com essa intencionalidade, além de marcar a distinção entre colonizador e colonizado (Leite e Franchetto, 2006, p. 12).

O projeto de homogeneização da diversidade étnica e linguística do Brasil foi implementado por meio de diversas políticas ideológicas e linguísticas, como já mencionado nas seções anteriores, que moldaram a atual composição dos troncos, famílias e

línguas indígenas presentes nos sertões do norte da Bahia. É importante notar que outros fatores também desempenharam um papel significativo nesse processo. Franchetto *et al.* (2019, p. 13) observam que "migrações, afastamentos ou aproximações entre povos, no tempo ou no espaço, geram novas variedades ou novas línguas, ou podem apagar variedades e línguas existentes".

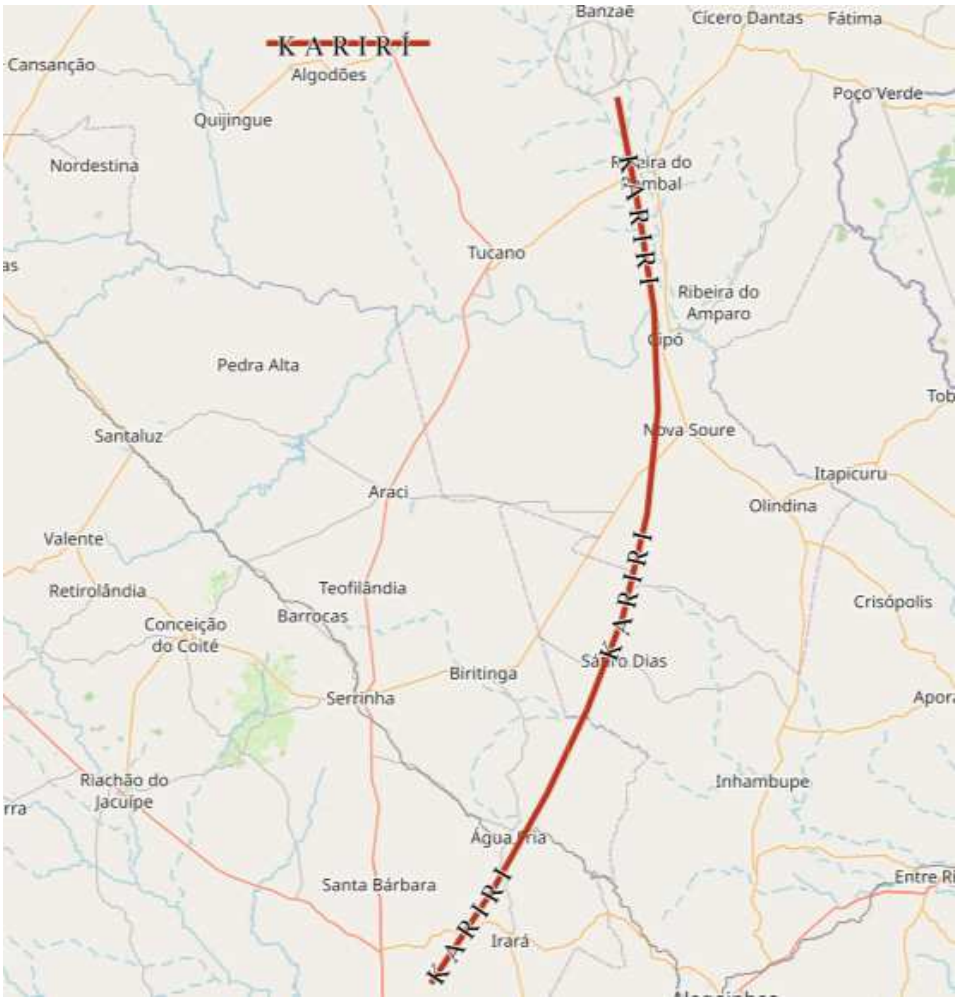
É interessante observar que o deslocamento geográfico dos povos indígenas é muito importante, por estar pautado em experiências de contato entre povos e línguas distintas, fato também responsável por caracterizar a rica diversidade linguística brasileira. Nesse sentido, é importante compreender o deslocamento da família linguística Kariri. Essa família se estabeleceu em maior número no interior do estado, embora sua presença nas regiões litorâneas não possa ser descartada. No entanto, a literatura aponta para uma maior incidência nos caminhos de dentro.

> Os cruzamentos intertribais deverão ter existido, pois não só havia em Canabrava outros Karirí, como em Massacará os Kaimbé, além de aldeamentos existentes em Jeremoabo e outras localidades. Remanescentes de outras tribos poderão ter se misturado aos Karirí, vez que no sertão da Bahia apenas Mirandela e Massacará resistiram até nossos dias. A miscigenação continuou se processando acrescida de elementos de raças branca e negra (Bandeira, 1972, p.31).

Nesse cenário, a distribuição da família Kariri na região norte da Bahia, conforme representada no *Mapa Etno-Histórico do Brasil* (1944)[9], de Curt Nimuendaju, oferece uma perspectiva visual valiosa. Esse registro contribui para a compreensão do enraizamento das línguas indígenas no solo fértil dos sertões baianos. No mapa, Nimuendaju descreve a distribuição geográfica das línguas Dzubukuá, Kamuru e Sapuyá, associando-as à família Kariri, com a inscrição "Dzubukuá-Kariri", por exemplo. No entanto, para a língua Kipeá, ele adota uma abordagem diferente. Nas áreas onde essa língua é manifestada, Nimuendaju se limita a classificá-la genericamente como "Kariri", sem especificar a língua. Abaixo, apresenta-se um recorte do mapa que ilustra a distribuição geográfica da língua Kipeá.

[9] Para visualizar a localização geográfica das línguas Sapuyá, Kamuru e Dzubukuá, consulte o mapa online disponível no site http://mapa-nimuendaju.eita.coop.br , pesquisando pela família Kariri.

Figura 1 – Disposição geográfica da língua Kariri-kipeá nos sertões-norte da Bahia



Fonte: Mapa etno-histórico de Curt Nimuendaju (1944).

Alguns dos aldeamentos evidenciados pelos documentos históricos do período colonial permanecem ativos até os dias atuais. A Aldeia Indígena Massacará, localizada em Euclides da Cunha – BA, situa-se a 360 km da capital, Salvador, enquanto o Território Indígena Kiriri, Mirandela, em Banzaê – BA, está a 312 km. Ambas pertencem à mesma família linguística, a Kariri. A experiência dos povos que habitam essas localidades ressalta a extrema importância de sua identidade linguística. Franchetto (2020) observa que os falantes de uma língua considerada extinta veem essa língua como um meio fundamental de reivindicar direitos territoriais, além de buscar o reconhecimento de sua existência e resistência, mediante a revitalização linguística.

A família linguística Kariri está, como já mencionado, hipoteticamente vinculada ao tronco Macro-Jê. Do ponto de vista territorial, ocupa uma região que se presume ter

sido anteriormente habitada por povos desse tronco. A compreensão da constituição desse tronco é desafiadora e ainda suscita debates entre pesquisadores renomados, principalmente devido à escassez de registros ao longo da história, especialmente quando comparada à abundância de registros do Tupi. Além disso, o tronco macro-jê foi profundamente afetado pela política de redução da diversidade linguística. Neste contexto, Aryon Dall'Igna Rodrigues destaca-se como um dos principais linguistas dedicados ao estudo desse tronco linguístico. Ele explorou a hipótese de parentesco genético a longa distância para caracterizá-lo, conforme discutido por Soares e Carvalho (2014), ao citar Campbell (2012, p. 135). Esses estudos também identificam as supostas famílias que compõem esse tronco, destacando a complexidade e a diversidade dentro desse agrupamento linguístico.

> **Além da família Jê**, que funciona como espécie de pivô do grupo, as outras famílias arroladas por Rodrigues como **membros putativos do tronco Macro-Jê seriam as famílias Krenák, Maxakalí, Boróro, Kamakã, Purí, Yatê, Guató, Karajá, Rikbatsá, Karirí e Ofayé**. (Soares e Carvalho, 2014, p. 551, grifo meu)

Os estudos de Nikulin e Carvalho (2019) corroboram com a percepção de que o tronco linguístico macro-jê abarca diversas famílias linguisticamente distintas, entre elas Chiquitano, Jabuti, Rikbaktsa, Ofayé, Karajá, Maxakali, Kamakã, Krenak e Jê. Ademais, os autores abordam a organização interna dessas famílias, bem como exploram as reconstituições fonológicas e sintáticas das proto-línguas, além de investigarem mudanças linguísticas por meio de análises filológicas. A análise diacrônica das línguas pertencentes ao tronco macro-jê tem recebido crescente destaque como um foco significativo de estudos e investigações, sobressaindo-se o recente trabalho de Andrey Nikulin (2020). O objetivo primordial dessas análises é desvendar a evolução histórica dessas línguas, abrangendo elementos fonológicos, morfológicos e sintáticos. Concomitantemente, busca-se estabelecer conexões de parentesco genético entre as diversas línguas do tronco macro-jê, enquanto se exploram as transformações linguísticas ocorridas ao longo do tempo. Essas pesquisas têm contribuído significativamente para uma compreensão mais profunda da riqueza linguística do Brasil, ao mesmo tempo que desempenham um papel crucial na preservação das línguas indígenas.

Nikulin (2020), em sua tese, após realizar estudos comparativos das proto-línguas vinculadas ao macro-jê, exclui as famílias Boróro, Yaathê, Purí, Guató, Kariri e Otí da proposta de pertencimento ao tronco macro-jê. Entretanto, o estudo não descarta com-

pletamente a possibilidade de uma relação genética distante entre as famílias Boróro e Kariri e o referido tronco. Portanto, não há uma confirmação conclusiva de que a família Kariri seja parte do tronco macro-jê, embora possa ter estado associada a ele em algum momento histórico, atualmente essa relação é considerada insustentável. Assim, a questão permanece aberta a investigações adicionais (Nikulin, 2020).

Ainda que a classificação linguística permaneça em debate, a compreensão da situação atual das línguas indígenas exige uma análise que vá além das filiações genealógicas. Afinal, para compreender como encontram-se as línguas indígenas na atualidade, é preciso compreender as políticas de dominação utilizadas pelo processo de colonização. Santos (2012, p. 4) retrata o contexto inicial da doutrinação política e religiosa dos povos originários, que ocorreu por meio de um "processo de conversão pautado na conquista de almas, terras e poder". Rodrigues (1993) menciona que a depreciação das Línguas Indígenas é consequência da força colonialista do século XVII e, consequentemente, de uma articulação política para deslegitimar os povos indígenas que se fazem presentes nos dias atuais de uma forma estrutural e, em alguns momentos, velada.

> A aversão pela diversidade étnica e pela diversidade linguística continuam sendo uma das mais fortes heranças que a sociedade brasileira recebeu dos colonizadores portugueses. Mesmo as línguas indígenas mais faladas no presente estão submetidas a fortes pressões dos mais diversos setores da sociedade majoritária, que as vêem como empecilho à integração, isto é, ao domínio ou domesticação das minorias indígenas (Rodrigues,1993, p. 99-100).

Sob essa perspectiva, é evidente que os sertões da Bahia, especialmente a região norte, refletem os processos de formação do Brasil colonial, cujas marcas perduram até os dias atuais. Ao focar nos estudos das línguas indígenas dos sertões baianos, emerge uma conexão intrínseca com a noção de retomada e preservação cultural. A análise das fundamentações teóricas permite identificar tanto os elementos socioculturais e históricos quanto os aspectos linguísticos envolvidos, iluminando a complexidade desse contexto e enriquecendo nossa compreensão das dinâmicas de continuidade e mudança presentes nessa região.

Nesse sentido, a revitalização das línguas ocorre de maneira emergente, apesar das políticas linguísticas estatais, que frequentemente se mostram insuficientes e fragmentadas. Assim, destaca-se a visão da língua como uma construção política. Uma língua que é proclamada para existir, resistir e reagir, conforme argumentam Franchetto e Balykova (2020). Essa perspectiva ressalta a importância de considerar as línguas não apenas como

instrumentos de comunicação, mas como expressões de identidade e resistência cultural em um contexto marcado por desafios históricos e contemporâneos.

5 CONSIDERAÇÕES FINAIS

No decorrer deste estudo, mergulhamos em um breve panorama histórico das políticas linguísticas coloniais que moldaram a trajetória linguística e cultural do Brasil desde os tempos do Grão-Pará até a Bahia em momentos atuais. Essa jornada nos permitiu compreender como a diversidade linguística foi profundamente afetada pelas forças colonizadoras e religiosas, que desempenharam um papel crucial na transformação das línguas e dos grupos étnicos indígenas presentes na região. Destacamos a importância das línguas da família Kariri, especialmente no contexto dos "tapuias" dos sertões baianos. A partir das análises de Demarquet (1991), evidenciamos a presença significativa dos indígenas desse grupo nas aldeias sob a administração dos padres jesuítas, o que ressalta a história sociolinguística de suas línguas durante o período colonial. Essa relação entre os povos indígenas e a atuação missionária não apenas moldou a dinâmica cultural da região, mas também resultou em profundas transformações nas línguas Kariri, refletindo a complexidade das interações sociais e linguísticas nesse contexto.

A compreensão dessas populações indígenas, suas línguas e sua (re)existência no interior da Bahia nos permite reconhecer a riqueza da diversidade linguística e cultural que permeou a região. As aldeias Natuba, Canabrava e Saco dos Morcegos tornaram-se marcos tangíveis dessa história compartilhada, em que os "tapuias" e suas línguas desempenharam um papel vital na configuração da paisagem linguística local. Ao reconhecer a resiliência dos povos indígenas, revelamos um capítulo fundamental da história do Brasil, que enriquece nossa compreensão sobre a formação étnica e linguística da nação.

A análise realizada ao longo deste trabalho evidencia que as línguas indígenas, especialmente as dos sertões baianos, foram alvo de um severo ataque à pluralidade linguística. Esse fenômeno foi exacerbado pelas missões jesuíticas, que proibiram os indígenas aldeados de praticarem seus hábitos culturais e impuseram o uso da língua do colonizador, tanto para os povos originários da costa quanto para aqueles do interior. Assim, as línguas em questão emergem como um reflexo desse cenário; ao dissecá-las, podemos identificar os vestígios sócio-históricos que revelam o panorama descrito.

Entretanto, é crucial reconhecermos que, apesar dos avanços acadêmicos na compreensão dessas questões, ainda há muito a ser explorado. É nosso dever continuar a investigar e respeitar a herança linguística e cultural desses povos, promovendo um diálogo

mais inclusivo e respeitoso em relação à história e identidade do Brasil. Ao fazê-lo, reconhecemos e valorizamos a sabedoria dos "tapuias" dos sertões baianos e o legado da família linguística Kariri, ampliando, assim, nossa compreensão da complexa tapeçaria linguística e cultural que forma a essência multiétnica e multilíngue do Brasil.

Referências

AZEVEDO, Gilda Maria Côrrea de. Língua Kiriri: descrição do dialeto Kipeá. Dissertação (Mestrado) – Instituto de Letras, Universidade de Brasília, Brasília-DF, 1965.

BANDEIRA, Maria de Lourdes. *Os Kariris de Mirandela:* Um grupo indígena integrado. 1 ed. Salvador : EDUFBA. 1972.

BORTONI-RICARDO, Stella Maris. *Português brasileiro, a língua que falamos* - São Paulo: Contexto, 2021.

CARDIM, Fernão. [1583-90]. *Tratados da terra e da Gente do Brasil*, transcrição, edição e notas de Ana Maria de Azevedo, Lisboa, Comissão Nacional para a Comemoração dos Descobrimentos Portugueses. 1997.

CÂMARA JR, Joaquim. Mattoso. *Introdução às Línguas Indígenas brasileiras*. Museu Nacional – Rio de Janeiro, 1965.

CARVALHO, Moacyr Ribeiro de. *Dicionário tupi (antigo) português*. Moacyr Ribeiro de carvalho - Salvador, 1987.

CONDE SILVESTRE, Juan Camilo. Problemas y principios. In: CONDE SILVESTRE, Juan Camilo. S*ociolinguística histórica*. Madrid, Gredos, 2007, p. 19-40.

COSTA, Vera Lúcia Anunciação. *A importância do conhecimento da variação lingüística*. Educar em Revista [online]. 1996, n. 12 [Acessado 3 Junho 2022], pp. 51-60. Disponível em: https://doi.org/10.1590/0104-4060.157. Epub 06 Mar 2015. ISSN 1984-0411. https://doi.org/10.1590/0104-4060.157.

CUNHA E SOUZA, Hirão Fernandes. *O Português Kiriri:* aspectos fônicos e lexicais na fala de uma comunidade do Sertão Baiano. Dissertação (Mestrado em Língua e cultura) Faculdade de Filosofia e Ciências Humanas, Universidade Federal da Bahia, Salvador/BA, 2011.

DANTAS, Beatriz Góis. et al. Os Povos Indígenas no Nordeste Brasileiro: um esboço histórico. In: CUNHA, Manuela Carneiro (org.) *História dos Índios no Brasil*. São Paulo: Companhia das Letras, Secretaria Municipal de Cultura, FAPESP, 1992.

DEMARQUET, Sonia de Almeida. *A resistência dos kariri de Mirandela à dominação*. In: Revista do serviço de ação cultural da terceira SUER - FUNAI. Série Etnohistória - nº 2 - Recife - janeiro, 1991.

DURAZZO, Leandro. *A garantia do seguimento indígena:* ciência ritual, rede proká e revitalização linguística no submédio São Francisco. Policromias – Revista de Estudos do Discurso, Imagem e Som, Rio de Janeiro, v. 6, n. 2, p. 421-460, 2021.

DURAZZO, Leandro; BONFIM, Evandro. A área etnolinguística das línguas encantadas: Repensando os conceitos de língua, falante, transmissão e território nas retomadas de língua indígena. *In*: SEVERO, Cristine; BUZATO, Marcelo. Cosmopolítica e linguagem. *Cosmopolítica e linguagem.-1. ed.-Araraquara: Letraria*, 2023. p. 129-141

FRANCHETTO, Bruna; BALYKOVA, Kristina (org.). *Índio não fala só tupi:* uma viagem pelas línguas dos povos originários no Brasil. 1. Ed. - Rio de Janeiro: 7Letras, 2020.

JESUS, Pollyana Macêdo de. *Notícias de viagem entre vilas e aldeias de indígenas sublevados na Bahia colonial:* edição semidiplomática, edição interpretativa e glossário. Dissertação (Mestrado em Língua e Cultura) - Universidade Federal da Bahia, Salvador-BA, 2022.

LEITE, Yonne & FRANCHETTO, Bruna. "500 anos de línguas indígenas no Brasil". In: Suzana A. M. Cardoso, Jacyra A. Mota, Rosa Virgínia Mattos e Silva (org.), *Quinhentos Anos de História Lingüística do Brasil*. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. Pp. 15-62. ISBN 85-232-0260-9.

MAMIANI, Luiz Vincencio. [1698]. *Catecismo da Doutrina Christãa na Lingua Brasilica da Nação Karirí*. Rio de Janeiro: Bibliotheca Nacional. (Edição fac-similar).

MATTOS E SILVA, Rosa Virgínia. *Caminhos da linguística histórica* - "ouvir o inaudível" - São Paulo, Parábola Editorial, 2008.

MATTOS E SILVA, Rosa Virgínia. *Ensaios para uma sócio-história do português brasileiro*. São Paulo: Parábola, 2004.

NANTES, Bernard de. *Catecismo da língua Kiriri*, publicado por Júlio Platmann. Leipzig, 1896.

NANTES, Martinho de. 1702. *Relação de uma missão no São Francisco*. São Paulo: Brasiliana, volume 368.Companhia Editora Nacional/MEC, 1979.

NASCIMENTO, Gabriel. *Racismo Linguístico*: os subterrâneos da linguagem e do racismo. Belo Horizonte: Letramento, 2019.

NIKULIN, Audrey; CARVALHO, Fernando Orphão. de. *Estudos diacrônicos de línguas indígenas brasileiras:* um panorama. MACABÉA – REVISTA ELETRÔNICA DO NETLLI, CRATO, V. 8., N. 2., 2019, p. 255- 305.

NIKULIN, Andrey. *Proto-macro-jê:* um estudo reconstrutivo. Tese (Doutorado em Linguística) – Universidade de Brasília, 2020.

NIMUENDAJU, Curt. (1944). Mapa Etno-Histórico de Curt Nimuendaju. [Mapa Interativo]. Disponível em: http://mapa-nimuendaju.eita.coop.br/. Acesso em: 01 maio 2024.

NOVAIS-CARNEIRO, Zenaide de Oliveira; ALMEIDA, Norma Lucia Fernandes de. Demografia e norma linguística no semi-árido baiano nos séculos XVIII e XIX: uma introdução. *In*: NEVES, Erivaldo Fagundes (org.). *Sertões da Bahia* - formação social, desenvolvimento Econômico, Evolução política e Diversidade cultural. Salvador: Editora Arcádia, 2011. p. 599-617.

RODRIGUES, Aryon Dall'Igna. *Línguas Brasileiras* para o conhecimento das línguas índigenas. Edições Loyola, São Paulo, Brasil, 1986.

RODRIGUES, Aryon Dall'Igna. *Línguas indígenas*: 500 anos de descobertas e perdas. São Paulo: DELTA, 1993. p. 99-100.

RODRIGUES, Aryon Dall'Igna. 2002. *Para o estudo histórico-comparativo das línguas Jê*. In Ludoviko dos Santos & Ismael Pontes (orgs.), *Línguas Jê: estudos vários*, p. 1-14. Londrina: UEL.

SANTANA, Rejane Cristine Carneiro; MENDES, Luan Oliveira. *O Legado Linguístico dos Tapuias no Sertão Baiano*. Légua & Meia, Brasil, n.11, v. 2, p. 78-93, 2020.Feira de Santana, v. 12, n. 1, 2020.

SANTANA, Letícia Suely de. *A herança linguística dos Kariri na aldeia Tumbalalá.* - Trabalho de Conclusão de Curso (TCC) - Departamento de Letras e Artes, Universidade Estadual de Feira de Santana, Feira de Santana-BA, 2023.

SANTOS, Fabricio Lyrio. *Da catequese à civilização:* colonização e povos indígenas na Bahia (1750-1800) – tese (Doutorado em História Social) - Faculdade de Filosofia e Ciências Humanas, Universidade Federal da Bahia, Salvador-BA, 2012.

SANTOS, Gustavo dos. *"Do multilinguismo generalizado ao multilinguismo localizado":* políticas de redução da diversidade linguística na História Social Linguística do Brasil. MACABÉA – REVISTA ELETRÔNICA DO NETLLI, CRATO, V. 8., N. 2., 2019, p. 237-254.

SIMAS, Hellen Cristina Picanço. LUCCHESI, Dante. *O português indígena sateré-mawé*: um projeto de pesquisa sociolinguística. Linguagem & Ensino, Pelotas, v. 23, n. 4, p. 1076-1096, out.-dez. 2020.

SOARES, Marília Facó. CARVALHO, Fernando Orphão de. *As hipóteses de Aryon Rodrigues:* validade, valor e papel no cenário dos estudos de línguas indígenas e de linguística histórica. D.E.L.T.A., 30 especial, 2014, p. 543-570.

SOARES, Paloma Reis; MATAREZIO FILHO, Edson Tosta. Terra árida, povo árido: desdobramentos sócio-históricos dos povos indígenas nas entranhas dos sertões-norte da Bahia. *In:* MATAREZIO FILHO, Edson Tosta; ARAÚJO, Silvana Silva de Farias; SANTOS, Gredson dos (org.). *Cultura e linguagem nos sertões baianos: narrativas afro-indígenas*. 1. ed. - Campinas, SP: Pontes Editores, 2025. p. 23-50.

VIVEIROS DE CASTRO, Eduardo. *Os pronomes cosmológicos e o perspectivismo ameríndio*. MANA 2(2):115-144, 1996.

http://doi.org/10.5007/1984-8420.2025e101203

CONSTRUÇÕES-WH NO PORTUGUÊS BRASILEIRO: "GÊMEOS" (DOUBLETS) SINTÁTICOS OU VARIAÇÃO ESTILÍSTICA[1]

WH-CONSTRUCTIONS IN BRAZILIAN PORTUGUESE: SYNTACTIC 'DOUBLETS' OR STYLISTIC VARIATION

Mary Aizawa Kato | Lattes | maryak@unicamp.br
Universidade Estadual de Campinas

Reumo: Este capítulo apresenta minha visão geral dos estudos diacrônicos e sincrônicos sobre as construções-wh no português brasileiro e discute o problema instigante dos aparentes "gêmeos" sintáticos, ou da opcionalidade, nesse domínio. Contudo, uma análise cuidadosa da ordem diacrônica em que esses padrões emergem, bem como da variação exibida em *corpora* contemporâneos, escritos e falados, leva-nos a concluir que os padrões coocorrentes não constituem "gêmeos" sintáticos.

Palavras-chave: Construções-wh; Português brasileiro; "Gêmeos" sintáticos; Opcionalidade; Diacronia; Movimento-wh; Clivagem.

Abstract: This chapter contains my overview of diachronic and synchronic studies on Brazilian Portuguese wh-questions, and discusses the intriguing problem of apparent syntactic doublets, or optionality, in this domain. However, a careful examination of the diachronic order in which these patterns appear, and the variation exhibited in contemporary written and spoken corpora lead us to conclude that the co-occurring patterns do not constitute syntactic "doublets."

Keywords: Wh-questions; Brazilian Portuguese; Syntactic "doublets"; Optionality; Diachrony; Wh-movement; Clefting.

1 Introdução

O Português Brasileiro (PB) contemporâneo apresenta variação nas construções-wh como no exemplo (1) (cf. Kato, 2018). À primeira vista, isto enfrenta um problema relativo ao parâmetro-wh, segundo o qual as línguas podem ser de dois tipos: com ou

[1] Este artigo é uma adaptação traduzida do capítulo *Morphological doublets in Brazilian Portuguese wh-constructions* de Kato (2018).

sem movimento-wh (Huang, 1982). Além disso, há um consenso de que a inversão VS é mais limitada no PB com verbos inergativos e transitivos (inter alia, Kato, 2020) mas os exemplos em (2) parecem todas ser perfeitamente aceitáveis.

(1) a. Você votou **em quem** em 1989?
b. **Em quem** você votou em 1989?

(2) a. Onde **dormem as crianças**? [Wh-VS] [2]
b. Onde **as crianças dormem**? [Wh-SV]
c. **As crianças dormem** onde? [Wh-in-situ]

(3) a. Afinal, quanto **vale a vida**? [Wh-VS]
b. Afinal, quanto **a vida vale**? [Wh-SV]
c. Afinal, **a vida vale quanto**? [Wh-in-situ]

O que temos no PB são construções-wh em variação, não só entre duas formas, mas até de três, o que não deveria ser permitido, dado o princípio "blocking effect" de Aronoff (1976), que exclui "doublets " sintáticos, que aqui traduzimos como "gêmeos" sintáticos.

De acordo com Kroch (1984, p. 181), "gêmeos" sintáticos são sempre reflexos de uma competição instável entre opções gramaticais mutuamente exclusivas, que normalmente acaba resultando em mudança gramatical. Também na visão minimalista, a variação/ opcionalidade em derivações gramaticais onde estão envolvidas operações de checagem não é esperada (Saito e Fukui, 1998). A tese que defendemos, todavia, é que "gêmeos" sintáticos podem ser encontrados atualmente em um indivíduo, não envolvendo mudança sintática.

As perguntas encaixadas, por outro lado, exibem um padrão uniforme, com constituintes frontais-wh, e sem inversão sujeito-verbo no Português Europeu (PE) ou no PB:

(4) a. A Ana está perguntando **onde** as crianças **comem**. PB PE
b. A Ana está perguntando **onde comem** as crianças. *PB *PE

Meu objetivo neste artigo é mostrar que as três variantes em (3) constituem reflexos de mudança, mas o que quero afirmar é que elas não constituem um "trigêmeo" e

[2] Sobre o exemplo (2) vale um comentário. Segundo Pilati (2016), a inversão VS é comum com verbos existenciais e inacusativos, mas não com verbos inergativos. Mas a autora mostra que a inversão é possível quando há um expletivo locativo (*aqui, aí* ou *lá).* Ex. *Alí Dormem os meninos./ Onde dormem os meninos?.*

podem ser analisadas como possíveis variantes estilísticas dentro da mesma gramática.

O artigo está organizado da seguinte forma: a Seção 2 contém uma visão geral dos dados diacrônicos em português; a Seção 3 contém a minha análise dos dados; e a última seção contém as conclusões.

2 O desenvolvimento diacrônico das construções-wh no PB

2.1 O período V2 no Português Clássico (PCl)

Até o final do século XVIII, ou período Clássico, as perguntas-wh em português tinham o padrão V2 (wh-VS) comum no Romance em geral, um padrão que continua a subsistir no Português Europeu Moderno (PE), mas só sobrevive como resíduo no PB. Além disso, a ordem VS que encontramos no século XX é apenas do tipo em (6) com sujeito DP (cf. Duarte e Kato, 2002), e não pronominal.

(5) a. Mas que me quer ele a mi? (XVI) PCl
b. Que fez vossa mercê? (XVIII) PCl
c. Que te disse ele? (XIX) PE
d. Mas que tenho eu com isso? (XIX) PB

(6) a. Onde andará a Neiva? (XIX - XX) PE PB
b. Que lhe disse o Honorato? (XIX - XX) PE ?PB

2.2 A construção clivada inversa do Português Moderno

A partir da segunda metade do século XVII, começou a surgir uma construção-wh a partir de uma numeração diferente, tanto no PB como no português europeu (PE), nomeadamente uma sentença clivada-inversa (ex. (7a)), tendo a cópula a segunda posição privilegiada e o verbo temático exibindo a ordem SV de sua oração encaixada (Duarte, 1992; Duarte e Kato, 2002). A gramática que produz a clivada inversa da pergunta-wh é a mesma da gramática V2, com movimento-wh, e a cópula na segunda posição. Considero isso uma espécie de gramaticalização, pois na gramática V2 qualquer verbo temático pode ser elevado a C, e no caso da clivada essa possibilidade é reduzida à cópula, um verbo funcional.

A partir do século XIX, a cópula deixa de apresentar consecutio temporum, exibindo a forma invariável (ex. (7b)), é novamente mais um caso de gramaticalização.

(7) a. Onde **foi que** as crianças dormiram? (XVIII) (clivada inversa PB/PE)
b. Onde **é que** as crianças dormiram? (XX) (clivada inversa PB/PE)

2.3 A construção wh-in-situ

Enquanto o PE tem uma ocorrência muito restrita de perguntas wh-in-situ (1,22% nos dados escritos), a segunda metade do século XIX vê, no PB, o aparecimento de uma quantidade muito maior (8,84 nos dados do corpus escrito de Kato e Mioto, 2005), alcançando 42% de ocorrências nos dados da fala de brasileiros do século XX (Lopes-Rossi, 1996). Neste ponto, devemos ressaltar que o PB distingue dois tipos de construções wh in-situ: a) aquele com entoação ascendente (⇑) (8a), com interpretação de eco, e aquele com entoação descendente (⇓) (8b), uma pergunta genuína (cf. Kato, 2013a; 2013b).

(8) a. Você é de **onde**?⇑ [Entoação ascendente] (PB falado)
b. Você pensava **em que** na sua vida?⇓ [Entoação descendente] (PB falado)

Embora o PB e o francês (FR) possam ser considerados tipos semelhantes de língua, licenciando tanto o wh- movido quanto o wh-in-situ, temos que explicar por que o PB permite (9b) e (9c), enquanto o francês não licencia os padrões paralelos (10b) e (10c) (Cheng e Rooryck, 2000; Vergnaud e Zubizarreta, 2005). Além disso, o francês tem a mesma entoação tanto para construções de eco quanto para construções genuínas in-situ.

(9) a. então você vai dizer **o que**?⇓ PB (XIX segunda metade)
b. E Maria pensa que João comprou **o que**?⇓ PB (XIX segunda metade)
c. Maria ama o livro que **quem** escreveu?⇓ PB [in-situ dentro de ilha]

(10) a. alors tu vas dire **quoi**?⇑ FR
b. *Marie pense que Jean a acheté **quoi**?⇑ FR [in-situ em sentença complemento]
c. *Marie aime le livre que **qui** a écrit?⇑ FR [in-situ dentro de ilha]

2.4 A construção-wh semi-clivada

Uma mudança que se segue à introdução da construção wh-in-situ, na segunda metade do século XIX, é o tipo de clivagem sem cópula, a semi-clivada, existente apenas no PB, e observada principalmente na língua falada do século XX por Lopes-Rossi (1996) e Kato e Duarte (2002).

(11) a. **O que que** você faz? (XX, PB falado)
b. **De quem que** é esse peixe? (XX, PB falado)

2.5 O aparecimento do padrão wh-SV

Na segunda metade do século XIX, o PB vê também o aparecimento da construção do tipo não-V2, nomeadamente o padrão wh-SV, que ocorre tanto na língua escrita como na falada, mas um tipo único entre as línguas românicas.

(12) a. **Onde** ele foi ? (XIX, PB escrito)
b. **Quanto** você ganha? (XX, PB falado)

Os brasileiros evitam a Clivada Reduzida (wh-que-SV) na linguagem escrita, enquanto a preferem na fala coloquial. O que os brasileiros usam, em vez disso, na linguagem escrita, é o tipo wh-SV, prescrito pelos gramáticos tradicionais como a norma brasileira prototípica. Este ponto será tratado em detalhes abaixo em 3.7.

Em suma, a diacronia das construções-wh no PB pode ser vista no quadro abaixo de Lopes Rossi (1996), com as seguintes abreviaturas (PA:Português Antigo; PCl Português Clássico):

Figura 1: Tipos de perguntas-wh no Português

PA & PCl XIV-XVIII	wh**VS**				
PE XVIII-XX	wh**VS**	wh **é que** VS/SV	% wh-**in-situ**		
PB XVIII-XX	-----	wh **é que** SV	wh-**in-situ**	wh **que SV**	Wh **SV**

Fonte: Adaptado de Lopes-Rossi (1996)

2.6 Uma célula ausente entre o wh-in-situ e o wh-que SV

Assumindo que as construções de Foco e as construções wh- têm derivações paralelas devido ao fato de ambas checarem suas características com o mesmo núcleo Focus, Kato e Ribeiro (2009) notaram uma lacuna entre a construção in-situ e o padrão semi-clivado. O que descobriram foi que no período V2 o tipo de clivagem era do tipo inverso, com a cópula em segunda posição, mas com o advento do wh-in-situ o tipo de clivagem

também muda do tipo inverso para um tipo canônico de clivagem, com a cópula não mais em segundo posição.

(13) a. E ISSO **é que** se chama postura, ou posição reta. (XVII, clivada invertida)
b. E QUANDO **é que** são relativas? (XVII, clivada invertida)

(14) a. **É** a Maria **que** tá tocando violão. (XX, clivada canônica)
b. **É** quem **que** tá tocando o violão? (XX, clivada canônica)[3]

Podemos modificar a Figura anterior de Lopes-Rossi, adicionando o padrão da clivada canônica entre o padrão wh-in-situ e o padrão wh-VS (ver Figura 2).

Figura 2: Tipos de perguntas-wh no PB através dos tempos

PA & PCl XIV-XVIII	wh**VS**					
PE XVIII-XX	wh**VS**	Wh é que VS/SV	%(wh-in-situ)			
PB XVIII-XX	% wh**VS**	Wh **é que** SV	wh-in-situ	**é** Wh **que**	Wh **que** SV	Wh **SV**

3 Análises sintáticas

3.1 Sentenças "téticas" e sentenças "categóricas"

Assumirei, com Kuroda (1972), que as línguas distinguem entre sentenças categóricas e téticas. No PE esta distinção é expressa através da ordem de palavras SV e VS (Martins, 1994), um contraste que se afirma ter sido perdido no PB (Britto, 2000), segundo a qual a ordem SV(X) expressa um julgamento tético, enquanto o o julgamento categórico é codificado como uma estrutura de deslocamento à esquerda (LD).

(15) a. O gato está a dormir SV (categórica) PE
b. Está o gato a dormir VS (tética) PE

(16) a. O gato, ele está dormindo DE (categórica) PB
b. O gato está dormindo SV (tética) PB

[3] O tipo canônico de pergunta wh- não foi encontrado nos corpora consultados. É por isso que encontramos a lacuna na Figura 1 de Lopes-Rossi (1996). Kato (2013b; 2014) encontrou exemplos em diálogos entre crianças e mães. Esses exemplos foram facilmente aceitos por meus alunos de graduação.

3.2 O tipo mais antigo de perguntas-wh no Português: a construção V2

Neste artigo, também presumo que as perguntas-wh são sentenças téticas e, portanto, refletirão a ordem VS das sentenças téticas declarativas no PCl. No PE o padrão ainda faz parte do vernáculo. Ver (18a)' para a derivação de uma sentença declarativa (Kato, 1987; Martins, 1994). Quanto a uma pergunta-wh, a derivação em (17b)' reflete aquela de uma sentença declarativa, relativa à posição do verbo finito.

(17) a. Está o gato a dormir no jardim. PCl PE
b. Onde está o gato a dormir? PCl PE

(17)'[4] a.[$_{FocP}$ [$_{F}$ está [$_{IP}$ ~~está~~ [$_{VP}$ o gato ~~está~~ [$_{VP}$ ~~está~~ a dormir no jardim]]]]
b.[$_{FocP}$ **onde**[$_{F}$ está [$_{TP}$ ~~está~~ [$_{vP}$ o gato ~~está~~ [$_{VP}$ ~~está~~ a dormir ~~onde~~]]]]]]

A minha análise assume, com Uriagereka (1995) e Kato e Raposo (1996), que o verbo finito no Romance Ibérico Ocidental se desloca para FocP, na periferia esquerda da sentença, o que também acontece no Português Europeu Moderno (PE)[5].

O PB, por outro lado, perdeu o tipo V2 nas construções-wh com verbos temáticos, e o aparente tipo V2 contemporâneo foi analisado em Kato (1993) como uma inversão estilística em francês (Kayne e Pollock 1978), ou um deslocamento à direita com sujeito nulo em Duarte e Kato (2002). Isso explica porque o sujeito geralmente é um nome ou um DP[6].

(18) Onde (ela) está essa mulher? XX BP[7]

(18)' [$_{FocP}$ onde [$_{TP}$ [$_{TP}$ (ela) está [$_{vP}$ (ela) ~~está onde~~]] essa mulher]

3.3. A pergunta clivada-wh inversa (wh-é que SV) como uma construção-V2 gramaticalizada

Até o período antigo e clássico (período V2), o elemento-wh em português era seguido por qualquer verbo temático ou auxiliar, fato que mudou no século XVIII na me-

[4] Vou ignorar a projeção ForceP em ambas as derivações.
[5] Uma versão diferente é oferecida em Barbosa (2001), para quem o verbo está em T e o sujeito *in-situ em Spec,V.*
[6] Na verdade, como o tipo deslocado à direita é um tipo de fenômeno similar a desdobramento clítico, a derivação real deve juntar o pronome *ela* com *essa mulher* (ela essa mulher) no sentido de um "big DP como em Kayne (2001), caso em que esse DP depende de *ela*, um nominativo.
[7] A presença ou a ausência do sujeito pronominal não faz diferença no sentido aqui.

dida em que a cópula passou a ter a segunda posição privilegiada, processo que podemos considerar um tipo de gramaticalização.

(19) a.	Que **compraste** (tu)?	PCl	PE	*PB
b.	Quando **chegaram** as crianças?	PCl	PE	%PB
c.	Onde **podem** as crianças estar a dormir?	PCl	PE	*PB

(20) a.	Que **foi** que tu compraste?	PE	*PB
b.	Quando **é** que chegam as crianças?	PE	PB

A ordem do verbo temático com o sujeito não importa mais, pois o primeiro ocorre dentro de uma sentença encaixada. O que agora passa para F é a cópula.

(20)' a. $[_{FocP}$ que $[_{F}$ foi $[_{TP}$ ~~foi~~ $[_{vP}$ ~~foi~~ $[$ ~~foi~~ $[_{CP}$ que tu compraste ~~que~~]]]]]]

b. $[_{FocP}$ quando $[_{F}$ é $[_{TP}$ ~~é~~ $[_{vP}$ ~~é~~ $[$ ~~é~~ $[_{CP}$ que chegam as crianças ~~quando~~]]]]]] ~~quando~~]]]]]]

Nossa previsão é que, embora a clivada inversa seja a forma vernácula no PE, no PB ela sobrevive na linguagem escrita, sendo substituída pela forma semi-clivada mais curta na linguagem falada.

3.4 Do movimento-wh para uma construção falsa de movimento

Neste artigo, assumirei o modelo cartográfico, nomeadamente uma organização hierárquica dos constituintes sintáticos, tanto ao nível das estruturas oracionais (IP/TP), como no nível periférico da sentença (CP) (21a1), e no nível sentencial médio (21a2):

(21) a1. $[_{Force\ P}$ ...$[_{TopP}$...$[_{FocP}$...$[_{TopP}$... $[_{IP}$ ..... (Rizzi 1997)

a2. $[_{CP}$.... $[_{IP}$.....$[_{TopP}$...$[_{FocP}$...$[_{TopP}$ ... $[_{vP}$ $[_{VP}$]]]]] (Belletti 2004)

Também assumirei, no ForceP, um Operador de Ato de Fala, que pode ter uma realização morfológica, como no japonês -ka para pergunta, ou um operador vazio $\sum 1$ que resulta em um padrão prosódico ascendente, como no francês e no português brasileiro. Também assumo que em PB pode haver dois operadores vazios em ForceP, um $\sum 1$ para entonação ascendente e um $\sum 2$ para entonação decrescente.

A mudança estrutural mais marcante que ocorreu no PB em relação às construções-wh foi o aparecimento das construções wh-in-situ como questões reais, no século XX (Lopes-Rossi 1996). Até o século XIX, a projeção FocP era exclusivamente na periferia sentencial, licenciando tanto a construção V2 quanto as sentenças-wh com clivagem inversa.

No século XX, a projeção FocP mais alta deixa de ser projetada e uma posição FocusP inferior, adjacente à esquerda de vP, é ativada (cf. Belletti, 2004). O aparente 'wh-in-situ' no PB é proposto aqui, segundo Kato (2013a), para derivar de um curto movimento do constituinte-wh para a posição FocP adjacente ao vP como em (22'a). É importante ressaltar que as construções que estamos analisando são questões reais, com entonação decrescente (⇓), determinada pelo operador Σ2 em ForceP, e não questões de eco, que possuem entonação ascendente (⇑), determinada por ∑1. A questão do eco em (22b)', por outro lado, é analisada como o resultado de todo o TP movido para Spec, ForceP (cf. Kayne, 1994) interpretado como o pressuposto da sentença.

(22) a. As crianças dormem **onde**? ⇓ [pergunta genuina]
b. As crianças dormem **onde**? ⇑ [pergunta-eco]

(22)' a.' [$_{ForceP}$ Σ2 [$_{TP}$ as crianças dormem [$_{FocusP}$ **onde**[$_{vP}$ ~~as criança~~ [VP ~~dormonde~~]]]]]]
b.' [$_{ForceP}$ Σ1 [$_{TP}$ as crianças dormem **onde**] ∑1 [$_{TP}$ [$_{TP}$ ~~as crianças dormem~~
[$_{vP}$ ~~as crianças dormem~~ [$_{VP}$ ~~dormem~~ ~~onde~~]]]]

A evidência de que as falsas construções wh-in-situ sofreram movimento é fornecida por Kato (2013a) usando múltiplos adjuntos adverbiais. Nas orações declarativas, vários adjuntos adverbiais vêm empilhados no final da frase. Quando vários adjuntos-wh são movidos, eles aparecem em coordenação. As palavras-wh em wh-in-situ falsas também aparecem em coordenação, o que indica que foram movidas. As questões de eco in-situ, por outro lado, podem ter apenas um constituinte wh in-situ (cf. (23e)), com duas entonações independentes.

(23) a. Maria nasceu **na Bahia em 1980**
b. **Onde e quando** Maria nasceu?
c. Maria nasceu **onde e quando**?⇓
d. * Maria nasceu **onde quando**?⇑
e. Maria nasceu **onde**? ⇑ **Quando**? ⇑

Mas as falsas questões "in-situ" como em (24a) parecem competir com o resultado do movimento-que como em (24b). Devemos chamar a atenção para o fato de que não precisa haver diferença de pressupostos entre as duas formas no PB. Essa perfeita equivalência funcional pode ser vista num verso de um samba brasileiro (24c).

(24) a. Você parou **por que**?
b. **Por que** você parou?
c. **Porque** (você) parou? (Você) Parou **porque**?

O que resta a ser respondido em relação a (24a) e (24b) é: os movimentos-wh curtos e longos competem gramaticalmente no vernáculo do PB? Mas lembremos que ambos são inovações do século XX e não podem ser atribuídos à instabilidade da mudança linguística. Discutiremos esse ponto mais adiante, na seção 3.6.

3.5 De clivadas-wh inversas a clivadas-wh canônicas

Ao mesmo tempo que surgem as falsas perguntas wh-in-situ, a pergunta clivada inversa como (25a), existente em ambas as variedades desde o final do século XVII, começa a exibir um tipo competitivo de clivada em BP, a clivada canônica (25b), raramente encontrada em corpora do século XX:

(25) a. **Onde é que** as crianças dormem? [clivada inversa] (PB/PE XVII a XX)
b. **É onde que** as crianças dormem? [clivada canônica em perguntas-wh) (PB XX)

Analisamos o novo tipo, a clivada canônica, como consequência da ativação da posição inferior de FocP na oração da cópula, para a qual o elemento-wh se move.

(25)' b [$_{ForceP}$ Σ2 [$_{TP}$ **É** [$_{FocusP}$ **onde** [$_{vP}$ ~~é~~ [$_{FiniteP}$ que [$_{TP}$ as crianças dormem [$_{vP}$ ~~as crianças dormem~~ [$_{VP}$ ~~dormem onde~~]]]]]]

3.6 As inovadoras clivadas reduzidas wh-que no século

Análises anteriores da clivada reduzida propuseram que elas exibiam uma espécie de "Comp duplo preenchido" como em (26)' (cf. Mioto e Figueiredo Silva, 1995; Hornstein, Nunes e Grohman, 2005).

(26) **Onde que** as crianças dormiam

(26)' [$_{CP}$ Onde [$_{C}$ que [$_{TP}$ as crianças dormiam [$_{vP}$ ~~as crianças dormiam~~ [$_{VP}$ ~~dormiam onde~~]]]]]

Kato e Raposo (1996) propuseram, ao contrário, que a clivada reduzida seria a redução da pergunta clivada inversa.

(26)'' [$_{FocP}$ Onde [$_{Foc}$ (é) [$_{TP}$ ~~é~~ [$_{vP}$ **é** [$_{CP}$ que [$_{TP}$ as crianças dormiam [$_{vP}$ ~~as crianças dormiam~~ [$_{VP}$ ~~dormiam onde~~]]]]]

Kato (2007) propõe, contudo, que a cópula no PB só pode ser apagada na posição inicial da sentença.

(27) a. O seu cabelo *(é) lindo.
b. (É) lindo o seu cabelo.

Uma conclusão natural é que a clivada reduzida não deriva de uma clivada inversa, mas sim de uma clivada canônica, onde a cópula aparece na posição inicial da frase. Podemos afirmar também que o apagamento seria possível depois que a cópula perde o consecutio temporum, tornando-se invariável, como acontece com a construção da clivada inversa:

(28) a. **Era** onde **que** as crianças dormiam? [Clivada canônica com consecutio Temporum]
b. **É** onde **que** as crianças dormiam? [Gramaticalização do tempo]
c. **Ɇ** Onde **que** as crianças dormiam? [Apagamento da copula em PF)

A derivação de (28c) se daria da seguinte forma:

(28)' c. [$_{ForceP\Downarrow}$ Σ2 [$_{TP}$ (**É**) [$_{FocusP}$ **onde** [$_{vP}$ **é** [$_{FiniteP}$ que [$_{TP}$ as crianças dormiam [$_{vP}$ ~~as~~ ~~crianças dormiam~~ [$_{VP}$ ~~dormiam onde~~]]]]]]

3.7 O aparecimento de WH-SV, o padrão mais recente

A análise apresentada em Kato e Raposo (1996) e também em Duarte e Kato (2002) relacionou o padrão WH-SV à perda do Sujeito Nulo. Enquanto nas línguas com

sujeito nulo o sujeito, quando expresso, era pós-verbal nas perguntas- wh, no PB o enfraquecimento do sistema flexional levou à criação de um paradigma de pronomes fracos (cf. Kato, 1999; Kato, 2000) que poderiam ocupar Spec, T[8].

(29) a. Onde está **ele** dormindo? EP
b. Onde **ele** está dormindo? BP

(29)' a. [$_{FocP}$ onde [está [$_{TP}$ ~~está~~ [$_{vP}$ ele ~~está~~ [$_{VP}$ ~~está dormindo onde~~]]]]]
b. [$_{FocP}$ onde [[$_{TP}$ ele está [$_{vP}$ ele ~~está~~ [$_{VP}$ ~~está dormindo onde~~]]]]]

Contudo, os fatos diacrônicos, e também a distribuição sincrônica dos padrões, permitem-nos considerar uma hipótese diferente. Lembremos que o padrão V2 era o mais antigo na história do Português, enquanto o padrão Wh-SV era o padrão mais recente, aparecendo concomitantemente ao padrão Wh-que-SV.

O interessante, aliás, é que os brasileiros evitam a Clivada-reduzida (WH-que-SV) na linguagem escrita, enquanto a preferem na fala coloquial. No corpus escrito Kato e Mioto (2005) encontraram apenas um caso de clivada reduzida, enquanto no corpus oral encontraram 18 casos. O que os brasileiros utilizam em vez disso, na linguagem escrita, é o tipo WHSV, reconhecido pelos gramáticos como a norma brasileira prototípica.

A utilização de WHSV como variante estilística de WH-que-SV me leva a considerar a possibilidade de que a partir do tipo clivada reduzida, uma regra estilística possa apagar o complementizador que em (30), resultando em (30)' em PF. Uma observação interessante em Duarte (1992) foi que a ordem SV aparece primeiro com perguntas adjuntas wh, e pode-se perguntar por quê. A minha hipótese é que com este tipo de questões temos o local ideal para aplicar a regra estilística da haplologia (apagar sílabas semelhantes).

(30) a. De **que que** os meninos riam? > (30)' a. **De que** os meninos riam?
b. Para **que que** você quer isso? > b. **Para que** você quer isso?
c. Por **que que** eles pararam? > c. **Por que** eles pararam?

Nos casos em que a palavra- wh não é semelhante ao complementizador que, podemos sugerir que o apagamento é o resultado da analogia com os casos adjuntos.

Mesmo a estrutura de deslocamento à direita em (31), com a ordem wh-VS, pode resultar da mesma estrutura gramatical, nomeadamente uma clivada canônica.

[8] Ordoñez e Olarrea (2006) oferecem uma explicação similar para o Caribenho, que também perdeu o padrão V2.

(31) a. **Onde** foi a Suzana?

b. (É) **onde** (que) (ela)foi a Suzana? [Deslocamento à direita). BP

4 Conclusões

A principal mudança estrutural do Português Antigo para o PB moderno foi o lugar da projeção do FocusP.:

Focus inicial

(32) a. **O que** comprou a Maria?

b. **O que é que** a Maria comprou?

(32)' a. [$_{ForceP}$ Σ2 [$_{FocP}$ **o que** comprou [$_{TP}$ ~~comprou~~ [$_{vP}$ a M. [$_{VP}$ ~~comprou o que~~]]]] EP

b. [$_{ForceP}$ Σ2 [$_{FocP}$ **o que é** [$_{TP}$ ~~**é**~~ [$_{VP}$ é [$_{CP}$ **que** [$_{IP}$ a M. comprou ~~**o que**~~ ?]]]]

PE PB

Focus medial/baixo

(33) a. A Maria comprou **o que**?

b. É **o que** que a Maria comprou?

(33)' a. [$_{ForceP}$ Σ2 [$_{TP}$ a Maria comprou [$_{FocP}$ **o que** [$_{vP}$ ~~a M.~~ [$_{VP}$ ~~comprou o que~~]]]]

PB %PE

b. [$_{ForceP}$ Σ2 [$_{TP}$ (**é**) [$_{FocP}$ o que [$_{VP}$ **é** [$_{CP}$ **que** [$_{IP}$ a M. comprou ~~**o que**~~]]]

PB *PE

Figura 3: Resumo das conclusões deste capítulo: Do Português Antigo para o PB, passando pelo Português Clássico e o Português Europeu

Movimento-wh-longo		Movimento-wh-curto			
V2	Clivada Inversa	Falso wh-in-situ	Clivada canônica		
wh-VS	wh-é que-SV	SVwh-	É wh-que-SV	wh-que-SV	wh-SV
PA PCl PE	PCl PE **PB**	PB			

O único caso de movimento-wh longo no PB é a clivagem invertida, que consideramos um resíduo da gramática antiga, enquanto essa mesma construção é a mais vernacular no PE.

Concluo, finalmente, que no vernáculo do PB não há movimento-wh longo, nem uma construção real de wh-in-situ, exceto em perguntas-eco. Temos apenas um movimento curto em todos os casos. Todas as outras variantes resultam de redução fonológica e, portanto, não produzem "gêmeos" sintáticos, mas variantes estilísticas.

Referências

ARONOFF, M. Word Formation in Generative Grammar. *Linguistic Inquiry Monographs 1*. Cambridge: MIT Press, 1976.

BARBOSA, P. On Inversion in Wh-questions. *In:* HULK, A.; POLLOCK, J.-Y (Ed.) *Romance Inversion*. New York: Oxford, 2001, p. 2-59.

BELLETTI, A. Aspects of the low IP area. *In:* RIZZI, L. (Ed.) *The Cartography of Syntactic Structures*. New York: OUP, 2004, p. 16-51.

BRITTO, H. Syntactic codification of categorical and thetic judgments in Brazilian Portuguese. *In: KATO*, M. A.; NEGRÃO, E. V. (Ed.) *Brazilian Portuguese and he Null Subject Parameter*. Frankfurt: Vervuert/Iberoamericana, 2000, p. 195-222.

CHENG, L.; J. ROORYCK, J. Licensing wh-in-situ. *Syntax*, 3,1, p. 1-19, 2000.

DUARTE, M. E. L. A perda da ordem V(erbo) S(ujeito) no português do Brasil. *DELTA*. Número Especial. P. 37-52, 1992.

DUARTE, M. E. L; KATO, M. A. A Diachronic Analysis of Brazilian Portuguese Wh-Questions. *Santa Barbara Portuguese Studies*, v. VI, p. 326-339, 2002.

HORNSTEIN, N.; NUNES, J.; GROHMAN, K. K. *Understanding Minimalism*. Cambridge/New York: Cambridge University Press, 2005.

HUANG, C. T. J. *Logical Relations in Chinese and the Theory of Grammar*. MIT: PH.D. Dissertation, 1982.

KATO, M. A. Inversão da ordem SV em interrogativas no português. *DELTA*.,3, 2, p. 243-252, 1987.

KATO, M. A. The distribution of null and pronominal objects in Brazilian Portuguese. *In: ASH*BY, W.; MITHUN, M.; PERISSINOTO, G.; RAPOSO, E. (Ed.) *Linguistic Perspectives on the Romance Languages*: Selected Papers from the XXI Linguistic Symposium on Romance Languages (Currents Issues in Linguistic Theory Series), Amsterdam: John Benjamins, 1993, p. 225-235.

KATO, M. A. Strong pronouns and weak pronominals in the null subject parameter. *PROBUS*, Berlin, Editora Mouton de Gruyter, 1999, p. 1-37.

KATO, M. A. The partial pro-drop nature and the restricted VS order in Brazilian Portuguese. *In:* KATO, M. A.; E.V. NEGRÃO, E. V. (Ed.) *Brazilian Portuguese and the Null subject Parameter*. Frankfurt: Vervuert/ Iberoamericana, 2000a, p. 223-258.

KATO, M. A. A restrição de mono-argumentalidade da ordem VS no português do Brasil. *Fórum Lingüístico*, Florianópolis, 2,1, p. 97-127, 2000b.

KATO, M. A. Free and dependent small clauses in Brazilian Portuguese. *DELTA*, 23 (Especial): Homenagem a Lucia Lobato, 85-111, 2007.

KATO, M. A. Deriving wh-in-situ through movement. *In:* CAMACHO-TABOADA, V.; GIMENEZ-FERNANCEZ, A.; MARTIN-GONZALES, J.; REYES-TEJEDOR, M. (Ed.) . 175-191. Amsterdam/Philadelphia: John Benjamins, 2013a.

KATO, M. A. WH-movement in the history of Brazilian Portuguese. STIL, vol 5. 59-76, 2013b.

KATO, M. A. Focus and WH-questions in Brazilian Portuguese. *In: DIO*N, N.; LAPIERRE, A.; CACOULLOS, R. T. (Ed.). *Linguistic variation*: Confronting fact and theory. New York/London: Routledge, 2014, 111-130.

KATO, M. A. Morphological doublets in Brazilian Portuguese wh-constructions. *In:* REPETTI, L.; ORDOÑEZ, F. (Ed.) *Romance Languages and Linguistic Theory*. Amsterdam/Filadelfia. John Benjamins, 2018, p. 135-152.

KATO, M. A.; MIOTO, C. A multi-evidence study of European and Brazilian wh-questions. *In: KEP*SER, S.; REIS, M. (Ed.) *Linguistic evidence*: empirical, theoretical and computational perspectives. Berlin & New York: Mouton De Gruyter, 2005, p. 307-328.

KATO, M. A.; RAPOSO, E. European and Brazilian word order: questions, focus and topic constructions. *In: PAR*ODI, C.; QUICOLI, A.C.; SALTARELLI, M.; ZUBIZARRETA, M. L. (Ed.) *Aspects of Romance Linguistics*. Washing ton: Georgetown U.Press, 1996, p. 267-277.

KATO, M. A.; RIBEIRO, I. Cleft sentences from Old Portuguese to Modern Brazilian Portuguese. *In:* DUFTER, A.; JACOB, D. (Ed.). *Focus and Background in Romance Languages.* Amsterdam /Philadelphia: John. Benjamins, 2009, p. 123-154.

KAYNE, R. *The Antysymmetry of Syntax*. Cabridge, Mass: The MIT Press, 1994.

KAYNE, R. Pronouns and their antecedents. *In: EPS*TEIN, S.; SEELY, D. (Ed.) *Derivation and explanations in in the Minimalist program*. Malden, M A: Blackwell, 2001, p. 133-166.

KAYNE, R.; POLLOCK, J-Y. Stylistic Inversion, Successive Cyclicity and Move NP in French. *Linguistic Inquiry*, 9: 595-621, 1978.

KROCH, A. Morphosyntactic variation. *In:* BEALS, K. et al (Ed.) Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory, 1994, p. 180-201.

LOPES-ROSSI, M. A. 1996. A Sintaxe Diacrônica das Interrogativas-Q do Português. UNICAMP: Tese de doutorado, 1996.

KURODA, S-Y. The categorical and the thetic judgment. *Foundations of Language*, 9, p. 153-185, 1972.

MARTINS, A. M. 1994. *Clíticos na História do Português*. Universidade de Lisboa: Tese de Doutorado, 1994.

MIOTO, C.; FIGUEIREDO SILVA, M. C. Wh que = Wh é que? *DELTA*, 11, Vol 2, p. 301-311, 1995.

ORDOÑEZ, F.; OLARREA, A. Microvariation in Caribbean/non Caribbean Spanish. *Probus*, v. 18.3. p. 59-97, 2006.

PILATI, E. Sobre a ordem verbo-sujeito no português brasileiro: 30 anos em mirada crítica. *Revista Linguística*, v. 12, n. 2. p.183-205, 2016.

RIZZI, L. The Fine Structure of the Left Periphery. *In: HAE*GEMAN, L. (Ed.) Elements of Grammar. Dordrecht: Kluver, 1997, p. 281-337.

URIAGEREKA, J. An F position in Western Romance. *In: KISS*, K. E. (Ed.) Discourse Configurational Languages, New York/Oxford: Oxford University Press, 1995, p. 153-175.

VERGNAUD, J. R.; ZUBIZARRETA, M. L. The Representation of Focus and its Implications: towards an alternative account of some 'intervention effects. *In: BRO*EKHUIS, H.; CORVER, N.; HUYBREGTS, R.; KLEINHENZ, U.; KOSTER, J. (Ed.) Organizing Grammar- Linguistic Studies in Honor of Henk van Riemsdijk, 641-60. Berlin/ New York: Mouton de Gruyter, 2005.

CC BY NC