



UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA E CULTURA

LEANDRO ALMEIDA DOS SANTOS

**PRA EU E PRA MIM FAZER: TRAÇOS SOCIODIALETAIS NAS
CAPITAIS BRASILEIRAS**

Salvador
2023

LEANDRO ALMEIDA DOS SANTOS

**PRA EU E PRA MIM FAZER: TRAÇOS SOCIODIALETAIS NAS
CAPITAIS BRASILEIRAS**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Língua e Cultura do Instituto de Letras da Universidade Federal da Bahia – UFBA – como requisito parcial para obtenção do grau de Doutor em Língua e Cultura.

Orientadora: Profa. Dra. Silvana Soares Costa Ribeiro

Coorientador: Prof. Dr. Rerisson Cavalcante de Araújo

Salvador

2023

Ficha catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA),
com os dados fornecidos pelo(a) autor(a).

Santos, Leandro Almeida dos

Pra eu e Pra mim fazer: traços sociodialetais nas
capitais brasileiras / Leandro Almeida dos Santos. --
Salvador, 2023.
285 f.

Orientadora: Silvana Soares Costa Ribeiro.

Coorientador: Rerisson Cavalcante de Araújo.

Tese (Doutorado - Língua e Cultura) -- Universidade
Federal da Bahia, Instituto de Letras, 2023.

1. Dialetoлогия.. 2. Projeto NURC.. 3. Projeto
ALiB.. 4. Pronomes.. 5. Sociolinguística.. I. Ribeiro,
Silvana Soares Costa. II. Araújo, Rerisson
Cavalcante de. III. Título.


LEANDRO ALMEIDA DOS SANTOS

PRA EU E PRA MIM FAZER: TRAÇOS SOCIODIALETAIS NAS CAPITAIS BRASILEIRAS


Tese de Doutorado apresentada ao Programa de Pós-Graduação em Língua e Cultura do Instituto de Letras da Universidade Federal da Bahia – UFBA – como requisito parcial para obtenção do grau de Doutor em Língua e Cultura.

BANCA EXAMINADORA

SILVANA SOARES COSTA RIBEIRO – Orientadora
Doutora em Letras e Linguística, Universidade Federal da Bahia
Universidade Federal da Bahia


Documento assinado digitalmente
 SILVANA SOARES COSTA RIBEIRO
Data: 20/03/2024 08:27:37-0300
Verifique em <https://validar.iti.gov.br>

RERISSON CAVALCANTE DE ARAÚJO – Coorientador
Doutor em Linguística, Universidade de São Paulo
Universidade Federal da Bahia


Documento assinado digitalmente
 RERISSON CAVALCANTE DE ARAUJO
Data: 25/03/2024 16:12:36-0300
Verifique em <https://validar.iti.gov.br>

JACYRA ANDRADE MOTA
Doutora em Letras Vernáculas, Universidade Federal do Rio de Janeiro
Universidade Federal da Bahia


GREDSON DOS SANTOS
Doutor em Letras e Linguística, Universidade Federal da Bahia
Universidade Federal da Bahia

Documento assinado digitalmente
 GREDSON DOS SANTOS
Data: 26/03/2024 11:20:42-0300
Verifique em <https://validar.iti.gov.br>

ELISÂNGELA DOS PASSOS MENDES
Doutora em Língua e Cultura, Universidade Federal da Bahia
Instituto Federal da Bahia

Documento assinado digitalmente
 ELISANGELA DOS PASSOS MENDES
Data: 02/04/2024 17:46:45-0300
Verifique em <https://validar.iti.gov.br>

CLÉZIO ROBERTO GONÇALVES
Doutor em Linguística, Universidade de São Paulo
Universidade Federal de Ouro Preto

Documento assinado digitalmente
 CLEZIO ROBERTO GONCALVES
Data: 29/03/2024 22:02:23-0300
Verifique em <https://validar.iti.gov.br>

MARCELA MOURA TORRES PAIM (Suplente)
Doutora em Letras e Linguística, Universidade Federal da Bahia
Universidade Federal Rural de Pernambuco

LILIAN TEIXEIRA DE SOUSA (Suplente)
Doutora em Linguística, Universidade Estadual de Campinas
Universidade Federal da Bahia

NORMA LÚCIA DE ALMEIDA (Suplente)
Doutora em Linguística, Universidade Estadual de Campinas
Universidade Estadual de Feira de Santana

VALTER PEREIRA ROMANO (Suplente)
Doutor em Estudos da Linguagem, Universidade Estadual de Londrina
Universidade Federal de Santa Catarina

AGRADECIMENTOS

A priori, essa parte da tese é, sem dúvidas, a mais complexa e, também, a mais prazerosa, devido aos movimentos de ida e de vinda feitos para recordar as *personas* e os momentos envolvidos em minha trajetória.

Escrever uma tese de doutorado é estar numa ilha, deserta e bonita. Não é uma tarefa simples, sobretudo, quando eclode um período pandêmico. A solidão e os medos se potencializaram, o que é natural, mas também as inspirações brotaram juntamente com a esperança. E, como o movimento feito pelas marés, houve momentos vazios e cheios.

Preciso agradecer, embora as palavras não sejam suficientes. Quero usá-las para expressar o mar de emoção, gratidão e alegria que fui tomado, ao fazer um mergulho retrospectivo íntimo.

Vale ressaltar que escrever uma tese em um período de pandemia foi um desafio, haja vista as dificuldades ocasionadas, devido à situação de saúde coletiva. Nesse momento, tivemos que enfrentar novos cenários, tais como o distanciamento social, fechamentos de universidades e de outras instituições, além de diversos protocolos de orientações e cuidados, a fim de que ficássemos protegidos. Por tais razões, muitas pesquisas tiveram os prazos dilatados, inclusive a minha.

Pelos mares afetivos

Deus, sem o Senhor, nada disso seria possível. Pela sabedoria e forças advindas do seu amor, muito obrigado.

Aos meus pais, sempre presentes, Lindaura e Raimundo, pelo carinho e pela paciência, além de todo o suporte necessário para que esse momento fosse concretizado. Com toda certeza, vocês são o meu porto seguro, o meu lugar de descanso e de paz.

Ao meu irmão, Romilson, por acreditar e, na curiosidade, incentivar com pequenos gestos e poucas palavras.

Ao meu sobrinho, Ravi, pelas visitas lindas, cheias de alegria e encanto, que me fizeram parar e prestar a atenção sobre a escrita de outra tese, a tese do amor, essa que deve ser contínua e diária.

Aos familiares que, mesmo distantes fisicamente, vibram com as minhas vitórias pessoais e acadêmicas.

Isamar, minha amiga, aquela que sempre foi meu ombro amigo acadêmico, desde a minha chegada na pesquisa. Eu aprendo muito contigo, muito e sempre.

Aos amigos que me conhecem bastante, por essa amizade bonita que nasceu nos recintos do Projeto ALiB e, felizmente, se estende até hoje: Vitor, Flávia e Geiele.

As meninas cuja amizade é fruto das salas de aula da graduação e das salas do Projeto ALiB: Márcia, Luciana, Isabela, Eliana, Gracy, Grazi e Vívian.

Aquelas que, com muito carinho e respeito, ultrapassaram as barreiras do universo acadêmico, com trocas valiosas e muitas horas de escuta e desabafos: Thamiris e Carla Elisa.

Aquelas que, desde 2019, criamos uma rede de apoio salvadora em muitos momentos, Ana Rita e Thaís.

A Paulo, meu terapeuta junguiano, pelas horas de investimento, paciência e fé psicológica, além dos sentidos dados aos signos, símbolos e sonhos que me acompanham.

A Ivana, pelos diversos momentos de amparo e aconselhamento, por sempre acreditar e me mostrar outras formas flexíveis de enxergar os movimentos que a vida faz.

A Josimar, pelos momentos de diversões e risos, além da confiança no compartilhar segredos e projetos.

A Ju, pelo carinho e cuidado, além do incentivo constante, suas colaborações foram ímpares.

A Diego Batista, pelo carinho e pelo cuidado. Você faz parte dessa conquista, tenha certeza disso.

A Salma, pelos conselhos, pelas escutas atentas, pela admiração e pelo carinho sempre muito evidentes em cada encontro e pelas mensagens diárias, com humor, desabafos e fofocas.

A Messias, pela paciência em preservar laços, peculiar às características astrológicas do seu signo, pelos diversos movimentos de saídas e diversões, desabafos e amparos mútuos.

Aos meninos que, em tão pouco tempo de história, vivemos momentos incríveis e divertidos, Heber, Marco e Iago.

Pelos mares acadêmicos

A primeira professora que ministrou aula na graduação, 2009.1, em LET A13 – Introdução à Língua Portuguesa, a primeira orientadora, Profa. Dra. Marcela Paim, que, desde aquele momento, me despertou para os estudos sobre a diversidade linguística. Pró, obrigado pela paciência, gentileza e sabedoria compartilhadas.

A professora que, em LET C12 – Introdução aos estudos dialetais, na graduação, consolidou o meu interesse pela beleza que é a variação linguística, a atual orientadora, Profa. Dra. Silvana Ribeiro. Pró, obrigado pelos incontáveis bons momentos vividos – das orientações, das

viagens de campo, dos ensinamentos, pelo exemplo profissional. Vida longa ao seu fazer didático/pedagógico, que outros estudantes tenham a oportunidade de tê-la pelos caminhos.

Ao professor, coorientador desta tese, Prof. Dr. Rerisson Cavalcante, que com bastante competência e com contribuições muito importantes, além de questionamentos ímpares, tanto me ajudou na construção desse trabalho.

As professoras do Projeto ALiB, sobretudo, a Profa. Dra. Jacyra Mota, a Profa. Suzana Cardoso – *in memoriam* – e a Profa. Dra. Ana Regina – *in memoriam* – pois fizeram a diferença na minha trajetória, sendo exemplos de transparência, dedicação, respeito e competência.

Ao Comitê Nacional do Projeto ALiB, pela autorização em ceder dados inéditos.

Aos professores e as professoras do ILUFBA, tanto da graduação, quanto da pós-graduação, uma vez que, se comparados a um time, vocês são excelentes e importantes para a formação dos estudantes.

As professoras que, gentilmente, aceitaram o convite para compor a banca: Profa. Dra. Jacyra Mota, Profa. Dra. Marcela Paim, Profa. Dra. Elisângela Mendes, Profa. Dra. Norma Lúcia Almeida e Profa. Dra. Lílian Teixeira. E aos professores Prof. Dr. Clézio Gonçalves, Prof. Dr. Gredson dos Santos e Prof. Dr. Valter Romano.

Aos colegas da graduação, que foram muitos, em especial, Daiane, Jéssica, Elias, Heide, Jacqueline, Ana Virgínia, Naiara, William e Rodrigo.

Aos colegas bolsistas do Projeto ALiB, pois, durante anos, compartilhamos juntos esse sonho, nos tornarmos doutores, em especial, Carolina Antônia e Mara Haab.

Aos colegas da pós-graduação, pelos conhecimentos e trocas durante os componentes cursados nessa jornada.

Aos colegas da pós-graduação alibiana, Amanda, Rosana e Diócles, os foneticistas *experts* no uso do programa estatístico.

A Weldon, pelo cuidado e pela presteza em me atender e traduzir os pensamentos de um linguista na mais pura e bela, arte cartográfica.

A Luís, pelo cuidado e pela presteza, sempre.

A Mariana e a Nicolas, que abraçaram a ideia de utilizar os mecanismos digitais nos produtos geolinguísticos.

Aos servidores do ILUFBA, pois são importantes peças para que o Instituto continue servindo qualidade, em especial: Camila, Hugo, Thiago, Ricardo, Cristiane e Almiro. Além destes, os funcionários terceirizados, em especial, Nancy.

Aos colegas professores da UFBA, em especial, ao grupo de Estágio de Língua Portuguesa da FACED/UFBA, professoras Emília, Dinea, Lícia, Noemi, Raquel, Ana Cristina, Sandra e Fátima, agradeço pelos momentos importantes e pelos exemplos no cuidado com a formação dos estudantes.

Aos colegas professores da UNEB, em especial, Girleide, *campus* IV, Lícia, Mariane, Charles, Joana, Lílian e Marcelo, *campus* XVIII, por todo carinho e acolhimento, além da parceria.

Aos colegas professores do Grupo Montessoriano, em especial, Carina, Vércio, Alzira, Semar, Daniele, Alexandra, Augusto e Diego. Além das diretoras, Rosana e Lúcia.

Aos estudantes das faculdades particulares, UNIRB, Montessoriano e FIEP. Além daqueles das universidades públicas, UFBA e UNEB. Vocês me estimulam a crescer a cada encontro.

A CAPES, pois, com o auxílio financeiro, por meio da bolsa, pude investir tempo e forças na pesquisa.

Pra mim brincar

Não há nada mais gostoso do que o mim sujeito de verbo no infinitivo. Pra mim brincar. As cariocas que não sabem gramática falam assim. Todos os brasileiros deviam de querer falar como as cariocas que não sabem gramática.

As palavras mais feias da língua portuguesa são quiçá, alhures e miúde (BANDEIRA, 1986, p. 19).

— Você é sempre muito boa observadora em relação às atitudes das pessoas — elogia Irene. — E talvez esteja certa também nesse caso. Quando queremos saber de que maneira os falantes reagem a determinadas formas lingüísticas, aplicamos testes que servem para medir a aceitabilidade dessas formas. Não sei se já foi feito algum teste em relação ao “para mim fazer”, mas é provável que os falantes cultos não aceitem essa construção com tranqüilidade, embora muitos a usem diariamente (BAGNO, 2006, p. 209).

SANTOS, Leandro Almeida dos. **Pra eu e pra mim fazer:** traços sociodialetais nas capitais brasileiras. 2023. 285f. Tese (Doutorado em Língua e Cultura). Universidade Federal da Bahia, Salvador. 2023.

RESUMO

Nesta tese, analisa-se o comportamento dos pronomes pessoais do caso reto *eu* e do caso oblíquo tônico *mim*, em posição de sujeito, precedidos da preposição *para*, em contextos como *Isto é para eu fazer* e *Isto é para mim fazer*, nas capitais brasileiras. O objetivo é verificar a variação entre o *eu* (forma de prestígio) e o *mim* (forma estigmatizada) e as possíveis influências de fatores intra e/ou extralinguísticos condicionadores ou não do uso variável, tais como: fatores diatópicos, sociais, linguísticos e históricos, que podem moldar a língua e, por vezes, interferir nos usos. Como fundamentos teóricos e metodológicos, adotam-se os pressupostos da Dialectologia, com ênfase no método geolinguístico pluridimensional – e da Sociolinguística Variacionista. Desse modo, acredita-se que a estrutura *para/pr+eu/mim+infinitivo* possui um comportamento que pode evidenciar diferentes perfis, quer seja dialetal/ horizontal, quer seja sociolinguístico/vertical. Os *corpora* de análise para o estudo foram recolhidos em amostras orais do Português Brasileiro (PB), nas décadas de 70 e 90, dados do Projeto de Estudo da Norma Linguística Urbana Culta do Brasil – Projeto NURC – e, nos anos 2000, dados do Projeto Atlas Linguístico do Brasil – Projeto ALiB. Ao tomar por base dois importantes bancos de dados sobre a língua portuguesa falada em solo brasileiro, ressaltam-se os critérios metodológicos distintos, a saber: i. Projeto NURC: 307 inquiridos, a partir do Diálogos entre Informante e Documentador (DID); Diálogos entre dois informantes (D2); e das Elocuções formais (EF); 342 informantes, oriundos de cinco capitais – Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre –, pertencentes as faixas etárias I – 25 a 35 anos; II – 36 a 55 anos; e III – 56 anos em diante, todos com nível universitário; ii. Projeto ALiB: 200 inquiridos, a partir de todo o Questionários (COMITÊ NACIONAL..., 2001), sobretudo, a questão 023 do Questionário Morfosintático (QMS); 200 informantes, oriundos das 25 capitais, pertencentes as faixas etárias I – 18 a 30 anos – e II – 50 a 65 anos; com nível de escolaridade fundamental e universitário. Após ampla revisão de literatura, por diferentes perspectivas – gramáticas, gerativista, sociocognitiva, sociolinguística e dialetológica – sobre o fenômeno ora analisado, audição dos áudios e consulta às transcrições grafemáticas, as ocorrências foram codificadas e submetidas à análise do programa estatístico *Goldvarb X* (SANKOFF; TAGLIAMONTE; SMITH, 2005). Os dados submetidos à tal análise demonstraram que a distribuição sob o eixo horizontal, no que tange à localidade, é um fator relevante para o entendimento e à compreensão desse fenômeno no PB. Assim, no que tange ao eixo vertical, embora a forma *mim* – estigmatizada, associada ao “erro” e atribuída à língua falada indígena – seja documentada em todas as capitais, independentemente do nível de escolaridade, nota-se que os informantes universitários, nos dois *corpora*, tendem ao uso da forma de prestígio. Ademais, um fator linguístico – função sintática da oração infinitiva – também exerceu influência sobre as escolhas dos pronomes em estudo. Os resultados foram dispostos em cinco cartas linguísticas, com dados percentuais e pesos relativos, a fim de demonstrar a variação diatópica e a diastrática, por meio de critérios que facilitam o acesso e a leitura dos dados, como a representação do não dado; a inserção de QR-code nas cartas, que levará à análise delas em uma plataforma digital aberta, visando a um fazer científico ao alcance de todos, fato que torna a geolinguística brasileira a pioneira nesse quesito.

Palavras-chave: Dialectologia. Projeto NURC. Projeto ALiB. Pronomes. Sociolinguística.

SANTOS, Leandro Almeida dos. **Pra eu e pra mim fazer:** traços sociodialetais nas capitais brasileiras. 2023. 285f. Tese (Doutorado em Língua e Cultura). Universidade Federal da Bahia, Salvador. 2023.

ABSTRACT

This thesis analyzes the behavior of the personal pronouns *eu* and *mim*, in subject position, preceded by the preposition *para*, in contexts such as *Isto é para eu fazer* and *Isto é para mim fazer*, in Brazilian capital cities. The aim is to verify the variation between *eu* (a prestigious form) and *mim* (a stigmatized form) and the possible influences of intra- and/or extralinguistic factors that may or may not condition variable use, such as: diatopic, social, linguistic and historical factors, which can shape language and sometimes interfere with use. The theoretical and methodological foundations are Dialectology, with an emphasis on the multidimensional geolinguistic method - and Variationist Sociolinguistics. Thus, it is believed that the structure *para/pr+eu/mim+infinitivo* has a behavior that can show different profiles, whether dialectal/horizontal or sociolinguistic/vertical. The corpora of analysis for the study were collected from oral samples of Brazilian Portuguese (BP) in the 1970s and 1990s, data from the Project for the Study of Adult Urban Linguistic Norms of Brazil - NURC Project - and, in the 2000s, data from the Linguistic Atlas of Brazil Project - ALiB Project. Based on two important databases on the Portuguese language spoken on Brazilian soil, the different methodological criteria stand out: i. NURC Project: 307 surveys, based on the Dialogues between Informant and Documenter (DID); Dialogues between two informants (D2); and Formal Elocutions (EF); 342 informants, from five capitals - Recife, Salvador, Rio de Janeiro, São Paulo and Porto Alegre -, belonging to the age groups I - 25 to 35 years; II - 36 to 55 years; and III - 56 years onwards, all with a university degree; ii. ALiB Project: 200 surveys, based on the entire Questionnaire (COMITÊ NACIONAL..., 2001), especially question 023 of the Morphosyntactic Questionnaire (QMS); 200 informants, from the 25 capital cities, belonging to age groups I - 18 to 30 years - and II - 50 to 65 years; with primary and university education. After a wide-ranging literature review from different perspectives - grammatical, generative, sociocognitive, sociolinguistic and dialectological - on the phenomenon being analyzed, listening to the audios and consulting the graphemic transcriptions, the occurrences were coded and submitted to analysis using the statistical program *Goldvarb X* (SANKOFF; TAGLIAMONTE; SMITH, 2005). The data submitted to this analysis showed that the distribution on the horizontal axis, in terms of location, is a relevant factor for understanding and comprehending this phenomenon in BP. Thus, with regard to the vertical axis, although the *mim* form - stigmatized, associated with "error" and attributed to the indigenous spoken language - is documented in all capitals, regardless of level of education, it can be seen that university-educated informants in both corpora tend to use the prestige form. In addition, a linguistic factor - the syntactic function of the infinitive clause - also influence the choices of pronouns under study. The results were laid out in five linguistic charts, with percentage data and relative weights, in order to demonstrate diatopic and diastratic variation, using criteria that make it easier to access and read the data, such as the representation of the not given; the insertion of QR-code in the charts, which will lead to their analysis on an open digital platform, aiming to make scientific work available to everyone, a fact that makes Brazilian geolinguistics a pioneer in this regard.

Keywords: Dialectology. NURC Project. ALiB Project. Pronouns. Sociolinguistics.

LISTA DE FIGURAS

Figura 1 – Periodização dos estudos dialetais brasileiros	36
Figura 2 – Os momentos da Geolinguística no Brasil	38
Figura 3 – Panorama dos atlas estaduais/regionais brasileiros (2020)	42
Figura 4 – Número de atlas de pequeno domínio (concluídos) no Brasil	43
Figura 5 – Mapa dialetal de Antenor Nascentes (1922)	47
Figura 6 – Mapa dialetal de Antenor Nascentes (1933)	48
Figura 7 – Mapa dialetal de Antenor Nascentes (1953)	48
Figura 8 – Mapa dialetal, a partir do georreferenciamento da Cartografia Automatizada	50
Figura 9 – Divisão dialetal de Minas Gerais proposta por Zágari ([1998] 2005)	52
Figura 10 – Representação da forma “para eu/mim fazer” na mídia digital	70
Figura 11 – Carta 87 Menstruação do APFB	73
Figura 12 – Carta 094 Menstruação do ALPB	74
Figura 13 – Carta 92 Menstruação do ALS	74
Figura 14 – Carta Pique	77
Figura 15 – Carta Sema Pegar	78
Figura 16 – Carta Amarelinha	79
Figura 17 – Carta Estrela da Tarde	79
Figura 18 – Genealogia da língua portuguesa	88
Figura 19 – Alguns fluxos migratórios ocorridos no Brasil entre 1950 a 1970	90
Figura 20 – Alguns fluxos migratórios ocorridos no Brasil entre 1970 a 1990	91
Figura 21 – Representação do fenômeno em estudo em uma gramática pedagógica	117
Figura 22 – Rede de Pontos do Projeto ALiB	153
Figura 23 – Etapas do processo de Informatização do Projeto ALiB	161
Figura 24 – Proposta de periodização do Atlas Linguístico do Brasil – ALiB	167
Figura 25 – Capitais pertencentes à Rede de Pontos do Projeto ALiB	175
Figura 26 – Descrição da amostra e das rodadas estatísticas – dados do Projeto NURC	201
Figura 27 – Carta experimental M01 – distribuição dos pronomes eu e mim quanto à localidade – dados do Projeto NURC	209
Figura 28 – Descrição da amostra e das rodadas estatísticas – dados do Projeto ALiB	215
Figura 29 – Carta experimental M02 – Distribuição dos pronomes eu e mim quanto à localidade – dados do Projeto ALiB	219

Figura 30 – Carta experimental M03 – distribuição dos pronomes eu quanto à escolaridade – dados do Projeto ALiB	223
Figura 31 – Descrição das amostras e rodadas estatísticas – dados dos Projetos NURC e ALiB – agrupamento dos corpora: informantes universitários	234
Figura 32 – Carta experimental M04 – Distribuição dos pronomes eu e mim quanto à localidade – dados do Projeto NURC e do Projeto ALiB – informantes universitários	241
Figura 33 – Carta experimental M05 – Distribuição dos pronomes eu e mim quanto às Eloquções Formais – dados do Projeto NURC	247

LISTA DE QUADROS

Quadro 1 – Periodização dos estudos dialetais brasileiros	33
Quadro 2 – Propostas de divisão dialetal do Brasil	45
Quadro 3 – Estudos lexicais com dados do Projeto ALiB sobre áreas dialetais	53
Quadro 4 – Cinco problemas propostos pela Teoria da Mudança Linguística	66
Quadro 5 – Ciclos econômicos marcantes para a formação do povo brasileiro	92
Quadro 6 – Panorama da história sociolinguística do PB	94
Quadro 7 – Capitais brasileiras quanto ao século e a data de fundação	96
Quadro 8 – Capitais brasileiras quanto ao surgimento do ensino universitário	96
Quadro 9 – Paradigma dos pronomes do latim vulgar	101
Quadro 10 – Paradigma de flexão de casos dos nomes da 1ª declinação do latim	102
Quadro 11 – Distribuição dos pronomes pessoais do português arcaico	104
Quadro 12 – Gramáticas de língua portuguesa dos séculos XX e XXI	111
Quadro 13 – Os pronomes pessoais do PB propostos por Castilho (2016)	114
Quadro 14 – Os pronomes pessoais, conforme a visão da gramática tradicional	115
Quadro 15 – Os pronomes pessoais no PB, conforme Carvalho (2008)	121
Quadro 16 – Formas pronominais, conforme Rocha (2010)	123
Quadro 17 – Os pronomes pessoais do PB, segundo Mendes (2016)	128
Quadro 18 – Pronomes pessoais de primeira pessoa do singular, conforme Mendes (2016)	129
Quadro 19 – Flexão de caso na norma culta do PB, conforme Lucchesi (2008)	138
Quadro 20 – Flexão de caso no português afro-brasileiro	139
Quadro 21 – Os pronomes pessoais do português de Cabo Verde	139
Quadro 22 – Os pronomes pessoais do português em Santo Tomé	141
Quadro 23 – Publicações da coleção do Projeto NURC	146
Quadro 24 – Proposta de periodização do Projeto da Norma Urbana Linguística Culta – Projeto NURC	148
Quadro 25 – Subdivisão do Questionário ALiB (2001)	154
Quadro 26 – Síntese da história do Projeto Atlas Linguístico do Brasil – Projeto ALiB	162
Quadro 27 – Proposta de periodização do Atlas Linguístico do Brasil – ALiB	163
Quadro 28 – Inquéritos consultados do Projeto NURC	172

Quadro 29 – Estratificação dos informantes do Projeto ALiB – capitais	173
Quadro 30 – Capitais pertencentes à Rede de Pontos do Projeto ALiB	175
Quadro 31 – Resumo dos resultados quanto ao pronome eu – dados dos corpora analisados	251

LISTA DE GRÁFICOS

Gráfico 1 – Uso pronominal na estrutura para/pra + eu/mim + infinitivo – dados das capitais do Projeto NURC: frequência geral das respostas válida	199
Gráfico 2 – Uso pronominal na estrutura para/pra + eu/mim + infinitivo – dados das capitais do Projeto ALiB: frequência geral das respostas	213
Gráfico 3 – Uso pronominal na estrutura para/pra + eu/mim + infinitivo – agrupamento dos corpora: frequência geral das respostas	232

LISTA DE TABELAS

Tabela 1 – Uso pronominal na estrutura para/pra + eu/mim + infinitivo – dados das capitais do Projeto NURC: todas as respostas	199
Tabela 2 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme função sintática da oração infinitiva – dados do Projeto NURC	202
Tabela 3 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme papel semântico do pronome sujeito – dados do Projeto NURC	203
Tabela 4 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme natureza semântica do verbo da oração principal – dados do Projeto NURC	204
Tabela 5 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme conjugação verbal – dados do Projeto NURC	205
Tabela 6 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme estilo de registro – dados do Projeto NURC	206
Tabela 7 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme a região geográfica – dados do Projeto NURC	208
Tabela 8 – O pronome eu em posição de sujeito, nas capitais do Brasil – distribuição conforme a localidade – dados do Projeto NURC	208
Tabela 9 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme faixa etária – dados do Projeto NURC	210
Tabela 10 – Uso pronominal na estrutura para/pra + eu/mim + infinitivo – dados das capitais do Projeto ALiB: todas as respostas	212
Tabela 11 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme a região geográfica – dados do Projeto ALiB	216
Tabela 12 – O pronome eu em posição de sujeito, nas capitais do Brasil – distribuição conforme a localidade – dados do Projeto ALiB	217
Tabela 13 – O pronome eu em posição de sujeito, nas capitais do Brasil – distribuição conforme a escolaridade – dados do Projeto ALiB	220
Tabela 14 – O pronome eu em posição de sujeito, nas capitais do Brasil – distribuição conforme escolaridade – dados do Projeto ALiB	222
Tabela 15 – O pronome eu em posição de sujeito, nas capitais do Brasil – distribuição conforme a faixa etária – dados do Projeto ALiB	224

Tabela 16 – O pronome eu em posição de sujeito, nas capitais do Brasil – distribuição conforme o sexo – dados do Projeto ALiB	225
Tabela 17 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme forma da preposição – dados do Projeto ALiB	226
Tabela 18 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme estilo de registro – dados do Projeto ALiB	227
Tabela 19 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme natureza semântica do verbo da oração principal – dados do Projeto ALiB	228
Tabela 20 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme função sintática da oração infinitiva – dados do Projeto ALiB	230
Tabela 21 – O pronome mim em posição de sujeito nas capitais do Brasil – distribuição conforme conjugação verbal – dados do Projeto ALiB	231
Tabela 22 – Uso pronominal na estrutura para/pra + eu/mim + infinitivo – agrupamento dos corpora: todas as respostas	232
Tabela 23 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme função sintática da oração infinitiva – agrupamento dos corpora ..	235
Tabela 24 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme natureza semântica do verbo da oração principal – agrupamento dos corpora	236
Tabela 25 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme papel semântico do pronome sujeito – agrupamento dos corpora ..	237
Tabela 26 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme conjugação verbal – agrupamento dos corpora	237
Tabela 27 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme a década – agrupamento dos corpora	239
Tabela 28 – O pronome eu em posição de sujeito nas capitais do Brasil – distribuição conforme a região geográfica – agrupamento dos corpora	240
Tabela 29 – O pronome eu em posição de sujeito, nas capitais do Brasil – distribuição conforme a localidade – agrupamento dos corpora	240
Tabela 30 – Uso pronominal na estrutura para/pra + eu/mim + infinitivo – dados das capitais do Projeto ALiB: respostas não obtidas	248

LISTA DE ABREVIATURAS

- ADDU – Atlas Diatopico y Diastratico del Uruguay
ADV – Advérbios inseridos na estrutura em análise
AIS – Atlas Italo-Suico
ALF – Atlas Linguistique de la France
ALiB – Atlas Linguístico do Brasil
A
LPB – Atlas Linguístico da Paraíba
ALS – *Atlas Linguístico de Sergipe*
APFB – *Atlas Prévio dos Falares Baianos*
BCIM – Base Cartográfica Contínua do Brasil, ao milionésimo
BNCC – *Base Nacional Comum Curricular*
CIDS – Congresso Internacional de Dialectologia e Sociolinguística
D2 – Diálogos entre dois informantes
DCI – Dativo com Infinitivo
DDB – Projeto Dicionário Dialetal Brasileiro
DID – Diálogos entre Informante e Documentador
DPs – *Determinant Phrases*
EALMG – *Esboço de um atlas linguístico de Minas Gerais*
ECM – *Exceptional Case Marking*
EF – Elocuções Formais
IBGE – Instituto Brasileiro de Geografia e Estatística
IC – Iniciação Científica
INDL – Inventário Nacional da Diversidade Linguística
INF – Informante
INQ – Inquiridor
LE – Texto para Leitura
NL – Não Lembra
NO – Não Obtida
NORMs – Nonmobile, older, rural, males
NS – Não Sabe
PA – Português africano
PAF – Português afro-brasileiro

PB – Português brasileiro
PCN – *Parâmetros Curriculares Nacionais*
PE – Português europeu
PEUL – Programa de Estudos sobre o uso da Língua
PILEI – Programa Interamericano de Linguística e Ensino de Idiomas
PM – Perguntas Metalinguísticas
POA – Porto Alegre
PPF – Português popular da cidade do interior de médio porte
PPGComp – Programa de Pós-Graduação em Computação
PPGLinC – Programa de Pós-Graduação em Língua e Cultura
PPI – Português popular de cidades do interior de pequeno porte
PPS – Português popular urbano
Projeto ALiB – Projeto Atlas Linguístico do Brasil
Projeto NURC – Projeto de Estudo da Norma Linguística Urbana Culta no Brasil
QFF – Questionário Fonético-Fonológico
QMS – Questionário Morfossintático
QP – Questões de Pragmática
QR Code – *Quick Response Code*
QSL – Questionário Semântico-Lexical
REC – Recife
RJ – Rio de Janeiro
SE – Sem a estrutura em análise
SGVCLin – Software para Geração e Visualização de Cartas Linguísticas
SIG – Sistemas de Informações Geográficas
SP – São Paulo
SSA – Salvador
TDIC – Tecnologias Digitais da Informatização e Comunicação
TDS – Temas para Discurso Semidirigido
TLI – Transmissão Linguística Irregular
UFBA – Universidade Federal da Bahia
UFMS – Universidade Federal do Mato Grosso do Sul

SUMÁRIO

1	ITINERÁRIOS INTRODUTÓRIOS	25
2	CAMINHOS TEÓRICOS	30
2.1	TRILHA DA VARIAÇÃO ESPACIAL: A DIALETOLOGIA	30
2.1.1	Propostas de periodizações dos estudos dialetais brasileiros	33
2.1.2	A pluridimensionalidade está na Dialetologia ou na Geolinguística?	39
2.1.3	Nos passos de Nascentes (1953): os estudos sobre as regiões dialetais do Brasil	45
2.2	TRILHA DA VARIAÇÃO SOCIAL: A SOCIOLINGUÍSTICA	60
2.2.1	Variação e mudança linguísticas: conceitos fundamentais	62
2.3	O ENTROCAMENTO DAS TRILHAS DA VARIAÇÃO	68
2.4	ROTEIROS POSSÍVEIS PARA OS CASOS DE “NÃO DADO” E POUCOS DADOS	72
2.5	SÍNTESE DA SEÇÃO	84
3	AS VIAS DOS PRONOMES	87
3.1	NAS ROTAS DA HISTÓRIA DA LÍNGUA PORTUGUESA	87
3.1.1	Trajetos sobre as hipóteses da formação do Português Brasileiro	97
3.2	OS ATALHOS DOS PRONOMES	100
3.2.1	Noções de pessoa e de caso: a natureza dos pronomes	102
3.3	SÍNTESE DA SEÇÃO	108
4	SENDAS SOBRE A ESTRUTURA “PARA/PRA + EU/MIM/ + VERBO INFINITIVO”	109
4.1	ESTUDOS SOBRE O PORTUGUÊS BRASILEIRO	109
4.1.1	Perspectiva das Gramáticas	110
4.1.2	Perspectiva Gerativa	118
4.1.3	Perspectiva Sociocognitiva	123
4.1.4	Perspectiva Sociolinguística	125
4.1.5	Perspectiva Dialetológica	134
4.2	ESTUDOS SOBRE O PORTUGUÊS BRASILEIRO, PORTUGUÊS EUROPEU E O PORTUGUÊS AFRICANO	137
4.3	SÍNTESE DA SEÇÃO	143
5	TRAJETÓRIAS DO PROJETO NURC E DO PROJETO ALiB	144
5.1	O PROJETO DE ESTUDO DA NORMA LINGUÍSTICA URBANA CULTA NO BRASIL – PROJETO NURC	144
5.1.1	Esboço de uma periodização do Projeto de Estudo da Norma Linguística Urbana Culta – Projeto NURC	147

5.2	O PROJETO ATLAS LINGUÍSTICO DO BRASIL – PROJETO ALiB	152
5.2.1	Esboço de uma periodização do Atlas Linguístico do Brasil – ALiB	162
5.3	SÍNTESE DA SEÇÃO	169
6	ROTAS METODOLÓGICAS	171
6.1	O LEVANTAMENTO DOS DADOS E OS CRITÉRIOS PARA O TRATAMENTO ESTATÍSTICO E INTERPRETATIVO	176
6.1.1	A variável dependente	178
6.1.2	As variáveis independentes	178
6.1.2.1	<i>Variáveis linguísticas</i>	178
6.1.2.1.1	Estilo de registro	179
6.1.2.1.2	Natureza semântica do verbo da oração principal	181
6.1.2.1.3	Papel semântico do pronome sujeito	183
6.1.2.1.4	Função sintática da oração infinitiva	184
6.1.2.1.5	Conjugação verbal	185
6.1.2.1.6	Forma da preposição	187
6.1.3	Variáveis extralinguísticas	188
6.1.3.1	<i>Variáveis extralinguísticas – Projeto NURC</i>	188
6.1.3.2	<i>Variáveis extralinguísticas – Projeto ALiB</i>	188
6.1.4	Variáveis linguísticas e extralinguísticas – agrupamento dos corpora	188
6.1.5	A codificação dos dados	189
6.1.6	O Goldvarb X	189
6.1.7	A cartografia	191
6.1.7.1	<i>Afinal, o que é e para que serve uma carta linguística?</i>	191
6.1.7.2	<i>As cartas linguísticas</i>	193
6.1.8	Metalinguística nos/dos dados	195
6.2	SÍNTESE DA SEÇÃO	197
7	RUMOS DA ANÁLISE DE DADOS	198
7.1	OBSERVAÇÕES INICIAIS – PROJETO NURC	198
7.2	OS CONDICIONADORES PARA A REALIZAÇÃO DO FENÔMENO – PROJETO NURC	200
7.2.1	Alguns comentários sobre as rodadas no programa estatístico	200
7.2.2	As variáveis independentes – intralinguísticas	201
7.2.2.1	<i>Função sintática da oração infinitiva</i>	201
7.2.2.2	<i>Papel semântico do pronome sujeito</i>	202
7.2.2.3	<i>Natureza semântica do verbo da oração principal</i>	203

7.2.2.4	<i>Conjugação Verbal</i>	205
7.2.2.5	<i>Estilo de registro</i>	205
7.2.2.6	<i>Forma da preposição</i>	207
7.2.3	As variáveis independentes – extralinguísticas	207
7.2.3.1	<i>A distribuição diatópica – região geográfica</i>	207
7.2.3.2	<i>A distribuição diatópica – localidade</i>	208
7.2.3.3	<i>A distribuição por faixa etária</i>	210
7.2.3.4	<i>A distribuição por década</i>	211
7.3	OBSERVAÇÕES INICIAIS – PROJETO ALiB	212
7.4	OS CONDICIONADORES PARA A REALIZAÇÃO DO FENÔMENO – PROJETO ALiB	214
7.4.1	Alguns comentários sobre as rodadas no programa estatístico	214
7.4.2	As variáveis independentes – extralinguísticas	215
7.4.2.1	<i>A distribuição diatópica – região geográfica</i>	216
7.4.2.2	<i>A distribuição diatópica – localidade</i>	217
7.4.2.3	<i>A distribuição por escolaridade</i>	220
7.4.2.4	<i>A distribuição por faixa etária</i>	224
7.4.2.5	<i>A distribuição por sexo</i>	225
7.4.3	As variáveis independentes – intralinguísticas	226
7.4.3.1	<i>Forma de preposição</i>	226
7.4.3.2	<i>Estilo de registro</i>	227
7.4.3.3	<i>Natureza semântica do verbo da oração principal</i>	228
7.4.3.4	<i>Papel semântico do pronome sujeito</i>	229
7.4.3.5	<i>Função sintática da oração infinitiva</i>	230
7.4.3.6	<i>Conjugação Verbal</i>	231
7.5	OBSERVAÇÕES INICIAIS – AGRUPAMENTO DOS <i>CORPORA</i>	232
7.5.1	Alguns comentários sobre as rodadas no programa estatístico	233
7.5.2	As variáveis independentes – intralinguísticas	234
7.5.2.1	<i>Função sintática da oração infinitiva</i>	234
7.5.2.2	<i>Natureza semântica do verbo da oração principal</i>	235
7.5.2.3	<i>Papel semântico do pronome sujeito</i>	236
7.5.2.4	<i>Conjugação Verbal</i>	237
7.5.2.5	<i>Forma de preposição</i>	238
7.5.3	As variáveis dependentes – extralinguísticas	238
7.5.3.1	<i>A distribuição por década</i>	238

7.5.3.2	<i>A distribuição diatópica – região geográfica</i>	239
7.5.3.3	<i>A distribuição diatópica – localidade</i>	240
7.6	METALINGUAGEM NOS/DOS DADOS	242
7.6.1	As percepções metalinguísticas	242
7.6.2	As respostas Não Obtidas	246
7.7	SÍNTESE DA SEÇÃO	250
8	ENCRUZILHADAS FINAIS	252
	REFERÊNCIAS	263
	APÊNDICES	281
	Apêndice A	281
	Apêndice B	282
	Apêndice C	283
	Apêndice D	284
	Apêndice E	285

1 ITINERÁRIOS INTRODUTÓRIOS

A todas essas perguntas só se pode responder depois de realizar inquéritos in loco. Qualquer afirmação que não parte deles é frágil e tem pés de barro...
(SILVA NETO, 1957 [1955], p. 43).

Culturalmente, na sociedade brasileira, a utilização do pronome *mim*, enquanto sujeito em frases infinitivas, é alvo de vários estigmas e preconceitos, muitos desses fomentados no interior das instituições formais de ensino. Além disso, essa utilização é, de forma pejorativa, atribuída à fala dos indígenas. Logo, em algumas comunidades linguísticas, a variação entre os pronomes *eu*, pronome pessoal do caso reto – (01) *Essa tarefa é para eu fazer* –, e o *mim*, pronome pessoal do caso oblíquo tônico – (02) *Essa tarefa é para mim fazer* –, em posição de sujeito, em estrutura tal como *para/prá + eu/mim + infinitivo*, não é bem acolhida por parte dos falantes, devido às diversas avaliações negativas atribuídas à forma oblíqua.

Nesta tese, essa variação entre os dois pronomes é o objeto de investigação, tomando como referência dados diferentes do português falado nas capitais do Brasil, a partir dos *corpora* do Projeto de Estudo da Norma Linguística Urbana Culta do Brasil – Projeto NURC – décadas de 70 e 90, e o *corpus* do Projeto Atlas Linguístico do Brasil – Projeto ALiB – anos 2000.

Ao tomar como referência os dados de dois grandes marcos para o estudo da língua falada no Brasil, Projeto NURC e Projeto ALiB, essa pesquisa vincula-se aos pressupostos da Dialetologia e da Sociolinguística Variacionista, duas áreas da macrolinguística que concebem a língua como sistema heterogêneo e multifacetado, cujo meio de realização elementar é a fala de seus usuários, conforme as suas prerrogativas socioculturais, suscetível à variação e à mudança linguísticas, sob diversas motivações e condicionantes.

Nessa perspectiva, compreende-se que a configuração atual do Português Brasileiro (PB) é fruto dos diversos fatos históricos desencadeados quando da difusão da língua portuguesa pelo globo e das idiossincrasias linguísticas e culturais dos seus diversos grupos de falantes, haja vista que de língua restrita ao espaço da Península Ibérica, até a Idade Moderna, passou a ocupar domínios variados. Assim, estabelecendo-se as mais diversas situações de contato entre o português e tantas outras línguas, permitindo que se considere o contexto de difusão, em terras brasileiras.

É possível, nesse sentido, afirmar que o PB reflete, certamente, as heranças desse cenário multilíngue, multiétnico e pluricultural, não encontrando pares no Português Europeu

(PE) de ontem e de hoje, nem, muitas vezes, em outras variedades da língua portuguesa, como no Português falado em África (PA).

Ademais, é de fundamental importância olhar para o passado, a fim de compreender o presente, no que tange à configuração das línguas. Assim sendo, alguns estudiosos pontuam que as mudanças na flexão de caso que ocorreram na passagem das duas variedades do latim, o clássico e o vulgar, afetaram as línguas românicas.

Dada essa situação, busca-se delinear o perfil dialetal e o perfil sociolinguístico de uso dos pronomes pessoais do caso reto *eu* e do caso oblíquo tônico *mim* no PB, para os contextos de posição de sujeito, precedidos pela preposição *para*, tais como *para/prá eu fazer*” e *“para/prá mim fazer*.

A fim de atingir tal meta, no que tange aos dados das capitais brasileiras, pertencentes aos *corpora* analisados, traçam-se os seguintes objetivos específicos:

- a. verificar, por meio da história das línguas, os resquícios da herança latina no que tange à utilização dos pronomes *eu* e *mim*;
- b. apresentar um panorama expositivo, a partir de uma ampla revisão de literatura sobre o fenômeno no PB, no PE e no PA;
- c. traçar um quadro comparativo das formas pronominais *eu* e *mim* na variedade brasileira do português, no contexto ora analisado, nos *corpora* elegidos, por meio das análises estatísticas do *Goldvarb X* (SANKOFF; TAGLIAMONTE; SMITH, 2005);
- d. identificar a influência de fatores linguísticos (estruturais e discursivos/estilísticos) na expressão do fenômeno, bem como evidenciar suas possíveis correlações com fatores extralinguísticos: espaciais (localidade e região geográfica); sociais (sexo, escolaridade e faixa etária), além dos fatos da história social e linguística das capitais;
- e. por meio das cartas linguísticas, demonstrar a variação areal do fenômeno, com vistas a observação de possíveis traçados de isoglossas morfossintáticas.

A partir da ciência de que a realidade sociocultural brasileira, de fato, é heterogênea, multiétnica e plurilíngue, bem como o conhecimento de que as vivências das localidades são historicamente diferentes por vários motivos, além das considerações de outros estudiosos que se debruçaram sobre o tema, em especial, no que se refere aos pronomes pessoais *eu* e *mim*, a hipótese principal é que o fenômeno em estudo apresenta um comportamento diferenciado, quando observado sob os prismas diatópico – horizontal – e diastrático – vertical. Nesse último, destaca-se o fator escolaridade.

Acredita-se que a mola propulsora deste trabalho é, justamente, atender à necessidade de se explorar a realidade do PB, com relação às formas pronominais *eu* e *mim*, quando

utilizadas em orações infinitivas, em posição de sujeito, de modo a conhecer a sua distribuição espacial, sendo possível separar as capitais brasileiras em grupos, com vistas a possibilitar o traçado de isoglossas e, quiçá, traçados que possam descortinar a realidade dialetal do Brasil, colaborando com a agenda de estudos dos dialetólogos brasileiros.

Outro fator que merece atenção, quando da eleição do fenômeno em estudo, é que a importância da investigação sobre as formas pronominais *eu* e *mim* é reconhecida pela própria escolha de se criar uma questão específica para ele nos Questionários do ALiB, Questionário Morfosintático (QMS)¹ (COMITÊ NACIONAL..., 2001, p.39-43), especialmente dentre um conjunto de, apenas, 49 questões que compõem o referido QMS, a QMS 23. Vale destacar que os dados utilizados, no que tange ao banco de dados do Projeto ALiB, são inéditos e, portanto, necessitam de prévia autorização do Comitê Nacional.

Torna-se necessário afirmar que a observação e a análise de variantes linguísticas no PB, língua materna de milhões de falantes, justificam-se pela necessidade de mais estudos, na perspectiva sociodialetoal, com vieses ainda a serem explorados, incluindo a observação do fenômeno em outros *corpora* e a comparação com outras variedades do português em que se observa o mesmo fenômeno.

Diante do que foi mencionado, na segunda seção desta tese, apresentam-se reflexões sobre os fundamentos teóricos de duas subáreas da Linguística que se debruçam sobre a variação e a mudança linguística, tendo como foco principal a língua falada, Dialetologia e Sociolinguística. Apresentam-se, ainda, nessa seção, uma discussão sobre o dissenso entre o qualificador pluridimensional, ora utilizado para a Dialetologia, ora para um dos seus métodos, a Geolinguística, e também, à luz de exemplos históricos, apresenta-se uma proposição sobre a utilização do não dado² para as pesquisas sociodialetais, sobretudo para a necessidade de aprimoramento da cartografia temática de dados linguísticos.

Na seção seguinte, a terceira, apresenta-se uma abordagem de caráter histórico voltada para os pronomes, desde as variedades do latim até o PB. Nesta, também, há uma discussão sobre a sociohistória de formação do PB, atentando-se para as hipóteses interpretativas dessa variedade. Também, destacam-se, nesse aspecto, os séculos e datas de fundação das capitais, bem como o século de surgimento do ensino universitário nelas. Além disso, é debatida a real natureza dos pronomes, por meio de diferentes noções de caso e pessoa, com base em alguns autores e em algumas gramáticas.

¹ A busca pelo fenômeno foi feita em todas as partes que compõem os Questionários ALiB (COMITÊ NACIONAL..., 2001).

² Nomeia-se como não dado a ausência de respostas catalogadas nos inquéritos.

A quarta seção traz à tona alguns estudos, no que concerne ao PB, que se debruçaram sobre os pronomes *eu* e *mim* em posição de sujeito em orações infinitivas precedidos pela preposição *para*. Com isso, através dessa revisão ampla de literatura, discute-se o dativo com o infinitivo, com base na visão de diferentes posicionamentos teóricos, tais como: o das gramáticas, o do Gerativismo, o do Sociocognitivismo, o da Sociolinguística e o da Dialetoлогия. Para além dessa, analisa-se, ainda, o fenômeno no PE e no PA.

Na quinta seção, optou-se pela apresentação e caracterização dos *corpora*, ou seja, os aspectos sobre o surgimento, desenvolvimento, frutos e estágios atuais tanto do Projeto NURC quanto do Projeto ALiB. Destaca-se, nessa seção, inéditos esboços de periodizações para os dois referidos Projetos, marcos relevantes para os estudos sobre a língua falada no Brasil.

Encontra-se, na sexta seção, a metodologia adotada para o tratamento dos dados orais na tese. Apresentam-se os critérios metodológicos estabelecidos para a tese, ao detalhar os dados dos referidos Projetos. Isso revelou aspectos ímpares sobre os *corpora*, o que, de fato, leva a um tratamento interpretativo e um tratamento estatístico diferenciados, com o uso do pacote de programas do *Goldvarb X*, considerando os propósitos específicos deles. Além disso, há esclarecimentos para o entendimento sobre a elaboração das cartas linguísticas da tese, essas possuem um *Quick Response Code* – QR Code – que, ao ser scaneado, leva à leitura e interpretação dos dados em uma plataforma digital aberta, permitindo um acesso mais facilitado aos dados linguísticos, com um recurso digital moderno. Desse modo, constituindo a geolinguística brasileira o pionerismo nessa empreitada.

Na seção seguinte, a sétima, expõem-se os resultados obtidos no estudo, a partir de análises quantitativas e qualitativas, quando couberem. Essas análises são consideradas e fundamentadas por meio da seleção feita pelo *Goldvarb X*. Ademais, a fim de verificar a distribuição espacial das formas pronominais em estudo, apresentam-se cartas linguísticas, demonstrando os percentuais e os pesos relativos obtidos em cada capital e nas cinco capitais coincidentes nos *corpora* – Porto Alegre, São Paulo, Rio de Janeiro, Salvador e Recife. Para além disso, são explorados os aspectos metalinguísticos dos dados, sejam através de recortes de falas dos informantes, sejam pela observação dos dados não obtidos, estes últimos contemplados na representação cartográfica, consoante Santos (2016/2020).

A última seção desse trabalho destina-se à apresentação geral dos resultados alcançados, após as análises, reflexões e direcionamentos para perspectivas futuras para o estudo e exploração do fenômeno.

Seguem-se, após a última seção, as referências utilizadas ao longo do trabalho, os apêndices.

Os dados aqui trazidos e analisados, no que tange ao *corpus* dos anos 2000, foram tratados de forma inédita pelo autor da tese com base em autorização de uso concedida pelo Comitê Nacional do Projeto ALiB.

Cabe salientar que, durante a pesquisa de doutoramento, a sociedade mundial enfrentou a pandemia do coronavírus, período em que todas as atividades humanas foram afetadas. No que tange ao universo acadêmico, as aulas, devido aos decretos de distanciamento social, passaram para o modo virtual. A pandemia, de fato, afetou o andamento da pesquisa ora desenvolvida.

2 CAMINHOS TEÓRICOS

Agora, se é preciso o maior cuidado no sentido de evitar-se a debilitação, a queda e a morte da língua, é preciso, de outro lado, não confundir polícia com ditadura, com o domínio dos ranhetas que se julgam donos dela e, espumando, arrancam os cabelos a qualquer necessário neologismo exótico, ou correm atrás da gente rogando pragas, açoitando regrinhas e espionando se o pronome não saiu deslocado ou se a preposição se mantém no recato de uma regência quinhentista.

(AMARAL, 1920, p. 21).

Apresentam-se, nesta seção, os pressupostos teóricos que são fundamentais e norteadores para os entendimentos e as análises sobre o fenômeno ora estudado e objeto da tese. Nesse sentido, a Dialetoлогия e a Sociolinguística Variacionista servirão como aportes teóricos e metodológicos, bem como as interfaces entre esses ramos dos estudos linguísticos.

2.1 TRILHA DA VARIAÇÃO ESPACIAL: A DIALETOLOGIA

De antemão, compreende-se que a gênese da Dialetoлогия remonta ao final do século XVIII e o início das décadas do século XIX, portanto, bem anterior à concepção de língua e novas visões estabelecidas pelas diversas vertentes da Linguística Moderna, enquanto ciência. A Dialetoлогия, desse modo, é o ramo dessa ciência que procura “identificar, descrever e situar os diferentes usos em que uma língua se diversifica, conforme sua distribuição espacial, sociocultural e cronológica” (CARDOSO, 2010, p. 15).

A parte da ciência responsável pela variação espacial, no bojo do seu espectro de análise e investigação, tem como prioridade o registro e a catalogação das diferenças e das semelhanças dos fenômenos, referentes à língua falada, circunscritas aos espaços geográficos, levando em consideração, também, as interferências dos fatores sociais sobre tais realizações. Desse modo, ao unir diatopia – aspecto prioritário/geográfico – ao social – aspecto agregador/sociocultural – tal ramo amplia as possibilidades para, cada vez mais, entender e elucidar como se dá a configuração dos usos linguísticos.

A preocupação com as diferentes formas de falar pode ser notada desde os séculos passados. Sendo assim, há diversos inícios que podem ser considerados como importantes para o desenvolvimento da Dialetoлогия. No século XIX, os estudos dialetológicos, bem como o método específico, a Geolinguística, ganham significação por meio de vários fatos, dentre eles, destacam-se: em 1804, a criação da *Academie Celtique* – foi a primeira manifestação com recolha sistemática de dados a se debruçar sobre a apuração das diferenças

dialetais de uma língua – e, em 1812, J. Grim, autor pioneiro da descrição de dialetos alemães, posiciona-se em defesa do *patois* – com objetivos gerais e enorme abrangência, que, também, incluiu dados sobre o PB, fato que data o início de estudos com o teor dialetológico brasileiro (CARDOSO, 2010).

Nesse sentido, também, devem ser realçadas as contribuições realizadas pelos responsáveis das delimitações que, de certo modo, acabaram por fixar a Dialetologia como uma disciplina separada dos estudos comparativistas: Gustavo Wenker, na Alemanha, e Jules Gilliéron, na França.

Em 1876, na Alemanha, por meio da recolha de dados linguísticos por correspondência, Wenker inicia a documentação dos falares populares, ao enviar para professores um questionário linguístico com frases em alemão literário, além de instruções para que os informantes fornecessem relatos acerca da localidade a que se referia. *A priori*, os propósitos investigativos recobriam uma pequena área, a Província do Reno, mas, com o apoio governamental, as investigações recobriram, *a posteriori*, todo o território alemão.

Jules Gilliéron, a partir de 1887, preocupado em salvar as formas de falar locais, pode ser considerado como o dialetólogo responsável por consolidar e solidificar os passos iniciais dados por Wenker, uma vez que a Geolinguística ganhou grande vulto, por meio da execução do *Atlas Linguistique de la France* – ALF – (GILLIÉRON; EDMONT, 1902-1910). A recolha dos dados para o ALF foi feita por um comerciante francês, Edmond Edmont, que, para isso, aplicou um questionário com 1.920 perguntas a 700 informantes, dos quais apenas 60 eram mulheres, e percorreu 639 localidades. As respostas eram transcritas de forma imediata pelo inquiridor. Instaura-se, assim, a metodologia dialetal de obtenção de dados *in loco*. A partir de então, houve uma consolidação da Dialetologia, enquanto responsável por investigar a língua, com o foco na variação espacial, ou seja, diatópica. Para tal, fazendo o uso do seu método primordial, a Geolinguística.

A Geolinguística ou Geografia Linguística é “[...] um método científico de recolha sistemática de dialetismo” (CASTILHO, 1972-1973, p. 7) que objetiva cartografar os dados linguísticos, de modo que se possam evidenciar as semelhanças e diferenças de um dialeto ou língua no espaço geográfico, sejam por representação cartográfica nacional, regional ou continental. Ainda, nesse sentido, a Geografia Linguística pode ser entendida como “[...] uma metodologia ou, mais precisamente, um conjunto de métodos para recolher de um modo

sistemático os testemunhos das diferenças dialetais”³ (CHAMBERS; TRUDGILL, 1994, p. 37) [tradução nossa].

De maneira mais específica, a Geolinguística destina-se a elaborar atlas linguísticos de diversas naturezas, a depender dos intentos do pesquisador. No tocante à cartografia, há diversas formas metodológicas que evidenciam os recursos técnicos da época. Nesse sentido, à medida que a Dialetoлогия foi se desenvolvendo, o seu método também foi ganhando outras faces e aprimorando as suas técnicas para uma melhor disposição dos produtos oriundos desse método cartográfico.

A metodologia proposta pela Geografia Linguística preconiza o estabelecimento de critérios para seleção das localidades e dos informantes; a recolha dos dados, por meio de entrevistas ou inquéritos, com base em questionários; a elaboração dos resultados em cartas e/ou mapas linguísticos, que, em muitos casos, vão compor os atlas, os estudos e/ou a interpretação dos dados dispostos pelos mapas linguísticos, objetivando o traçado de isoglossas demarcadoras de um dado fenômeno e suas fronteiras no espaço e, por fim, a publicação desses resultados.

No Brasil, observa-se que os primeiros comentários que se referem à língua portuguesa em solo brasileiro, em 1725, foram realizados por meio de D. Jerônimo, século XVIII, e tocavam em aspectos sobre as divergências entre os usos do PB e do PE. O contador de Argote, D. Jerônimo, afirmava serem os aspectos lexicais do PB caracterizados por formas oriundas das línguas autóctones e do português antigo, conforme assinala Cardoso (1999).

Semeavam-se, assim, em paralelo, com o Visconde de Pedra Branca, Domingos Borges de Barros, as primeiras considerações acerca da descrição do PB: “Trata-se de uma lista de palavras que apresenta um rol de oito nomes que mudam de significação e outro de cinquenta nomes usados exclusivamente no Brasil” (CARDOSO, 1999, p. 234), ao contrastar, no século XIX, aspectos lexicais e fônicos entre o PE e o PB, em 1826, na introdução da obra *Atlas Ethnographique du globe...*⁴, de Adrien Balbi. A partir desse marco, outros, aqui no Brasil, também importantes, são observados, no que concerne ao início, desenvolvimento e estágio atual desse ramo de estudos que é responsável pela variação espacial.

³ [...] una metodología – más exactamente – un conjunto de métodos para recopilar de un modo sistemático los testimonios de las diferencias dialectales.

⁴ O *Atlas Ethnographique du Globe, ou classification des peuples anciens et modernes d’après leurs langues... et suivi Du tableau phisique, moral et politique des cinq parties du monde...*

No intuito de compreender o percurso histórico acerca dos estudos dialetais, a partir desses fatos, vale mencionar que os tais foram deveras importantes para o desenvolvimento, solidificação e aprimoramento das pesquisas em Dialetologia em solo brasileiro. Faz-se, então, necessário verificar, ainda pelos vieses histórico e cronológico, algumas propostas de periodizações dos estudos dialetais brasileiros, conforme item 2.1.1. É notável que, quando se observam as periodizações propostas, os estudos sobre a morfologia, a sintaxe e a morfossintaxe são bem ausentes, de modo geral. E, quando presentes, são em períodos e/ou fases pontuais.

2.1.1 Propostas de periodizações dos estudos dialetais brasileiros

De modo a continuar pelos vieses histórico e cronológico sobre os estudos dialetais brasileiros, no Quadro 1, expõem-se as cinco fases, seus marcos temporais de início e término e a autoria das propostas quanto à periodização.

Quadro 1 – Periodização dos estudos dialetais brasileiros

FASES	MARCOS		AUTORIA(S)
	Início	Término	
1 ^a	1826 Considerações do Visconde de Pedra Branca sobre o PE e o PB.	1920 <i>O dialeto Caipira</i> , de Amadeu Amaral.	Nascentes (1952/1953).
2 ^a	1920 <i>O dialeto Caipira</i> , de Amadeu Amaral.	1952 Decreto presidencial que institui a criação de atlas nacional.	início: Nascentes (1952/1953) final: Ferreira; Cardoso (1994).
3 ^a	1952 Decreto presidencial que institui a criação de atlas nacional.	1996 Instituição do Projeto ALiB.	início: Ferreira; Cardoso (1994). final: Mota; Cardoso (2006).
4 ^a	1996 Instituição do Projeto ALiB.	2014 Publicação dos dois primeiros volumes do ALiB.	início: Mota; Cardoso (2006). final: Teles (2018).
5 ^a	2014 - Fase vigente Publicação dos dois primeiros volumes do ALiB.		Teles (2018).

Fonte: Elaborado pelo autor (2023).

A partir do exposto, citam-se, detalhadamente, as propostas de periodização de Nascentes (1952, 1953), o autor reconhece duas fases, a primeira, que se inicia em 1826, com a publicação do texto do Domingos Borges de Barros, e se estende até 1920, com a publicação de *O dialeto caipira*, de Amadeu Amaral. A principal característica dessa fase inicial é que as obras se voltavam para o estudo do léxico, devido às publicações de obras lexicográficas, tais como: glossários; vocabulários; dicionários e léxicos regionais.

Nessa fase, a atenção dos dialetólogos é voltada para a dimensão lexical, como se nota, por meio das obras e produções que retratam as intencionalidades dos pesquisadores. A dimensão morfossintática não aparece nesse primeiro momento.

No que concerne à segunda fase, esta se estende de 1920 até a de publicação do decreto presidencial que previa a elaboração de um atlas nacional, em 1952. Essa fase é caracterizada pelo foco nos aspectos lexicais e pelo início da descrição dos aspectos fonético-fonológicos e morfossintáticos. Além disso, a 2ª fase se caracteriza por produções diversas, tais como: em 1922, *O linguajar carioca em 1922*, de Antenor Nascentes, que, após a 2ª edição, passou a ser chamada de *O linguajar carioca*. Nessa obra, encontra-se uma das propostas de divisão dialetal do Brasil, apresentada pelo autor. Registra-se também a publicação de obras, tal como *A língua do Nordeste*, em 1934, de Mário Marroquim.

Nessa fase, a segunda, os aspectos morfossintáticos ganham mais importância, adicionados aos aspectos lexicais e fonéticos que, desde o início, são alvos mais frequentes dos interesses dos pesquisadores e estudiosos da área.

Ferreira e Cardoso (1994), baseadas em fatos históricos da Dialetologia, reanalisam a proposta de Nascentes e a reestruturam, propondo a 3ª fase. Para tal empreendimento, as autoras consideraram os pontos estabelecidos por Nascentes e propuseram as seguintes fases: a) 1826 a 1920, primeira fase; b) 1920 a 1952, ano de publicação do decreto presidencial que previa a elaboração de um atlas nacional, a segunda fase; elas que delimitam o fim desta fase, vale ressaltar, e c) 1952 em diante (fase que permanece “aberta” ou vigente até a data da publicação das autoras). Essa fase, a terceira, delineia-se por apresentar algumas características, sobretudo pela publicação de obras muito valiosas para o desenvolvimento dos estudos dialetais em solo brasileiro, tais como: em 1957, obra de Serafim da Silva Neto, *Guia para os estudos dialetológicos*; em 1958, obra de Antenor Nascentes, *Bases para elaboração do Atlas Linguístico do Brasil*; em 1955, a obra de Antenor Nascentes, *Divisão dialectológica do território brasileiro*.

Para além dessas obras que marcaram a terceira fase, também merece destaque o pioneiro passo concreto da Geolinguística no Brasil, a geolinguística baiana, com a

publicação do *Atlas Prévio dos Falares Baianos* – APFB – (ROSSI, 1963) – em 1963, sob a coordenação de Nelson Rossi, tendo como colaboradoras principais, Carlota Ferreira e Dinah Iseense, e colaboração de recém-licenciadas do curso de Letras da Universidade Federal da Bahia (UFBA). Também, menciona-se o surgimento, em 1969, do Projeto NURC, cujos pormenores estão descritos na subseção 5.1.

A morfossintaxe, na terceira fase, passa a ter um pouco mais de relevância, mas ainda continua restrita e até ausente em alguns trabalhos/atlas linguísticos estaduais e regional.

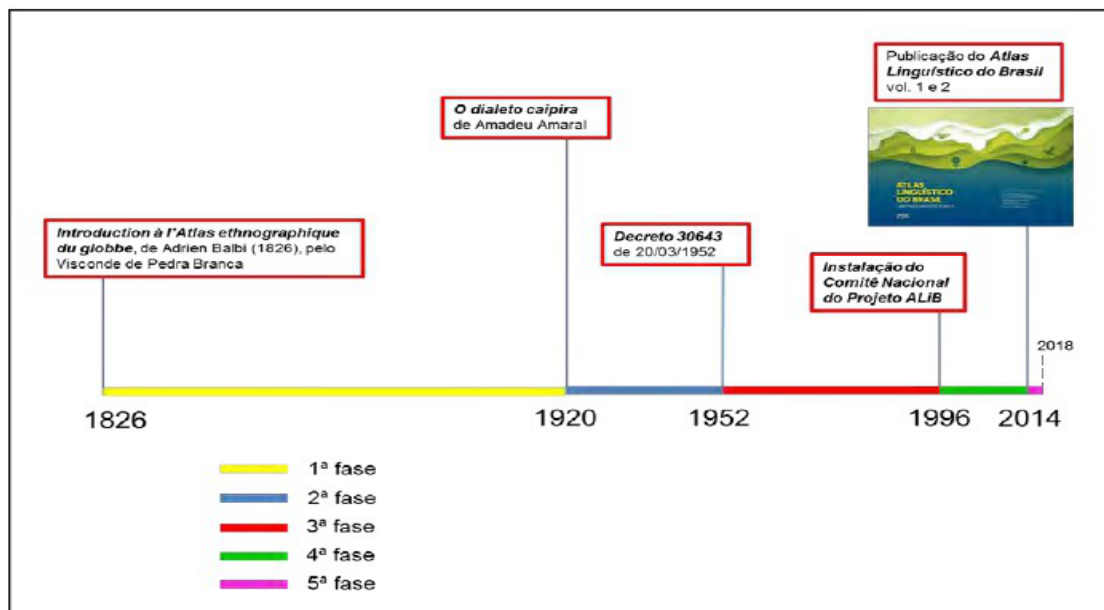
Em 2006, Mota e Cardoso revisitam a proposta de Ferreira e Cardoso (1994), e estabelecem, para a história dos estudos dialetais brasileiros, uma nova fase, a quarta. Segundo as autoras, a 3ª fase se finda em 1996, momento em que se inicia um novo momento, na Bahia, com a retomada dos intentos de elaboração do atlas nacional, no âmbito da língua portuguesa, com o surgimento do Projeto ALiB, com algumas características notáveis, tais como: a implementação e a consolidação, em várias universidades do Brasil, dos estudos no campo da Geografia Linguística. Deve-se destacar que, segundo as propositoras desta fase:

Do ponto de vista metodológico, essa nova fase coincide com a incorporação dos princípios implementados pela Sociolinguística a partir da década de 60 do século passado, abandonando-se a visão monodimensional – monoestrática, monogeracional, monogenérica, monofásica etc. – que predominou na geolinguística hoje rotulada de tradicional (MOTA; CARDOSO, 2006, p. 21).

Nota-se, na quarta fase, um aumento da importância da morfossintaxe, por exemplo, com a inclusão de questionário especificamente morfossintático nos Questionários do Projeto ALiB. Logo, é perceptível que o foco dos pesquisadores se amplia, não somente priorizando aspectos lexicais e fonéticos, conforme a tradição, mas também incluindo aspectos outros que revelam facetas sobre a língua falada no Brasil.

Retomando o que foi exposto do Quadro 1, Teles (2018), em sua tese de doutoramento, destaca alguns aspectos que, de certo modo, sinalizam para o surgimento, no que tange à periodização dos estudos geolinguísticos brasileiros, de uma nova fase, neste caso, “[...] propõe-se agora a inclusão de uma nova fase, encerrando-se a quarta fase em 2014” (TELES, 2018, p. 79). Ainda, segundo os argumentos da referida autora, com a publicação dos dois primeiros volumes do atlas nacional, novos rumos foram tomados pelos estudos dialetais brasileiros. Tal proposição pode ser visualizada na Figura 1:

Figura 1 – Periodização dos estudos dialetais brasileiros.



Fonte: Teles (2018, p. 81).

Nota-se que, a partir da apreciação da Figura 1, Teles (2018) considerou as propostas anteriores, validando, assim, os marcos estabelecidos pelos respectivos dialetólogos propositores (NASCENTES, 1952/1953; FERREIRA; CARDOSO, 1994; MOTA; CARDOSO, 2006). A autora, ainda, ressalta os avanços da Dialetologia, quanto às produções científicas, ao crescimento dessa área da Linguística no Brasil, evidenciados pelos atlas estaduais e regional publicados e afirma que

[...] Mas nada é tão marcante quanto a interdisciplinaridade e a transdisciplinaridade que a caracterizam. Na área linguística isso fica claro com os trabalhos que foram e vêm sendo desenvolvidos com o *corpus* do Projeto ALiB, em estudos lexicográficos e sociolinguísticos, por exemplo. O fato de se utilizar, como qualquer ciência moderna, dos novos recursos tecnológicos, tanto torna a linguística uma usuária em potencial desses recursos (pela grande quantidade e variedade de dados que manipula) como desperta a atenção de cientistas de outras áreas que nela veem um vasto celeiro de oportunidades para aplicarem e desenvolverem suas ciências, como pode ser exemplificado com o envolvimento de pesquisadores de outras áreas (notadamente os especialistas em tecnologia da informação e os da ciência cartográfica) tanto na equipe do Projeto ALiB quanto nas de outros projetos regionais (TELES, 2018, p. 78-79).

Para a adição de uma nova fase na historiografia da Dialetologia brasileira, a autora argumenta que há algumas razões, a saber:

De lá para cá, independentemente de iniciar-se – ou não – uma nova fase, é indiscutível a argumentação de que esse é um marco da Geolinguística brasileira: tanto pelo fato de inúmeros trabalhos terem sido desenvolvidos, desde então, a partir do seu conteúdo, quanto pela extensão alcançada, seja pela disponibilização do atlas em bibliotecas não apenas das Universidades integrantes do Projeto, seja pela

divulgação em âmbito nacional que o atlas teve a partir de telejornais de coberturas locais e nacionais, além de uma série composta de vários programas, abordando praticamente todas as áreas de estudos linguísticos que o ALiB contempla (TELES, 2018, p. 79-80).

Além disso, Teles (2018) apresenta outros motivos que fundamentam a proposição do alvorecer desta nova fase, tais como: o aumento progressivo e exponencial do número de alunos interessados pelos estudos geolinguísticos/dialetais nos cursos de pós-graduação nos diversos programas espalhados pelo país; o aspecto cartográfico e:

Não apenas do ponto de vista do conteúdo, o ALiB representa um marco: do ponto de vista cartográfico, também se trata de uma publicação inovadora, especialmente por ser o primeiro atlas linguístico no Brasil a ter sido concebido para utilização em Sistemas de Informações Geográficas, não somente pelo fato de ter sido utilizada uma base cartográfica digital, oficial, mas também porque todas as feições geográficas estão perfeitamente construídas para esse fim. Com isso, outra característica ímpar é o georreferenciamento de todas as localidades, tornando todas cartas publicadas passíveis de reedições e complementações a qualquer época. **Diante do exposto e pelas razões elencadas, propõe-se que a publicação dos dois primeiros volumes do ALiB seja considerada o marco inicial de uma quinta fase dos estudos da Dialetologia no Brasil** (TELES, 2018, p. 80) [grifos nossos].

Até então, nesta quinta e última fase, nota-se um reflexo do que, historicamente, foi construído na Dialetologia brasileira, inúmeros trabalhos e pesquisas nas perspectivas lexicais e fonéticas, ao passo que as morfossintáticas são em número reduzido, embora, cada vez mais, tenham despertado o interesse por parte dos pesquisadores.

Para além da periodização da Dialetologia do ponto de vista histórico, cabe explorar ainda alguns avanços da Dialetologia, no que se tange ao método geolinguístico. Nessa direção, Romano (2013/2020) apregoa que a Geolinguística brasileira está segregada em dois grandes e distintos momentos, que, respectivamente, evidenciam-se pela metodologia e pelas características das obras produzidas nesses, principalmente relacionando-os com os atlas publicados. Para ele, o primeiro momento tem início com a publicação do APFB, em 1963, e se estende até 1996, com o alvorecer do Projeto ALiB, consoante argumentos do referido autor:

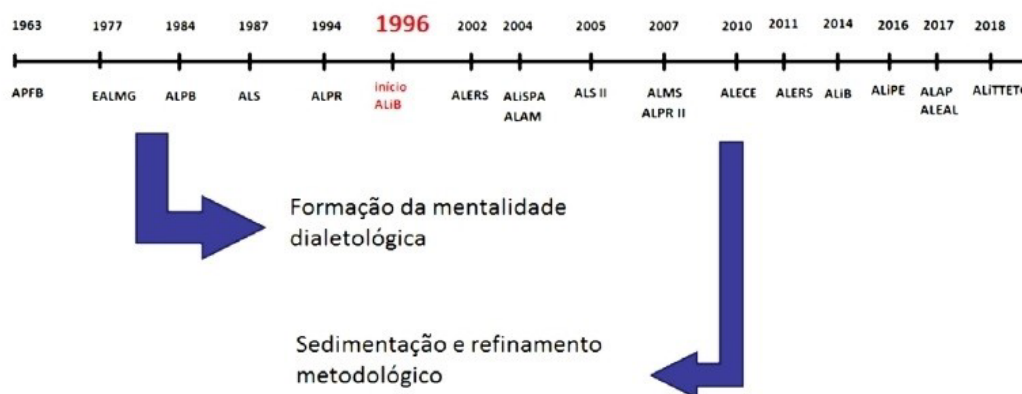
[...] Os atlas linguísticos desenvolvidos ou projetados nesse período de 33 anos seguem pressupostos metodológicos anteriores aos do ALiB, portanto divergem quanto à metodologia, ora priorizando a diatopia [...] ora incluindo outras dimensões da variação, como a diagenérica [...] Portanto pertencem a esse primeiro momento da Geolinguística brasileira, basicamente, nove atlas estaduais: APFB (ROSSI, 1963), EALMG (RIBEIRO et al., 1977), ALPB (ARAGÃO; BEZERRA DE MENEZES, 1984), ALS (FERREIRA et al., 1987), ALPR (AGUILERA, 1994), ALS II (CARDOSO, 2005), ALPR II (ALTINO, 2007), ALECE (BESSA et al., 2010) ALERS (KOCH; ALTENHOFEN; KLASSMAN, 2011) (ROMANO, 2020, p. 13).

Por sua vez, o segundo momento tem início em 1996, a partir da implantação do Projeto ALiB e se estende até os dias atuais. Desse modo, portanto, apresenta outros aspectos, a saber:

Os trabalhos desse período já apresentam influência dos procedimentos metodológicos do ALiB, caminhando para uma visão pluridimensional da variação, ou seja, adentram os “veios sociolinguísticos”, segundo Cardoso (2010). Os atlas ou trabalhos geolinguísticos desenvolvidos e em desenvolvimento, atualmente, têm buscado acrescentar à dimensão diatópica variáveis diversas, tais como faixa etária, sexo, religião, etnia, origem dos informantes, comunidades bilíngues/multilíngues, níveis de interlocução, entre outras, com o intento de contemplar de forma mais ampla possível as diferentes dimensões. Houve, portanto, a sedimentação de uma metodologia geolinguística por conta do controle de diferentes variáveis além da diatópica (ROMANO, 2020, p. 15).

A segmentação em dois momentos da geolinguística brasileira pode ser entendida assim, segundo Romano (2020):

Figura 2 – Os momentos da Geolinguística no Brasil.



Fonte: Romano (2020, p. 15).

Em concordância com a última proposta de periodização dos estudos dialetais, sob a égide da quinta fase, e com dois momentos da geolinguística brasileira, vislumbra-se um cenário de desenvolvimento e com perspectivas diversas para a Dialectologia e a Geolinguística brasileiras, não somente pela expectativa dos novos volumes do atlas nacional, com análises das cartas publicadas no volume dois e nos volumes com dados do interior, mas também pelo crescimento gradativo dos estudos e, conseqüentemente, do aprimoramento dos métodos e das ferramentas desse ramo da Linguística, a cada dia, mais interdisciplinar e mais agregadora.

2.1.2 A pluridimensionalidade está na Dialektologia ou na Geolinguística?

A fim de responder ao questionamento que nomeia esta subseção, vale, de antemão, pensar sobre as dimensões dos trabalhos e frutos geolinguísticos, os atlas linguísticos, fazendo alusões ao contexto externo e ao contexto interno, com intuito de encontrar possíveis caminhos para o entendimento e, quiçá, para uma resposta que, até então, não é consensual entre os dialetólogos.

Os atlas são instrumentos valiosos para salvaguardar formas linguísticas, além disso: “[...] Os atlas permitem ainda isolar áreas dialektais, definidas pela concentração de fenômenos linguísticos idênticos...” (FERREIRA et al., 1996, p. 484). No Brasil, nesse sentido,

É pertinente, portanto, considerar no território brasileiro a existência do atlas nacional, o ALiB, um atlas regional, o ALERS, os atlas estaduais, já publicados ou em desenvolvimento e os atlas de pequeno domínio. Assim, na terminologia proposta por Alinei (1994) e remodelada para a realidade brasileira, ter-se-ia: a) atlas continentais; b) grupos de língua; c) nacionais; d) regionais; e) estaduais e f) pequeno domínio/locais, os quais se aplicam na realidade brasileiras os do tipo c, d, e, f (SILVA, ROMANO, 2022, p. 22).

Os atlas brasileiros, por exemplo, conforme postulações de Thun (2005), podem ser classificados segundo a quantidade de dimensões de variações contempladas. As perspectivas monodimensional, bidimensional e pluridimensional são bastante esclarecedoras quanto à natureza dos trabalhos, as intenções dos pesquisadores, bem como dos meios fornecidos pela época para a tarefa de cartografar dados linguísticos. Cabe frisar que a fotografia sobre os dados da língua dependerá da intencionalidade e do ângulo de observação do dialetólogo, não admitindo a ideia de que uma perspectiva é superior a outra, mas todas, de algum modo, ajudam a entender um pouco mais e melhor a língua e os fenômenos observados.

A primeira, a perspectiva monodimensional, pode ser compreendida por ter o foco apenas na dimensão diatópica, além de ter como meta, por vezes, a identificação de áreas dialektais, a partir da presença e/ou ausência de dados linguísticos e o traçado de linhas virtuais, as isoglossas. Para obter esses resultados, a fim de preservar a fala isolada, buscavam-se informantes que tivessem o perfil *NORMs* (CHAMBERS; TRUDGILL, 1994, p. 58-59), sendo este sem nenhuma mobilidade, idoso, rural, com nível de escolaridade baixo e homem. Dessa maneira, caracterizando os estudos como monostráticos, monogeracionais e monofásicos.

Aludindo ao contexto externo, têm-se como exemplos de atlas que contemplam uma única dimensão os atlas⁵ de Wenker, e de Gilliéron, em 1887, ambos em fins do século XIX. Ainda, consoante à visão e aos produtos geolinguísticos de uma única dimensão, agora, no contexto interno, têm-se com exemplos: o *Atlas Prévio dos Falares Baianos* – APFB – (ROSSI, 1963); o *Esboço de um atlas linguístico de Minas Gerais* – EALMG – (RIBEIRO et al., 1977) e o *Atlas Linguístico da Paraíba* – ALPB – (ARAGÃO; BEZERRA DE MENEZES, 1984). Assim, os atlas supramencionados possuem uma visão espacial dos dados linguísticos, contemplam a variação diatópica, enquadrando-se em atlas, essencialmente, monodimensionais.

A segunda perspectiva, a bidimensional, abarca dois aspectos, geralmente, o espaço geográfico aliado ao aspecto social – sexo, escolaridade, faixa etária, etc – por exemplo. Os atlas considerados de duas dimensões controlam e retratam, de forma sistemática, essas duas variáveis: a variável espacial e a variável social. Como exemplos de atlas que se enquadram na referida bidimensionalidade tem-se, para o contexto interno, publicado em 1987, o *Atlas Linguístico de Sergipe* – ALS – (FERREIRA et al., 1987). Este contemplou, além da diatopia, a variação diassexual.

A perspectiva pluridimensional, oriunda da implementação dos princípios norteadores da Sociolinguística Laboviana, a partir da década de 60, trouxe essa possibilidade de observação, para além das dimensões anteriormente mencionadas, haja vista que

Assim, hoje, torna-se imperativo, por exemplo, incluir, entre os critérios de escolha dos indivíduos que servirão de informantes para a formação do *corpus* de um atlas linguístico, variáveis como idade, sexo, nível de instrução, ou mesmo situação socioeconômica, a fim de que se revelem ao máximo as particularidades do sistema dialetal focalizado e se possam melhor conhecer os condicionamentos socioculturais que presidem a distribuição geográfica dos fenômenos linguísticos (BRANDÃO, 1991, p. 26).

Também nomeada por contemporânea (THUN, 1998), a visão pluridimensional centra-se na observação de mais de duas variáveis, quais sejam as dimensões: a) diatópica – que envolve informantes nativos de determinada região geográfica; b) diastrática – que investiga informantes de classe sociais distintas; c) diassexual – que contempla informantes dos dois sexos, homem e mulher; d) dialingual – quando o atlas investiga duas ou mais línguas; e) diatópico-cinética – contempla informantes naturais da área pesquisada e dos informantes que migraram para ela; f) diageracional – abarca informantes de faixas etárias diferentes; g) diafásica – envolve graus de interlocução distintos; h) diarreferencial – diz

⁵ Respectivamente, *Sprachatlas Von Nord- und Mitteldeutschland, auf Grund Von systematisch MIT Hilfe der Volksschuler gesammeltem Material aus circa 30.000 Orten*; *Atlas Linguistique de la France*.

respeito ao tipo de questões utilizadas e tratamento dos dados, explorando aspectos metalinguísticos, por exemplo.

Com o progresso da ciência dialetal e, conseqüentemente, desenvolvimento dos métodos e das ferramentas, faz-se necessário levar em consideração que:

Do ponto de vista linguístico, o esquadramento da estrutura das línguas naturais levou ao conhecimento detalhado dos elementos que as constituem, das estruturas segundo as quais se organizam, das relações históricas que entre grupos se estabelecem, dos processos que marcam seu funcionamento. Do ponto de vista sócio-histórico e político, verifica-se que o mundo vem passando – como, aliás, sempre passou – por profundas e sucessivas transformações que lhe oferecem nova configuração e novo perfil (CARDOSO, 2010, p. 62).

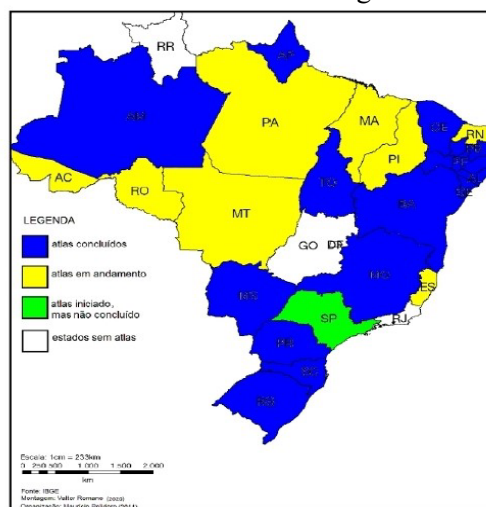
Dessa maneira, configuram-se como atlas que se enquadram na pluridimensionalidade, no contexto externo, destaca-se o *Atlas Diatopico y Diastratico del Uruguay* – ADDU – (THUN; FORTE; ELIZAINCÍN, 2000); e, no contexto interno, o *Atlas Linguístico do Brasil* – ALiB (CARDOSO et al, 2014).

No tocante ao Brasil, conforme assevera Cardoso (2010), o princípio da Geolinguística brasileira está atrelado ao começo de estudos sistemáticos e à publicação do APFB, de Nelson Rossi, em 1963. Desse modo, portanto, na Universidade Federal da Bahia – importante celeiro dialetológico – nascem a primeira manifestação concreta desse método; o segundo atlas a ser iniciado em solo brasileiro – ALS; e, por fim, a retomada do tão sonhado atlas nacional – ALiB. Conforme se nota, os três atlas – APFB, ALS e ALiB – são bem representativos para a Geolinguística do Brasil, bem como ilustram as dimensões e os avanços do método cartográfico dialetal. Nesse sentido, podem ser considerados como:

[...] o tripé fundamental para os estudos geolinguísticos brasileiros, a saber: o **APFB** (ROSSI, 1963); o **ALS** (FERREIRA et al., 1987); e o **ALiB** (CARDOSO et al., 2014). A proposição do tripé elementar explica-se pelos seguintes motivos: aquele, foi o atlas pioneiro elaborado no país, foi o pontapé inicial dos estudos geolinguísticos e, certamente, tornou-se referência para o desenvolvimento e aprimoramento dos sucessores frutos da Geolinguística brasileira, atlas **monodimensional**; esse, por sua vez, instaura o caráter **bidimensional** no âmbito da pesquisa dialetal, embora tenha sido publicado muito depois, foi o segundo atlas iniciado, com intencionalidade de continuação da pesquisa feita na Bahia; por fim, este, por ser uma meta alcançada, o sonho idealizado e concretizado por gerações de dialetólogos e dialetólogas do Brasil. Para além disso, constitui-se como um material valioso, e é obra de caráter **pluridimensional** (SANTOS; RIBEIRO, 2021, p. 310) [grifos nossos].

É necessário, hoje, notar o desenvolvimento da pesquisa Geolinguística brasileira, conforme se percebe na Figura 3, por meio da expansão dos atlas estaduais/regionais publicados e dos trabalhos que estão sendo desenvolvidos.

Figura 3 – Panorama dos atlas estaduais/regionais brasileiros (2020).

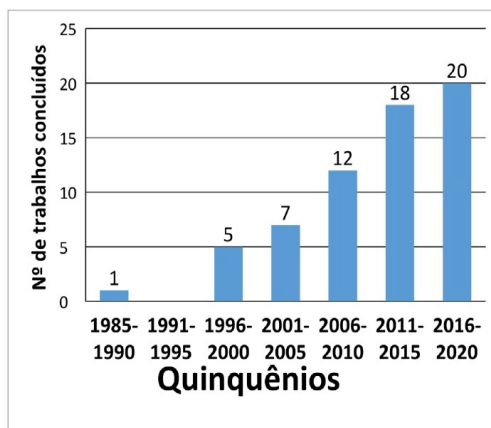


Fonte: Romano (2020, p. 16).

Nessa perspectiva, ainda sobre a referida expansão e consolidação dos métodos geolinguísticos brasileiros, sumariza Romano (2020):

Atualmente, são 14 os estados que têm seu atlas linguístico concluído (Rio Grande do Sul, Santa Catarina, Paraná, Mato Grosso do Sul, Minas Gerais, Tocantins, Bahia, Sergipe, Alagoas, Pernambuco, Paraíba, Ceará, Amapá e Amazonas), sendo dois deles com mais de um atlas, Paraná e Sergipe. Oito são os estados cujo atlas está em andamento (Acre, Rondônia, Mato Grosso, Pará, Maranhão, Rio Grande do Norte, Piauí e Espírito Santo). Há um estado em que o atlas foi iniciado, mas não foi concluído, São Paulo, e três estados sem atlas linguístico (Goiás, Roraima e Rio de Janeiro) (ROMANO, 2020, p. 17).

Vale frisar que os atlas de pequeno domínio – geralmente, no Brasil, são atlas que recobrem pequenas áreas geográficas, no que tange à extensão e ao número de pontos, são frutos de dissertações e teses defendidas em Programas de Pós-graduação –, também, fornecem fotografias valiosas sobre aspectos do falar brasileiro em dada localização geográfica e, para além disso, ilustram o desenvolvimento da pesquisa dialetal, conforme o Figura 4:

Figura 4 – Número de atlas de pequeno domínio (concluídos) no Brasil

Fonte: Romano (2020, p. 19).

Ainda, no tocante a esses atlas,

[...] os atlas de pequeno domínio funcionam como um tipo de zoom para as realidades maiores e são importantes instrumentos na análise e na interpretação de nuances dialetais de um território, pois lançam um olhar mais aprofundado, por vezes não contempladas em macroatlas, tais como as línguas minoritárias, as atividades artesanais que demarcam a economia local e que são decorrentes do contato do homem com seu meio físico, questões essas que caracterizam os diferentes pontos de uma grande região (SILVA, ROMANO, 2022, p. 23).

Em constante progresso, a Dialectologia brasileira dispõe de muitas obras e várias pesquisas vinculadas aos Programas de Pós-Graduação do país, em virtude, sobretudo, da colaboração e inspiração das obras mencionadas anteriormente, que fomentaram e subsidiaram os estudos sobre o conhecimento da realidade linguística e prepararam os caminhos para o tão sonhado atlas nacional, ajudando a descortinar, desse modo, as reais facetas das regiões dialetais do Brasil.

Ressaltando o caráter essencial dos frutos geolinguísticos, quer sejam de grande extensão, quer sejam de pequeno domínio, vale destacar que os atlas linguísticos revelam muitas histórias, por exemplo, contam sobre as transformações pelas quais as localidades passaram e, conseqüentemente, contam também sobre a língua falada pelos habitantes dessas, haja vista que a língua possui um vínculo indissociável com a sociedade. O testemunho dos atlas, representado pelas cartas linguísticas, guarda inúmeras variantes, arcaicas e inovadoras, e fornece fotografias simbólicas de períodos sincrônicos de uma língua. Portanto, um atlas é dotado de valores imensuráveis, com potencial de cunho linguístico, cultural e histórico, pois, por meio de dados coletados e tratados com o rigor científico, pode revelar os usos linguísticos, as facetas sociais de um povo e as itinerâncias dos homens pelo espaço geográfico.

Após o preâmbulo, que, de certo modo, ilustra e conduz a argumentação para responder à pergunta que dá nome a esta subseção – a pluridimensionalidade está no ramo da Linguística responsável pela variação diatópica ou no método geolinguístico? –, volta-se à perspectiva para a atribuição do adjetivo pluridimensional, ora empregado para caracterizar a Dialetologia, ora empregado para caracterizar a Geolinguística. Em virtude da falta de consenso na vasta literatura dialetal, no que concerne a essa divergência, fazer-se necessário trazer à tona alguns pensamentos e argumentos de dialetólogos, a fim de elucidar a referida questão.

A pluridimensionalidade refere-se ao método, haja vista que, conforme Cardoso (2010, p. 88): “Essa geolinguística pluridimensional, que se consolida ao final do século XX, vai ocupar-se de controle sistemático de variáveis sociais, mas terá que administrar, competentemente, a pluralidade de dados, sobretudo na cartografia dos resultados obtidos”. Ainda, nesse sentido: “A geolinguística pluridimensional deve pensar em uma organização inteligente e racional da grande quantidade de dados que recolhe” (THUN, 2000, p. 408). Logo, entende-se que o método, portanto, a Geolinguística, é quem deva receber o qualificador, pluridimensional, extinguindo, assim, a confusão terminológica, quando aplicado indistintamente tanto para a ciência dialetal, Dialetologia, quanto para o seu método, Geolinguística.

Com intuito de ilustrar sobre a adjetivação que deve ser aplicável ao método, a geolinguística, ressalta-se que os produtos da Dialetologia podem ser de dois tipos: Estudos monográficos (AMARAL, 1920; NASCENTES, 1922; 1953, MARROQUIM, 1934) ou atlas linguísticos (ROSSI, 1963, CARDOSO et al. 2014), por exemplo.

É consensual que a Dialetologia surge, eminentemente, com a preocupação única em catalogar a fala rural e com ênfase em apenas uma dimensão, a diatópica, esta última sendo a prioridade dessa ciência até os dias atuais. Com o passar do tempo, sobretudo, com o desenvolvimento, ao incorporar os princípios apregoados pelos estudos sociolinguísticos, os estudos dialetológicos passaram a observar, também, a fala urbana e a considerar, controlando de forma sistemática, as variáveis sociais, tais como: sexo, grau de escolaridade, faixa etária, profissão, etc., aliando, por vezes, várias dimensões na forma de representar os dados encontrados. Assim, segundo Cardoso (2000, p. 415),

Creio que a Geolinguística hoje, neste final de milênio, deve continuar a **priorizar a variação diatópica**, abrindo, porém, espaço para o **controle de outras variáveis como gênero, idade e escolaridade, sem a busca obcecante da quantificação**, mas tomando-as, de forma exemplificativa e não exaustiva, de modo a complementar os próprios dados areais (CARDOSO, 2000a, p. 415) [grifos nossos].

Historicamente, não são notados os qualificadores monodimensional e bidimensional atribuídos à Dialectologia, desse modo, o termo pluridimensional deve acompanhar o método, isto é, a Geolinguística, uma vez que os produtos oriundos desse fazer dialetológico possuirão as faces mono, bi ou pluri, conforme intentos do pesquisador e orientação do desenho metodológico da pesquisa.

Ao perseguir e examinar os traços das escolhas pronominais, em posição de sujeito, nas estruturas como *para/prá + eu/mim + infinitivo*, feitas pelos informantes das capitais brasileiras, em tempos distintos, sejam com os dados do Projeto NURC, sejam com os dados do Projeto ALiB, intenciona-se mapear a configuração dessas ocorrências conforme a sua distribuição espacial, utilizando os pressupostos teóricos da responsável pela variação areal, a Dialectologia. Além disso, tal mapeamento estende-se aos aspectos sociais – sexo; escolaridade; faixa etária –; dentre outros, quando possíveis, aludindo à responsável pela variação social, a Sociolinguística Variacionista. Quanto possível, tais aspectos serão representados por meio das cartas linguísticas, com orientação da Geolinguística Pluridimensional.

2.1.3 Nos passos de Nascentes (1953): os estudos sobre as regiões dialetais do Brasil

No que tange ao estabelecimento de áreas dialetais brasileiras, as primeiras considerações são do final do século XIX, vários autores propuseram divisões⁶, que podem ser, resumidamente, visualizadas no Quadro 2.

Quadro 2 – Propostas de divisão dialetal do Brasil.

ANO	AUTOR	PROPOSTA	CRITÉRIOS
1981	Júlio Ribeiro	Quatro áreas: norte; leste; centro e sul.	Exclusivamente geográfico.
1900	João Ribeiro	Cinco áreas: o extremo norte – Amazônia, Maranhão, Piauí e Ceara; o norte – Alagoas, Pernambuco, Paraíba e Rio Grande do Norte; o centro – Sergipe, Baía, Ilhéus e Porto Seguro; o interior – São Paulo, Minas Gerais, Goiás e Mato Grosso; o sul – Espírito Santo, Rio de Janeiro, Santa Catarina	Exclusivamente histórico.

⁶ Maiores informações e detalhamentos dos critérios utilizados por esses autores, bem como os mapas propostos podem ser vistos no texto de Nascentes (1955) que está em Barbadinho Neto (2003) e Teles (2018).

		e Rio Grande do Sul.	
1915	Rodolpho Garcia	Cinco zonas linguísticas: norte; norte-oriental; centralmarítima; meridional e altiplana-central.	Geográficos e históricos.
1922	Antenor Nascentes	Quatro áreas: nortista; fluminense; sertaneja; e, sulista.	Linguísticos.
1933	Antenor Nascentes	Cinco áreas: norte; nordeste; fluminense, sul; e sertão.	Linguísticos.
1936	Renato Mendonça	Nove áreas: amazônica; cearense; nordestina; baiana; fluminense; caipira; mineira; gaúcha; e sertaneja.	Étnicos e históricos, com base nos glossários de localismos e monografias dialetológicas.
1941	Eugenio de Castro	Três geografias: do gado; das bandeiras paulistas; e a geografia da canôa.	Toponímicos e históricos.
194-	Joaquim Ribeiro	Cinco áreas: costeira; agrícola; pastoril; mineração; e amazônica.	Geográfico, histórico, orientação sócio-cultural.
1950	Maximino Maciel	Tripartida: basilo-guianense ou setentrional; idiodialetos estaduais ou centrais e basilocastelhano ou meridional.	Exclusivamente geográfico.
1953	Antenor Nascentes	Dois grandes falares: Norte e Sul, contendo subfalares: Amazônico e Nordeste; Sulista, Mineiro, Fluminense e Baiano, respectivamente.	Linguísticos.
1960	Manuel D. Júnior	12 regiões culturais: nordeste agrário do litoral; nordeste mediterrâneo; Amazônia; mineração do planalto; centrooeste; extremos-sul; área de colonização estrangeira; área do café; faixa industrial; cacau; sal; e pesca.	Processo de ocupação humana do território.
1975 [1963]	Silvio Elia	Nove áreas: amazônica; cearense; nordestina; baiana; fluminense; caipira; sulina; centro leste; e centro-oeste.	Linguísticos.

Fonte: Elaborado pelo autor (2023) com base em Barbadinho Neto (2003) e Teles (2018).

Nota-se que, após a contemplação cronológica exposta no Quadro 2, algumas tentativas de delimitação dialetal para o Brasil foram apresentadas. No entanto, a última proposição de Nascentes, em 1953, tornou-se a referência para os estudiosos interessados no assunto. Desse modo, as propostas do referido autor, 1922/1933/1953⁷, serão detalhadas a seguir.

Na primeira proposta de divisão elaborada por Antenor Nascentes, em 1922, publicada na obra *O linguajar carioca em 1922*, há um mapa dialetal em que constam quatro áreas, a saber: 1) nortista; 2) fluminense; 3) sertaneja; e 4) sulista; delineadas assim:

o nortista (Amazonas, Pará, litoral dos Estados desde o Maranhão até a Baía), o fluminense (Espírito Santo, Rio de Janeiro, sul de Minas e zona da mata, Distrito Federal), o sertanejo (Mato Grosso, Goiás, norte de Minas, sertão dos Estados litorâneos desde o Maranhão à Baía) e o sulista (S. Paulo, Paraná, Santa Catarina, Rio Grande do Sul e triângulo mineiro) (NASCENTES, [192-], p. 20-21).

O mapa dialetal, Figura 5, ilustra a configuração dessas áreas:

Figura 5 – Mapa dialetal de Antenor Nascentes (1922).



Fonte: Nascentes (1955, p. 216).

Tal proposição recebeu algumas críticas, das quais se destacam as feitas por Lindolfo Gomes, na Revista de Filologia e História II, validadas por Nascentes, uma vez que “Quando fizemos aquela divisão, havíamos percorrido pequena parte do nosso território. Conhecíamos Minas Gerais, Espírito Santo, estado do Rio de Janeiro, São Paulo, Paraná, Santa Catarina, Rio Grande do Sul, Recife e Salvador” (NASCENTES, 1955, p. 217).

A segunda proposta, a 1933, é pouco conhecida até mesmo entre os dialetólogos, haja vista que há somente um mapa, conforme Figura 6, que faz alusão a essa proposição. Nascentes (1953) faz uma referência a ela: “Atendendo às ponderações de L. Gomes, alterei

⁷ São três obras diferentes, todas estão indicadas na seção de referências desta tese.

minha divisão de 1922 em 1933, quando fiz o mapa dialectológico do Brasil, constante da página 244 do quarto volume da minha serie, *O Idioma Nacional*” (NASCENTES, 1953, p. 23).

Figura 6 – Mapa dialetal de Antenor Nascentes (1933).



Fonte: Nascentes (1933, p. 244).

Nascentes (1953), ao observar a cadência e abertura das vogais médias em posição pretônica, divide o Brasil em dois grandes grupos de falares – os falares do Norte e os falares do Sul – que foram subdivididos em seis subfalares – Amazônico e Nordestino, compondo os do Norte; Baiano, Fluminense, Mineiro e Sulista, que pertencem aos do Sul. Além de um território incharacterístico, conforme Figura 7.

Figura 7 – Mapa dialetal de Antenor Nascentes (1953)



Fonte: Nascentes (1953, p. 18-19).

Nas palavras do referido autor:

Hoje que já realizei o meu ardente desejo de percorrer todo o Brasil, do Oiapoc ao Xuí, de Recife a Cuiabá, fiz nova divisão que não considero nem posso considerar

definitiva, mas sim um tanto próxima da verdade. [...] Dividimos então o falar brasileiro em seis subfalares que reuni em dois grupos a que chamei do norte e do sul. O que caracteriza êstes dois grupos é a cadência e a existência de protônicas abertas em vocábulos que não sejam diminutivos nem advérbios terminados em –mente. Basta uma singela frase, ou mesmo uma simples palavra, para caracterizar as pessoas pertencentes a cada um dêstes grupos (NASCENTES, 1955, p. 217).

Na proposição de 1953, Nascentes divide o Brasil em dois grandes grupos, subdivididos em seis subfalares, que compreendem:

Os subfalares do norte são dois: o amazônico, que abrange o Acre, o Amazonas, o Pará e a parte de Goiás que vai da foz do Aiqui à serra do Estrondo, e o nordestino, que compreende os estados do Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e a parte de Goiás, que vai da serra do Estrondo à nascente do Parnaíba (NASCENTES, 1955, p. 217).

E, por sua vez:

Os subfalares do sul são quatro: o baiano, intermediário entre os dois grupos, abrangendo Sergipe, Bahia, Minas (Nordeste, Norte e Noroeste), Goiás (parte que vem da nascente do Parnaíba, seguindo pelas serras dos Javaés, dos Xavantes, do Fanha e do Pilar até a cidade do Pilar, rio das Almas, Pirenópolis, Santa Luzia e Arrendidos); o fluminense, abrangendo o Espírito Santo, o estado do Rio de Janeiro, o Distrito Federal, Minas (Mata e parte Leste); o mineiro (Centro, Oeste e parte do Leste de Minas Gerais); o sulista, compreendendo São Paulo, Paraná, Santa Catarina, Rio Grande do Sul, Minas (Sul e Triângulo), Goiás (Sul) e Mato Grosso (NASCENTES, 1955, p. 217).

Além disso, dos dois grandes pólos linguísticos, há uma área assim descrita e, conseqüentemente, nomeada:

O perímetro compreendido entre a parte da fronteira boliviana (rios Verde, Guaporé, Mamoré até o Abunã), a fronteira de Mato Grosso com Amazonas e Pará (rios Madeira, Juparaná, Uruguatás, Tres Barras, Cariaí, Xingu, Fresco e Aiqui e a serra do Estrondo), é praticamente despovoado e portanto incaracterístico (NASCENTES, 1953, p. 25-26).

Nessa perspectiva, ainda, de acordo com o mapa de 1953, o próprio autor não a considera como definitiva, uma vez que, nas palavras dele: “Nosso trabalho, repetimos, não é e nem podia ser definitivo. Aguardemos o Atlas Lingüístico do Brasil (até quando?), para um trabalho definitivo” (NASCENTES, 1955, p. 99).

Embora duas outras divisões, conforme o Quadro 4, tenham sido publicadas, por diferentes autores, é a última proposição de Nascentes (1953) que serve de parâmetro para os estudiosos da área, quando se referem à divisão dialetal do Brasil, mesmo apresentando uma imprecisão dos limites de cada falar e subfalar, pela carência de elementos de localização geográfica e/ou uma escala que permitisse atestar, de forma mais precisa, tais contornos.

Teles (2018), ao utilizar aparatos tecnológicos modernos, mais precisamente, da Cartografia automatizada, a base BCIM, do Instituto Brasileiro de Geografia e Estatística

(2016) – IBGE – e o banco de dados Sistemas de Informações Geográficas – SIG – em sua tese de doutoramento, investigou textualmente e cartografou com precisão, pelo viés geográfico/cartográfico, o mapa dialetal proposto por Nascentes em 1953.

A referida autora faz diversas considerações sobre o mapa de 1953, tais como:

Acerca da divisão dialetal, a importância do primeiro documento está no fato de que, apesar de outras propostas terem sido formuladas, apresentando não necessariamente todas em conjunto ou mesmo nessa ordem, critérios geográficos, históricos e linguísticos, a de Antenor Nascentes (NASCENTES, [192-]; 1933; 1953) é a mais citada, referenciada e tem servido como base para quase todos os estudos de natureza geolinguística que buscam confirmar e atestar a pertinência dessa divisão nos dias atuais, identificando novas áreas dialetais e, conseqüentemente, novas divisões ou subdivisões. Apesar da inserção de um *mapa* às páginas 18-19 (NASCENTES, 1953), Figura 1 [vide Figura 7, nesta tese], este é de pouca valia já que não há como nele identificar, com segurança, os elementos contidos na descrição textual dos limites demarcadores de regiões ou de áreas dialetais, salvo os contornos dos limites políticos dos Estados da Federação à época da sua elaboração (TELES, 2018, p. 40-41) [grifos nossos].

Além disso, Teles (2018), conforme Figura 8, publica um mapa atualizado – com todos os vértices definidores – para o mapa dialetal elaborado por Nascentes (1953).

Figura 8 – Mapa dialetal, a partir do georreferenciamento da Cartografia Automatizada⁸



Fonte: Adaptado de Teles (2018, p. 485).

As pesquisas sobre áreas dialetais brasileiras ganham um novo parâmetro, quando a alusão for feita à divisão proposta em 1953, por Nascentes, uma vez que o mapa apresentado por Teles (2018) se configura como o mais atual, por meio dos critérios da Cartografia

⁸ Além desse mapa, há uma tabela com mais detalhes em Teles (2018).

Automatizada, conferindo, desse modo, mais cientificidade e precisão para os limites dos falares e subfalares.

A divisão de Nascentes (1953) é responsável por fomentar estudos de diversas naturezas, embora abordem os subfalares, dos quais podem ser citados: Cardoso (1986/1999); Ferreira, Mota e Rollemberg (1994); Ferreira e Cardoso (1995); Ferreira (1998); Mota (1999); e Zágari ([1998] 2005), todos com uma investigação voltada para o aspecto fonético. Dentre esses e inúmeros trabalhos nessa perspectiva, destacam-se três, a saber: Cardoso (1986/1999); e Zágari ([1998] 2005), que serão detalhados a seguir.

Em 1986, ao confrontar dados do APFB (ROSSI, 1963) e do EALMG (RIBEIRO et al., 1977), Cardoso (1986) afirma existir uma unidade dialetal para a área do falar baiano, por ter encontrado a presença de vogais abertas pretônicas (na Bahia e em parte de Minas Gerais). Além disso, as delimitações em que o Brasil é dividido em dois grandes grupos, Norte e Sul, são atestadas, segundo o cotejo feito pela autora.

Cardoso (1999), ao analisar dados referentes à realização das vogais médias pretônicas, abertas e fechadas, em 16 estados do Brasil, conclui que, no que tange à divisão proposta por Nascentes (1953), os dados afirmam tal divisão, uma vez que as vogais fechadas foram documentadas nas áreas dos falares do Sul (fluminense, mineiro e sulista) ao passo que, as vogais abertas foram documentadas, de forma majoritária, em estados que pertencem aos falares do Norte (amazônico e nordestino), mas também em partes do falar baiano.

Para a área de Minas Gerais, com base na pesquisa empreendida para elaboração do EALMG – (RIBEIRO et al., 1977), há uma proposição de Zágari ([1998] 2005) que, indo de encontro das ideias propostas por Nascentes (1953), afirma que o estado de Minas Gerais apresenta três grupos de falares, o falar baiano, o falar mineiro e o falar paulista. Dessa maneira, tal proposição não leva em consideração os falares fluminense e sulista da proposta de 1953. Sendo assim, para o estado de Minas Gerais, até então, não há um consenso, no que tange aos limites estabelecidos por Nascentes (1953) e Zágari ([1998] 2005).

Figura 9 – Divisão dialetal de Minas Gerais proposta por Zágari ([1998] 2005).



Fonte: Zágari ([1998] 2005, p. 64).

Ainda, no que tange à área de Minas Gerais, Alves (2022), a partir dos dados do Projeto ALiB, se debruçou a analisar o abaixamento das vogais médias pretônicas em 23 localidades do referido estado. Foram considerados os dados de 92 informantes (4 informantes por localidade), controlados por variáveis sociais, faixa etária e sexo, a partir das respostas fornecidas para os Questionários Fonético-Fonológico e Semântico-Lexical.

Após todo o levantamento de dados e utilização do programa estatístico *Goldvarb X*, com os suportes teóricos da Dialectologia e da Sociolinguística Variacionista, os resultados apontados foram:

Os dados da pesquisa evidenciaram tanto no tempo aparente (ALiB) quanto no tempo real (ALiB x EAMLG) duas áreas dialetais e uma área de transição do ponto de vista das vogais médias pretônicas. A realização das vogais abertas é minoritária, mesmo nas cidades do falar baiano que lideram o abaixamento. Na área setentrional ocorre a abertura das vogais abertas e a na área meridional o predomínio das vogais fechadas. Com exceção das cidades de Viçosa, Muriaé que, inseridas na área meridional, registraram a abertura da vogal, embora em frequência relativamente menor que as cidades inseridas na região Norte, o que consiste em uma zona de transição da vogal (ALVES, 2022, p. 252).

Os estudos sobre a realidade linguística do PB revelam a necessidade de uma nova proposição, porém com base em pesquisa empírica, a fim de confirmar ou contestar tal proposta em sua totalidade, conforme aponta Mota (2006). Ainda, a referida autora, ao refletir sobre a proposta de divisão feita por Nascentes (1953) e as análises contemporâneas, sinaliza para alguns fatos: informações insuficientes e para a dificuldade de intercomparação de dados, à época, recolhidos com metodologias distintas dos atlas regionais/estaduais, haja vista que

Analisados os dados hoje disponíveis, verifica-se a necessidade de um maior conhecimento das áreas dialetais brasileiras, especialmente daquelas que ainda não dispõem de atlas regionais, assim como de uma amostra atualizada, recolhida

simultaneamente, com a mesma metodologia e sob coordenação geral em todo o País, como a que se programa para a elaboração do Atlas Linguístico do Brasil. Somente a partir de amostra desse tipo será possível uma proposta cientificamente justificável de divisão do País em áreas dialetais (MOTA, 2006, p. 351).

Urge que, dessa maneira, estudos feitos com base em dados recolhidos para o atlas nacional retratem a atual configuração dialetal brasileira, como pontuaram Nascentes (1955) e Mota (2006). E, pelo que se tem de notícias, a nível lexical, a equipe de dialetólogos do referido Projeto ALiB vem se aproximando, cada vez mais, do esboço dessa configuração. Há diversos trabalhos, segundo Quadro 3, dos quais se citam os de: Aguilera (2009); Ribeiro (2012); Portilho (2013); Romano (2015); Cardoso (2016); Santos (2016); Santos (2018), Alencar (2018) e Souza (2021).

Quadro 3 – Estudos lexicais com dados do Projeto ALiB sobre áreas dialetais.

Ano	Autor (a)	Gênero textual	Título do trabalho
2009	Vanderci Aguilera	Artigo	Léxico e áreas dialetais: o que podem demonstrar os dados do ALiB.
2012	Silvana Ribeiro	Tese	Brinquedos e brincadeiras infantis na área do “Falar Baiano”.
2013	Danyelle Portilho	Dissertação	O falar amazônico: uma análise da proposta de Nascentes (1953) a partir de dados do Projeto ALiB.
2015	Valter Romano	Tese	Em busca de falares a partir de áreas lexicais no centro-sul do Brasil.
2016	Suzana Cardoso	Artigo	Áreas dialetais do português brasileiro: o que dizem as primeiras cartas do ALiB.
2016	Leandro Santos	Dissertação	Brincando pelos caminhos do Falar Fluminense.
2018	Graziele Santos	Dissertação	O léxico dos jogos e diversões infantis no <i>corpus</i> do Projeto ALiB: visitando o Falar Nordeste.
2018	Beatriz Alencar	Tese	O léxico de brinquedos e brincadeiras infantis em São Paulo
2021	Ana Rita Souza	Dissertação	As estrelas no <i>corpus</i> do Projeto ALiB: revisitando o Falar Amazônico.

Fonte: Elaborado pelo autor (2023).

Aguilera (2009), ao se debruçar sobre a última proposição de Nascentes (1953), verifica tal proposta quanto à validade e atualidade, apontando a inexistência de linhas virtuais, as isoglossas, que pudessem atestá-la, com base nos dados do Projeto ALiB, dando

as primeiras notícias referentes às áreas dialetais brasileiras por meio de dados do léxico. Por isso, a referida autora tece algumas considerações sobre os falares do norte:

(i) A divisão de falares do Brasil em Norte e Sul de Nascentes, ainda que decorridos quase 90 anos depois da proposta de 1923, durante os quais nosso país sofreu profundas mudanças sociais, demográficas, econômicas e políticas, ainda pode ser confirmada, principalmente em relação ao Falar do Norte, com a distribuição de igarapé (Norte) e córrego (Sul), tangerina, galinha picote, bagana e bombom/confeito (Norte). (ii) Algumas variantes têm uma distribuição bem marcada que pode ser relacionada com a subdivisão do falar nortista, como o amazônico, com igarapé, picote e bostela; e o nordestino com laranja cravo, capote, tô fraco, cataraca e confeito, por exemplo. (iii) O subfalar nordestino demonstra ora convergências, ora divergências nos nove estados que compõem a região Nordeste, principalmente se compararmos a distribuição de algumas variantes como: tanja, exclusiva de São Luís; capote, registrada apenas em São Luís, Teresina e Fortaleza; tô fraco produtiva em João Pessoa, Maceió e Aracaju; cochia, em Fortaleza; piruba, em Natal, piola, em João Pessoa e góia em Recife e Maceió (AGUILERA, 2009, p. 4231- 4232).

E sobre os falares do Sul:

O falar do Sul, que Nascentes dividiu em subfalares, porém, além de demonstrar menor uniformidade ainda não traz, pelo menos nas variantes aqui analisadas, a divisão nos subfalares fluminense, mineiro e sulista. A denominação deste último eu substituiria por subfalar paulista que considero mais apropriada para designar a variante de grande parte dos territórios dos Estados de São Paulo, Paraná, Mato Grosso, Mato Grosso do Sul, e partes menores de Goiás e Santa Catarina. Nesta minha proposta, reservaria a denominação sulista para a variante do Rio Grande do Sul, que, pelo menos no nível lexical que aqui analisamos, se constitui diferentemente dos demais Estados. Justifico a denominação de subfalar paulista considerando a influência dos paulistas que adentraram o Brasil a partir do século XVII, intensificada nos séculos posteriores em direção ao sul de Minas Gerais, Goiás, Mato Grosso (inclusive o atual Mato Grosso do Sul), Paraná e parte de Santa Catarina (AGUILERA, 2009, p. 4232).

A tese de doutoramento de Ribeiro (2012), seguindo a linha de investigação dialetal com dados lexicais, englobou 57 localidades pertencentes à rede de pontos do Projeto ALiB, com foco na área dialetal denominada como *Falar Baiano*, pertencente aos Falares do Sul. Tal área corresponde a 11 estados (os estados completos de Sergipe e Bahia) e (parte dos estados de Goiás, Tocantins, Minas Gerais, Maranhão, Piauí, Pernambuco, Alagoas, Mato Grosso e Espírito Santo), distribuídos em quatro regiões administrativas brasileiras. Foram analisadas as respostas de 244 informantes para as 13 perguntas da seção temática Jogos e Diversões Infantis do Questionário Semântico-Lexical (QSL) do Projeto ALiB, a saber:

155 – CAMBALHOTA: “... a brincadeira em que se gira o corpo sobre a cabeça e acaba sentado? (*Mímica*)”; 156 – BOLINHA DE GUDE: “... as coisinhas redondas de vidro que os meninos gostam de brincar”; 157 – ESTILINGUE/SETRA/BODOQUE: “... o brinquedo feito de uma forquilha e duas tiras de borracha (*mímica*), que os meninos usam para matar passarinho?”; 158 – PAPAGAIO DE PAPEL/PIPA: “... o brinquedo feito de varetas cobertas de papel que se empina no vento por meio de uma linha?”; 159 – PIPA/ARRAIA: “... um brinquedo parecido com o ____ (*cf. item 158*) também feito de papel, mas sem varetas, que se empina ao vento por meio de uma linha?”; 160 –

ESCONDEESCONDE: "... a brincadeira em que uma criança fecha os olhos, enquanto as outras correm para um lugar onde não são vistas e depois essa criança que fechou os olhos vai procurar as outras?"; 161 – CABRA-CEGA: "... a brincadeira em que uma criança, com os olhos vendados, tenta pegar as outras?"; 162 – PEGA-PEGA: "... uma brincadeira em que uma criança corre atrás das outras para tocar numa delas, antes que alcance um ponto combinado?"; 163 – FERROLHO / SALVA / PICULA / PIQUE: "... esse ponto combinado?"; 164 – CHICOTE-QUEIMADO / LENÇO ATRÁS: "... uma brincadeira em que as crianças ficam em círculo, enquanto uma outra vai passando com uma pedrinha, uma varinha, um lenço que deixa cair atrás de uma delas e esta pega a pedrinha, a varinha, o lenço e sai correndo para alcançar aquela que deixou cair?"; 165 – GANGORRA: "... uma tábua apoiada no meio, em cujas pontas sentam duas crianças e quando uma sobe, a outra desce?"; 166 – BALANÇO: "... uma tábua, pendurada por meio de cordas, onde uma criança se senta e se move para frente e para trás?"; 167 – AMARELINHA: "... a brincadeira em que as crianças riscam uma figura no chão, formada por quadrados numerados, jogam uma pedrinha (*mímica*) e vão pulando com uma perna só?" (COMITÊ NACIONAL DO PROJETO ALiB, 2001, p. 34-35).

Em nível de pós-graduação, esta pesquisa se tornou uma referência para os estudos posteriores, uma vez que a autora atestou ser possível traçar áreas dialetais por meio desse nível de análise da língua; criou área de controle, estabelecendo um limite a ser observado, para além dos pontos inclusos dentro do falar estudado. O trabalho está subdividido em três volumes e, no volume 2, encontram-se 40 cartas, dentre elas, 23 cartas são semântico-lexicais, nove cartas são resumo, sete cartas são introdutórias e uma carta é fonética. Além disso, no que se refere ao traçado do dialetólogo, especificamente o *Falar Baiano*, assevera a autora:

O produto cartográfico apresentado e a identificação do *Falar Baiano* e das subáreas demarcadas demonstram que o léxico pode revelar áreas dialetais. A proposta de Nascentes (1953) tem vitalidade na realidade presente, mas o limite traçado pelo autor, ainda não pode ser alargado ou reduzido sem que antes seja conhecida, com maior profundidade, a área circunvizinha ao *Falar Baiano*. As subáreas dialetais A, B, C e D apresentadas evidenciam a *diversidade* na *unidade* (RIBEIRO, 2012, p. 449) [grifos da autora].

Portilho (2013), em sua pesquisa, contemplou 26 localidades da rede de pontos do Projeto ALiB, também incluindo uma área de controle a ser observada, observando aos limites do Falar Amazônico, pertencente aos Falares do Norte. Dentre essas localidades, 20 estão distribuídas entre pontos situados no Acre, Amazonas, Pará, Roraima e Amapá e área de controle, seis pontos, localizados em Mato Grosso, de Rondônia, do Maranhão e do Tocantins. Na pesquisa, a autora analisa os dados por duas perspectivas: a diatópica e a léxico-semântica, sob os princípios teóricos da Dialetologia e Lexicologia.

Foram coletados os dados de fala de 128 informantes em relação às 13 perguntas também sobre a área temática dos Jogos e Diversões Infantis do QSL, o que resultou na

documentação de 196 itens lexicais, que estão distribuídos em 12 cartas linguísticas. Ao concluir a dissertação, sobre o *Falar Amazônico*, a autora afirma que:

Foi possível atestar parcialmente a vitalidade da área dialetal do falar amazônico [...] Pelo exposto, pode-se afirmar que, apesar das interinfluências entre os falares, especialmente entre as localidades fronteiriças e a área dialetal investigada nesta pesquisa, foi atestada uma relativa vitalidade do falar amazônico no nível lexical, considerando que o léxico dessa área mostrou-se peculiar em relação ao de outras regiões do Brasil. [...] Estudos mais amplos envolvendo as demais áreas dialetais propostas por Nascentes (1953) poderão ratificar ou retificar as conclusões obtidas a partir dos dados lexicais examinados (PORTILHO, 2013, p. 137-139).

Romano (2015), em tese de doutoramento, contemplou o *Falar Sulista* a partir dos dados de 118 pontos de inquérito, totalizando 472 informantes, que estão localizados em 118 municípios, de oito estados brasileiros, Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Mato Grosso do Sul, parte do estado do Mato Grosso, centro-sul de Goiás, parte do sul de Minas Gerais e do Triângulo mineiro. Para o empreendimento, o estudo considerou mais de um campo lexical do QSL, diferentemente dos estudos anteriormente mencionados, em nível de pós-graduação:

001 – CÓRREGO: “O rio pequeno de dois metros de largura”; [...] 039 – TANGERINA: “...as frutas menores que a laranja, que se descascam com a mão, e, normalmente, deixam um cheiro nas mão?”; [...] 132 – MENINO: “Criança pequena de 5 a 10 anos, do sexo masculino”; [...] 156 – BOLINHA DE GUDE: “... as coisinhas redondas de vidro que os meninos gostam de brincar”; [...] 177 – GELEIA: “pasta feita de frutas para passar no pão, biscoito” (COMITÊ NACIONAL DO PROJETO ALiB, 2001, p. 21-36).

O referido autor, por meio da ferramenta computacional *Software para Geração e Visualização de Cartas Linguísticas* [SGVCLin]⁹, apresenta um conjunto de 71 cartas lexicais, relatórios de produtividade dos itens analisados e, além disso, para a área ora analisada, afirma que:

Considerando que o PB, em meados da segunda década do século XXI, já se encontra documentado em áudio pela equipe do *Projeto ALiB*, pode-se afirmar que, sob o ponto de vista do léxico, a divisão de Nascentes (1953), para o subfalar sulista, não é mais válida. Seria um equívoco considerar estados linguística e sóciohistoricamente distintos como SP e RS pertencentes ao mesmo falar, uma vez que ambos os estados apresentam aspectos que os definem e os diferenciam [...] O território investigado pode ser dividido em duas grandes áreas: (i) a meridional (*falar sulista*) e (ii) a setentrional (*falar paulista*), caracterizadas anteriormente e que apresentam limites virtuais e fluidos (ROMANO, 2015, p. 265).

Em 2016, Cardoso descreve e analisa 10 cartas lexicais do atlas nacional, que fazem parte do volume 2 (CARDOSO et al., 2014b): L01- Granizo; L05-Tangerina, L06-Penca da Banana; L07- Extremidade da Inflorescência da bananeira; L11- Galinha D’Angola;

⁹ Mais informações podem ser obtidas no site <http://sgvclin.altervista.org/>.

L13Bicho da Goiaba; L14- Pernilongo; L16- Cigarro de Palha; L17- Cambalhota; e L18- Bolinha de Gude, trazendo esboços areais, com base nos dados das primeiras cartas do ALiB, dados obtidos nas 25 capitais brasileiras, consoante critérios metodológicos do Projeto ALiB.

Cardoso (2016), após uma análise contrastiva, afirma a existência de cinco subáreas dialetais, assim denominadas: área A, constituída por capitais da região Norte, o que se atesta com as presenças das formas *carapanã* (L14) e *capote* (L11); área B, formada também por capitais da região Norte com o acréscimo de duas capitais da região Nordeste, São Luís e Teresina, identificada pelo registro das formas *carambola/carambela*, *peteca* e *porronca*; área C, que apresenta unidade dialetal entre Norte e Nordeste em detrimento aos demais espaços geográficos em direção ao Sul, comprovada pela coexistência das formas *chuva de neve*, *mangará*, *palma* e *tapuru*; área D, que engloba as capitais da região Centro-oeste e São Paulo, identificada pela forma *maricote*; e a área E, caracterizada pela região Sul, por meio da presença de *bergamota*. Por fim, ainda, conforme Cardoso (2016):

O que se pode dizer como palavra de finalização? Avançar na análise dos dados do Projeto ALiB, examinando o que se delinea na voz e na expressão dos falantes – 900 entre homens e mulheres de diferentes idades – das 225 cidades que integram a rede de pontos com localidades distribuídas pelo interior de 26 estados brasileiros, para que se junte aos já explorados 200 informantes das 25 capitais. Esses dados não de permitir, embora reconhecendo as dificuldades que o estabelecimento de áreas dialetais comporta, o traçado de isoglossas que vão delinear, com a segurança da pesquisa in loco, as atuais áreas dialetais do português brasileiro. O caminho está aí. Necessário se faz continuar a percorrê-lo para atingir a meta alcançada no que diz respeito à divisão dialetal do Brasil (CARDOSO, 2016, p. 46-47).

Santos (2016) analisou, em sua dissertação de mestrado, os limites da área do *Falar Fluminense*. O estudo abarcou os dados de 35 localidades, pontos pertencentes ao falar estudado e pontos de controle, perfazendo um total de 152 informantes, que estão em cinco estados: Bahia, Minas Gerais, Espírito Santo, Rio de Janeiro e São Paulo.

A análise diatópica foi primordial, uma vez que o estudo visava à descrição de uma área dialetal, no entanto, os aspectos sociais também foram passíveis de análises. O *corpus* utilizado contemplou as 13 perguntas da seção temática dos Jogos e Diversões Infantis. E, para concluir a pesquisa, como produtos das análises, Santos (2016) catalogou 2208 itens lexicais, que estão cartografados em um conjunto de 17 cartas linguísticas. Além disso, sugere à cartografia linguística a inclusão da representação do não dado (conforme detalhado na seção 3.4). E, por fim, o referido autor constata que:

É oportuno afirmar a precisão de Nascentes (1953), ao dividir as terras brasileiras em dois grandes grupos, fato que se comprova ao serem cotejados os dados resultantes de pesquisas com os dados do Projeto ALiB, (Ribeiro, 2012; Portilho, 2013; Romano, 2015), bem como os dados desta dissertação, sob o ponto de vista

lexical. Logo, ratifica-se que, por meio deste nível de análise, é possível identificar e caracterizar áreas linguísticas. No entanto, pelo que se observa, no que tange às subdivisões dos falares do Sul, em especial à área do *Falar Fluminense*, tal proposição não pode ser considerada como válida, pois ora os dados evidenciam uma área linguística comum ora negam tal fato, não podendo, de fato, atestar uma unidade dialetal. A partir desse cenário, afirma-se que não foi possível estabelecer subáreas linguísticas nem traçar isoglossas, mas destacam-se o norte de Minas Gerais e o norte do Espírito Santo, pois eles se aproximam (SANTOS, 2016, p. 189-190).

Em 2018, analisando a área que corresponde ao *Falar Nordestino*, Santos selecionou 53 localidades, distribuídas assim na da região Nordeste - Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco e Alagoas. No total, foram elocuições dos 240 informantes. Assim como os estudos sobre os falares *Baiano*; *Amazônico* e *Fluminense*, também foram utilizadas as 13 perguntas dos jogos e diversões infantis.

No bojo das análises empreendidas sobre os itens lexicais documentados na área em estudo, Santos (2018) apresenta 17 cartas linguísticas e tece algumas considerações propositivas:

1) É possível delinear a área do dialetal do *Falar Nordestino* através do léxico? 2) Com base em estudo centrado no campo semântico de *jogos e diversões Infantis*, pode-se considerar atual a proposta de divisão dialetal de Nascentes (1953), em relação à área denominada *Falar Nordestino*? [...] Observando a primeira questão, os dados coletados mostraram que foi possível, por meio de itens lexicais, delinear diatopicamente a área do *Falar Nordestino* de Nascentes (1953). Verificando toda a área pesquisada responde-se negativamente à segunda questão, pela ausência de atualidade na proposta de Nascentes (1953), no que se refere ao *Falar Nordestino* como uma área homogênea. A falta de unidade linguística pode ser explicada pela extensão e pela diversidade sócio-histórica da área [...] A heterogeneidade linguística demonstrada na área examinada comprova a hipótese admitida na Introdução: a proposta de divisão dialetal de Nascentes, datada de 1953, para a área do *Falar Nordestino*, não corresponde, atualmente, à realidade linguística observada, considerando o campo semântico dos *jogos e diversões Infantis* nos dados do Projeto ALiB e uma descrição baseada no léxico (SANTOS, 2018, p. 200-201).

Em 2018, analisando a área que corresponde ao *Falar Sulista*, restringindo-se ao estado de São Paulo, Alencar (2018) selecionou 37 localidades, pertencentes ao referido estado, mas também 10 localidades, pertencentes aos estados vizinhos, um ponto no Mato Grosso do Sul, dois pontos no Rio de Janeiro, três pontos no Paraná e quatro pontos em Minas Gerais. No total, foram elocuições dos 180 informantes. Também foram utilizadas as 13 perguntas dos jogos e diversões infantis.

Após as análises, Alencar (2018) apresenta 26 cartas linguísticas, afirma que a área em estudo apresentou bastante vitalidade, quanto aos nomes dados para os brinquedos e brincadeiras infantis e teceu algumas considerações, tais como:

a) Não é possível, por ora, atestar a vitalidade da área dialetal do falar sulista considerada por Nascentes (1953) no estado de São Paulo, corroborando-se a hipótese da existência de um falar paulista com características próprias que o distingue do falar sulista na concepção do eminente dialetólogo brasileiro [...] b) Os dados analisados possibilitaram o delinear de três subáreas dialetais na área investigada, marcadas, sobretudo, pelas influências entre a cultura e a história social de São Paulo. Essas subáreas estão situadas: i) na proximidade do Vale do Ribeira associada aos diferentes trajetos percorridos pelas tropas oriundas do sul do Brasil; ii) na região Oeste e Noroeste do estado de São Paulo, fenômeno associado à cultura do café implantada no século XX e que se estendeu até a região norte do Paraná em virtude de terras propícias para essa cultura; iii) expansão do falar paulista para além dos limites estaduais MG-SP atingindo o sul de Minas Gerais e o Triângulo Mineiro, como já vislumbrado por Zágari (1998), entendido aqui como um prolongamento do falar paulista e não simplesmente como um dos três falares mineiros... (ALENCAR, 2018, p. 517-518).

Souza (2021), seguindo a trilha das pesquisas sobre o léxico, a partir dos dados do Projeto ALiB, com vista à apuração de áreas dialetais, analisou, em sua pesquisa de mestrado, a área do *Falar Amazônico*. Para o empreendimento, a referida autora observou dados de 30 localidades, sendo 20 da região Norte e, também, área de controle, que abrange 10 localidades inclusas nas regiões Nordeste, Centro-Oeste e na própria região Norte. Ademais, foram utilizadas três questões do QSL, referentes às estrelas, da seção temática *Astros e Tempo*, a saber:

029 – ESTRELA MATUTINA/ VÊNUS/ ESTRELA DA MANHÃ/ ESTRELA-D’ALVA: “De manhã cedo, uma estrela brilha mais e é a última a desaparecer. Como chamam esta estrela?”; 030 – ESTRELA VESPERTINA/ VÊNUS/ ESTRELA DA TARDE: “De tardezinha, uma estrela aparece antes das outras, perto do horizonte, e brilha mais. Como chamam esta estrela?”; 031 – ESTRELA CADENTE/ ESTRELA FILANTE/ METEDORO/ ZELAÇÃO: “De noite, muitas vezes pode-se observar uma estrela que se desloca no céu, assim, (mímica) e faz um risco de luz. Como chamam isso?” (COMITÊ NACIONAL DO PROJETO ALiB, 2001, p. 23).

Ao findar o estudo, a referida autora apresenta um conjunto de seis cartas linguísticas, a partir do produto cartográfico atual, disponibilizado por Teles (2018), e tece algumas considerações, dentre as quais se destacam:

A cartografia foi produzida com base na distribuição das respostas mais produtivas, em seguida, as menos produtivas em conjunto com as não respostas. O traçado das isoléxicas não foi feito porque a distribuição das respostas indica que o traçado de Nascentes (1953), referente ao Falar Amazônico, não pôde ser delimitado, a partir dos nomes atribuídos às estrelas e, também, porque não registramos a presença de nenhuma subárea dialetal dentro dos limites estabelecidos (SOUZA, 2021, p. 203).

Nesse sentido, ao observar o transcorrer dos anos, de 2009 a 2021, percebe-se que os dialetólogos brasileiros, mais especificamente, os pesquisadores do Projeto ALiB, vêm, há anos, buscando descortinar a real faceta areal brasileira, desejo que persegue gerações de estudiosos da área. Dos trabalhos lexicais supramencionados, sugere-se a reunião desses

dados, por meio de critérios de tratamento únicos e bem delimitados, a fim de que se possam ter, sob o prisma lexical, notícias hodiernas da configuração desses falares. Para tal, nessa perspectiva:

[...] o trabalho em conjunto torna-se de fundamental importância para o sucesso da empreitada brasileira em alcançar o atual mapa dialetológico [...] Logo, então, nas trilhas de um novo mapa dialetal, seguem os dialetólogos brasileiros. Assim que o país dispor dos dados do ALiB publicados, tanto das capitais quanto do interior, algumas análises poderão ser feitas, com o fito de delimitação de áreas dialetais, traçando, assim, isoglossas que permitam verificar unidade e/ou diversidade, nos mais diferentes níveis da língua (SANTOS, 2020, p. 60).

Tal labor é árduo, mas, aos poucos, vem se desenhando a proximidade de esboçar um novo mapa dialetal, com georrefenciamentos atuais, por meio da Cartografia Automatizada, a partir de um material linguístico recolhido *in tottum e in loco*, analisados por dialetólogos especialistas no assunto, sobre a realidade da língua falada do Brasil, no tocante à língua portuguesa.

Outro aspecto notável é que os estudos sobre os pronomes e os rearranjos na estrutura pronominal do Brasil, tais como os pronomes: *eu x mim* em posição de sujeito em orações infinitivas, precedidos pela preposição *para*; *tu x você*; *nós x a gente*; dentre outros, podem auxiliar na descrição dos caminhos para o entendimento da realidade dialetal brasileira. Espera-se que os dados aqui examinados possam contribuir para a agenda de estudos sobre as configurações atuais do PB, em virtude da carência de dados morfossintáticos sendo utilizados para essa tarefa, haja vista que os estudos fonéticos e os estudos lexicais são mais proeminentes nesse tipo de empreendimento.

2.2 TRILHA DA VARIAÇÃO SOCIAL: A SOCIOLINGUÍSTICA

A Sociolinguística, concebida enquanto um dos ramos da Linguística, surge na década de 60, sendo esta uma vertente dos estudos linguísticos que possui uma natureza interdisciplinar, tem por meta a observação do sistema linguístico a partir do componente social. Também conhecida como a Teoria da Variação e/ou Sociolinguística Variacionista, ganha força por meio dos estudos feitos por William Labov, linguista norte-americano, sobre a mudança linguística na ilha de Martha's Vineyard (1962), ao observar os ditongos, e a variação social do inglês, em 1966, na comunidade de fala de Nova York, observando o sistema sonoro do inglês.

Os postulados labovianos revolucionaram os estudos linguísticos da época, pois a variação e a mudança linguísticas são compreendidas como frutos de alternância de uso de

duas ou mais formas linguísticas ao longo do tempo. A premissa desse ramo da Linguística é de que a “[...] percepção de que a base do conhecimento intersubjetivo na linguística tem de ser encontrada na fala – a língua tal como usada na vida diária por membros da ordem social...” (LABOV, 2008 [1972], p. 13), dentro de uma comunidade de fala, que pode ser definida

[...] pela participação num conjunto de normas compartilhadas; estas normas podem ser observadas em tipos de comportamentos avaliativos explícitos e pela uniformidade de padrões abstratos de variação que são invariantes no tocante a níveis particulares de uso. De igual modo, por meio de observações do comportamento linguístico, é possível fazer estudos detalhados da estrutura de estratificação de classe numa dada comunidade (LABOV, 2008 [1972], p. 150).

De modo a compreender os pressupostos do modelo teórico da Sociolinguística, observa-se uma nova abordagem sobre a língua, bem como a tarefa de descrever e analisar a fala, antes vista como caótica e não sistematizável, conforme a teoria estruturalista. Desse modo, a Sociolinguística observa os padrões linguísticos em uma comunidade, tendo a variação e a mudança linguística como frutos oriundos da condição natural da própria língua, e podem ser passíveis de descrição e sistematização, haja vista que

[a]s gramáticas em que ocorre a mudança linguística são gramáticas da comunidade de fala. Como estruturas variáveis contidas na língua são determinadas por funções sociais, os dialetos não oferecem a base para as gramáticas autônomas ou internamente consistentes (WEINREICH, LABOV; HERZOG, 2006 [1968], p. 126).

Para a presente pesquisa, visa-se a observar se as formas pronominais *eu* e *mim* estão em processo de variação ou mudança, quando em contextos de oração reduzida de infinitivo, precedidas pela preposição *para* (e sua forma contraída *pra*), levando em consideração fatores da estrutura linguística e fatores sociais, pois

[...] não se pode entender o desenvolvimento de uma mudança linguística sem levar em conta a vida social da comunidade em que ela ocorre. Ou, dizendo de outro modo, as pressões sociais estão operando continuamente sobre a língua, não de algum ponto remoto do passado, mas como uma força social imanente, agindo no presente vivo (LABOV, 2008 [1972], p. 21).

O modelo teórico da Sociolinguística Variacionista prioriza o estudo do vernáculo, que, para os sociolinguistas, são “[...] momentos em que o mínimo de atenção é prestado à língua, ao *como* da enunciação. Essas partes do discurso falado, caracterizadas aqui como o vernáculo, constituem o material básico para a análise sociolingüística” (TARALLO, 2001, p. 19), ou seja, “[...] a língua tal como usada na vida diárias por membros da ordem social, este veículo de comunicação com que as pessoas discutem com seus cônjuges, brincam com seus amigos e ludibriam seus inimigos (LABOV, 2008 [1972], p. 13). Além disso, defende

que o sistema linguístico é heterogêneo, ou seja, as formas que variam e podem ocasionar mudança linguística não são aleatórias, mas podem ser descritas e analisadas.

Ao tomar por parâmetro a observação do vernáculo, os estudos sociolinguísticos permitem a obtenção de dados para a categorização de regras variáveis em uma comunidade de fala, essa, por sua vez, pode ser definida como uma comunidade de um grupo de indivíduos que partilham normas, uma vez que elas podem ser observadas em diferentes tipos de avaliação (LABOV, 2008 [1972]). No que tange à escolha da amostra dos informantes, esta deve priorizar um quantitativo de pessoas que representem o perfil social das comunidades de fala. Também, essa eleição dos informantes deve ser pautada em uma estratificação, com base nos aspectos escolhidos segundo os intentos do pesquisador. Contemplando informantes tanto das áreas rurais quanto das áreas urbanas, quando for possível, conforme critérios pré-estabelecidos pela pesquisa.

2.2.1 Variação e mudança linguísticas: conceitos fundamentais

Para o modelo da Teoria da Variação e da Mudança, os processos que dizem respeito à variação e à mudança linguística devem, de certo modo, ser associados ao seu encaixamento na estrutura social, uma vez que a significação social dada pelos utentes da língua, ao escolherem determinada forma em detrimento de outra, também se constitui como alvo de investigação para os pesquisadores dessa vertente teórica.

Ao considerar os fatores da estrutura social, sexo, faixa etária, classe social e escolaridade, o pesquisador sociolinguista vai correlacioná-los aos fatores linguísticos, a fim de poder delinear o perfil daquela comunidade de fala, por meio de regras variáveis. Assumem-se, desse modo, nessas pesquisas, as noções de variantes linguísticas, que, segundo Tarallo (2001), são duas formas linguísticas que, em um mesmo contexto, funcionam como alternantes para representar o mesmo significado com mesmo valor de verdade.

Por exemplo, as formas *eu* e *mim* são consideradas como variantes no presente estudo. E o conjunto de duas ou mais variantes recebe o nome de variável, por exemplo, a alternância entre *eu* e *mim* para ocupar o lugar de sujeito na primeira pessoa do singular. Além do nível morfosintático, as variantes linguísticas podem ser encontradas em outros níveis de análise da língua. Para além das discussões sobre o mesmo valor de verdade das variantes, acredita-se que elas possuem valores e referenciais idênticos, mas divergem quanto ao sentido social e estilístico (LABOV, 2008 [1972]).

Assim, alguns tipos de variação são observados pelas análises do método sociolinguístico, a variação diafásica, a variação diatópica, a variação diastrática e a variação diamésica.

A variação diafásica, conhecida também por variação estilística, diz respeito ao controle de variáveis que evidenciem as diferentes formas de falar a partir do contexto comunicativo. Ao considerar esse pressuposto, evidencia-se que os comportamentos linguísticos das pessoas são alterados mediante aos papéis sociais que essas desenvolvem, com quem interagem e em qual ambiente elas estão. O monitoramento linguístico, desse modo, pode e deve ser alvo de investigações, pois ele, de certa maneira, determina escolhas linguísticas que, certamente, perpassam pela dicotomia das formas: prestigiadas x desprestigiadas.

Outro tipo de variação, observado tanto pela Sociolinguística quanto pela Dialetoлогия, é a espacial, geográfica ou areal. Nesse eixo, o local natalício dos indivíduos é alvo de análise, isto é, acredita-se que origem geográfica das pessoas é denunciadora de aspectos de normas de fala. Como a língua deve ser estudada associando-a aos aspectos sócio-históricos, as diferenças regionais e espaciais são frutos das diferentes maneiras de povoamento, colonização e formação dos povos. Além disso, os hábitos culturais, as tradições também podem moldar como a língua é falada em determinado local, ditando formas fonéticas, lexicais, morfossintáticas e prosódicas.

A variação diastrática, ou variação social, é aquela responsável por analisar como alguns fatores sociais podem delinear a forma de fala de pessoas e grupos. Geralmente, os fatores elencados para esse tipo de variação são: o nível socioeconômico, a profissão, a faixa etária, o grau de escolaridade e o sexo. As três últimas são as mais utilizadas nas diversas pesquisas, por isso merecem destaque, a saber:

- i. A faixa etária, ou variação diageracional, evidencia aspectos linguísticos e até sociais, uma vez que a fala, quando observada por esse ângulo, por meio da idade, revela surgimento de formas, manutenção delas ou desuso de outras.
- ii. Em relação ao nível de escolaridade, em geral, nota-se que a quantidade de anos em que o falante passa exposto a uma cultura de educação formal influencia bastante nas suas escolhas linguísticas, fazendo grande diferença quando comparado com outro falante que não teve ou teve pouco contato com a cultura de educação formal, a escola.
- iii. Por muito tempo, sobretudo no alvorecer dos estudos sociolinguísticos, advogou-se que as mulheres se comportavam linguisticamente de forma mais

conservadora. A variação diagenérica, desse modo, é o tipo de variação que busca analisar aspectos referentes às falas dos homens e das mulheres, oferecendo índices e resultados que, por vezes, desmistificam alguns pensamentos fundamentados pelo preconceito social.

As diferenças entre a fala e a escrita são avaliadas pela variação diamésica. Historicamente, difundiu-se que são modalidades diferentes de usar a língua, e que, por isso, constituem-se em dois polos distintos. No entanto, estudos demonstram a existência de um *continuum* tipológico e que existem pontos em comum entre elas (MARCUSCHI, 2008). Sobretudo, vale dizer que a fala interfere na produção escrita e, muitas vezes, a escrita interfere na fala.

No que tange à análise de regras variáveis, a Sociolinguística compreende que todo o processo de variação se caracteriza como uma regra variável, ou seja, há formas para sistematizar a variação e podem ser utilizados métodos quantitativos para tal mensuração. Através de tais métodos, as verificações podem ser empreendidas e os valores obtidos podem revelar diversas relações entre os fatores da estrutura linguística e os fatores da estrutura social.

As variáveis dependentes correspondem às formas da língua em variação, que, por sua vez, são consideradas formas dependentes, haja vista que a utilização delas está condicionada por grupos de fatores ligados à estrutura, à estilística e ao social, que são as variáveis independentes.

Nesse contexto, nas pesquisas sociolinguísticas brasileiras, os cálculos são feitos por pacotes de programas estatísticos, tais como: o VARBRUL, *Goldvarb* e o Rbrul. Vale destacar os ganhos das pesquisas linguísticas com os avanços tecnológicos, bem como a utilização de programas dessa natureza, visto que conferem aos trabalhos mais confiabilidade e precisão das afirmações sobre o estado da língua em determinada comunidade.

Vale mencionar que, de modo a favorecer a compreensão para os estudos em Sociolinguística Variacionista, são elencados fatores de variáveis independentes, de natureza intralinguística, e fatores de variáveis independentes, de natureza extralinguística. Os primeiros são entendidos como fatores de natureza estrutural, que dizem respeito ao contexto linguístico em que a variável dependente ocorre, sobretudo em situações em que há pressões na escolha entre uma variante de prestígio e outra estigmatizada. Por sua vez, as variáveis independentes definem-se como os fatores sociais que podem interferir na escolha da forma linguística pelos falantes.

Ao escolher os dados das capitais do Brasil, o estudo aqui proposto, no que tange aos fatores sociais, observará como as variáveis sociais, idade, sexo e escolaridade podem moldar a escolha pronominal para ocupar a posição de sujeito em estruturas como *para/pra* + *eu/mim* + *infinitivo*, feita pelos informantes desses grandes espaços urbanos, correlacionando-os aos condicionamentos da estrutura linguística. Aludindo a que o pronome *eu*, canonicamente, é considerado como a variante que goza de prestígio, ao passo que o pronome *mim*, considerado como “erro”, quando utilizado na posição de sujeito, é a variante estigmatizada. Nesse sentido,

Entre as variedades faladas num território, uma delas, por diversas razões, pode adquirir maior prestígio e impor-se como norma ou língua padrão. Os factores que determinam essa escolha são normalmente sócio-políticos, históricos, comunicativos e até pedagógicos. Nada, de um ponto de vista estritamente linguístico, leva a que uma determinada variedade seja preferida como norma de uma língua. Só factores extralinguísticos influem nessa escolha... (FERREIRA et al., 1996, p. 483).

Com o alvorecer dos estudos sociolinguísticos, a mudança também passa a ser observada no momento em que ainda há a concorrência entre as diferentes variantes, isto é, no momento presente. Por isso, é pertinente trazer as definições sobre mudança linguística em progresso – disputa entre variantes com a vitória de uma delas observada no presente. Tal visualização só poderá ser empreendida no chamado tempo aparente – comparação entre dados obtidos da fala de indivíduos de diferentes faixas etárias, sistematicamente controladas – o qual se contrapõe com o tempo real – ou seja, investigação ao longo de um espaço de tempo determinado – e variação estável – concomitância estável entre variantes no sistema.

Portanto, conforme Weinreich, Labov e Herzog (2006, p. 126): “Nem toda variabilidade e heterogeneidade na estrutura linguística implica mudança, mas toda mudança implica variabilidade e heterogeneidade”. É necessário encarar a mudança e a variação como fenômenos interligados e compreendidos enquanto próprios à dinâmica da língua, haja vista que são fatos que corroboram os pressupostos dos estudos sociolinguísticos. Ainda, a mudança linguística “[...] não deve ser identificada como deriva aleatória procedente da variação inerente na fala...” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 125).

Assim, Weinreich, Labov e Herzog (2006) apresentam cinco problemas que devem ser analisados, quando observada à língua falada, a partir da complexidade que envolve os processos de mudança linguística, conforme Quadro 4:

Quadro 4 – Cinco problemas propostos pela Teoria da Mudança Linguística

Nº	Problemas
1	Problema dos fatores condicionantes;
2	Problema do encaixamento;
3	Problema da avaliação;
4	Problema da transição;
5	Problema da implementação.

Fonte: Weinreich; Labov; Herzog (2006 [1968], p. 121-125).

Para a Teoria da Variação e da Mudança, o problema dos fatores condicionantes atesta que a variação não é livre e nem ocorre de maneira aleatória, mas sofre condicionamentos por alguns fatores, tais como: linguísticos, estilísticos e sociais. Por isso, tal problema determina “o conjunto de mudanças possíveis e condições possíveis para a mudança” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 121). Sendo assim, as pesquisas sociolinguísticas visam a identificar os princípios gerais e/ou universais que podem favorecer ou restringir as mudanças.

Para o problema do encaixamento, o entendimento é consensual, haja vista que a mudança deve ser compreendida como encaixada no sistema linguístico, cabendo ao pesquisador sociolinguista resolver a natureza e a extensão do encaixamento (WEINREICH; LABOV; HERZOG, 2006 [1968]). Para tal, procura-se observar o encaixamento na estrutura linguística e o encaixamento na estrutura social.

O problema da avaliação é visto como a tarefa de constatar, de maneira empírica, através dos testes de reação subjetiva, a avaliação social frente aos fenômenos linguísticos em variação, pois a reação subjetiva que o falante faz de uma determinada forma linguística pode favorecer a implementação da mudança ou retardar o processo. Diante disso, as ocorrências em que os informantes fizeram uma avaliação, frente ao fenômeno analisado nesta tese, serão alvos de análises e interpretações.

O problema da transição pode descrever o percurso que uma variante pode fazer até alcançar o *status* de ser incorporada no sistema linguístico, pois a mudança linguística não ocorre de forma abrupta, mas por meio de estágios, de um *continuum*, a saber: a origem, “[...] uma mudança é uma das inúmeras variações confinadas ao uso de algumas pessoas...” a propagação, “[...] a mudança é adotada por números tão amplos de falantes que passa a contrastar com a forma mais antiga...” e, por fim, o término, “[...] a mudança alcança regularidade pela eliminação de variantes concorrentes” (LABOV, 2008 [1972], p.152).

Por fim, o problema da implementação pode ser esquematizado como aquele que possui características peculiares dos outros problemas elencados, mas possui etapas, tais

como: a) um dos traços em variação se difunde no seio da comunidade; b) uma significação social é atribuída a esse traço, que possui associação aos valores fornecidos pela tal comunidade; c) de forma gradual, a mudança é generalizada a outros elementos do sistema; d) o progresso da mudança pode ser observado na elevação da consciência social com relação a esta mudança; e, e) há a completude no estágio da mudança, quando o traço se torna regular e, conseqüentemente, perde a sua significação social (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 121-125).

Os dois últimos problemas elencados correspondem aos aspectos relacionados entre indicadores, marcadores e estereótipos. Aqueles são considerados como os fenômenos que estão abaixo do nível de consciência dos falantes de uma dada língua, isto é, o fenômeno pode sofrer variação por influência de aspectos sociais, mas não variação estilística, uma vez que os falantes vão utilizar em contextos formais e informais. Esses são considerados como situados mais ou menos abaixo do nível de consciência do falante, mas sofrem além da variação social, a variação estilística, uma vez que o falante difere entre os contextos de usos formais e informais. Estes estão no nível de consciência do falante, correspondendo às variantes que sofrem estigmas sociais.

Entende-se que as mudanças podem surgir em qualquer grupo social, conforme Labov (2008 [1972], p. 152): “As forças sociais exercidas sobre as formas linguísticas são de dois tipos distintos, que podemos designar como pressões vindas de cima e pressões vindas de baixo”.

Para apuração do estado de variação linguística, o pesquisador deve adotar alguns critérios, tais como: a) deve ser frequente na fala espontânea, b) quanto mais integrada ao sistema, maior o interesse sobre ela, c) as investigações prévias devem informar sobre sua possível estratificação social, desse modo, fornecendo interesse às variáveis que refletem maiores estratificações.

Para além das delimitações conceituais e o desenvolvimento de uma metodologia própria, ao associar os aspectos sociais às análises linguísticas, o modelo laboviano superou as ideias saussurianas, que afirmavam que as mudanças só poderiam ser estudadas após sua conclusão.

2.3 O ENTROCAMENTO DAS TRILHAS DA VARIAÇÃO

Ao observar o percurso dos dois ramos responsáveis por sistematizar a variação e a mudança linguísticas na língua falada, Dialetoлогия e a Sociolinguística, alguns fatores devem ser destacados. Inicialmente, há diferenças dos focos de análises, pois a primeira delas ocupa-se, prioritariamente, na identificação das diversas manifestações linguísticas da língua falada levando em consideração o aspecto horizontal/espacial. Ao passo que a segunda se ocupa de correlacionar os fenômenos linguísticos com os aspectos da estrutura social, portanto, o aspecto vertical/social, conforme mencionados nas seções 2.1 e 2.2, respectivamente.

Assim, embora tenham o mesmo objeto de estudo, a Dialetoлогия possui um caráter eminentemente diatópico, tendo como base de sua descrição a localização espacial. Por sua vez, a Sociolinguística possui uma essência social, posto que centra a sua investigação nas possíveis relações entre os fenômenos linguísticos e os fatores sociais, tais como: idade, sexo, escolaridade, profissão, religião, etc. (CARDOSO, 2010, p. 26).

São ramos que possuem algumas semelhanças e diferenças; na verdade, elas são “[...] duas perspectivas de observação e análise da língua que não se opõem, mas sim se encontram e se complementam” (CALLOU, 2010, p. 35). E podem, até certo ponto, serem consideradas como sinônimas, pois se detêm aos aspectos inerentes à língua falada. Nas palavras de Silva-Corvalan (1989):

Por um lado, a preocupação básica da Dialetoлогия é estabelecer as fronteiras geográficas de certos usos linguísticos [...]. E, por outro lado, a preocupação central da Sociolinguística é identificar processos de mudança linguística em curso e estabelecer as fronteiras sociais de certos usos linguísticos... (SILVA-CORVALAN, 1989, p. 9)¹⁰ [tradução nossa].

Sendo assim, a Dialetoлогия, por ser mais antiga, a partir dos anos 60, à luz dos pressupostos teóricos da Sociolinguística, passou a ter outra vertente considerada sinônima. Em princípio, os estudiosos chegaram a cogitar uma crise no ramo mais percussor, bem como do seu método, a Geografia Linguística, mas houve aprimoramento em ambos.

Desde então, historicamente, muitas pesquisas vêm sendo desenvolvidas e estão colaborando para o entendimento e sistematização da variação e da mudança linguísticas, pois: “Estes estudos demonstram que é possível combinar perspectivas distintas de análise e também que é necessário considerar uma gama de fatores linguísticos e extralinguísticos na busca de explicações para a variação e mudança...” (CALLOU, 2010, p. 45). Aqui, nesta

¹⁰ Por un lado, la preocupación básica de la dialectología por establecer las fronteras geográficas de ciertos usos lingüísticos [...] Y por otro lado, La preocupación central de La sociolingüística por identificar procesos de cambio lingüístico en marcha y por establecer faz fronteras sociales de ciertos usos lingüísticos...

tese, pretende-se unir as perspectivas teóricas dessas duas áreas, a fim de poder retratar, da melhor maneira, o fenômeno ora estudado, uma vez que “[...] Todos os dialetos são tanto espaciais como sociais, visto que todos os falantes têm um entorno social igual a uma localização espacial¹¹...” (CHAMBERS; TRUDGIL, 1994, p. 82).

Ainda no que se refere aos aspectos que intersectam essas duas áreas da Linguística, cabe mencionar, de forma sucinta, as metodologias. Para as pesquisas dialetais, conforme Ferreira e Cardoso (1994), são propostos alguns passos, dentre eles: a) a preparação da pesquisa, b) a execução dos inquéritos, c) exegese e análise dos materiais coletados, e, por fim, d) a divulgação dos resultados. Cardoso (2010) afirma que existe um tripé básico que sustenta as pesquisas dialetais, representado pelo questionário, o informante e a rede de pontos. Tal tripé carece de revisão e adequação, conforme será mencionado na seção 5.2.


Nas pesquisas sociolinguísticas, por seu turno, o pesquisador deve a) selecionar os informantes e a área geografia que pretende investigar; b) criar uma metodologia da coleta dos dados; c) o envelope da variação; d) levantamento de questões e hipóteses; e) codificação dos dados obtidos e análise estatística.

Basicamente, há aspectos convergentes, ao observar as metodologias adotadas, mas também há pontos que divergem, sobretudo no tratamento dos dados, sendo de prioridade dos estudos sociolinguísticos a busca exaustiva pelo fenômeno e a utilização de pacote de programas de análise estatística. Por seu turno, para os estudos dialetológicos, não se faz necessária a busca exaustiva pelo dado. Tais aspectos serão discutidos na subseção 2.4.

Dialetologia e Sociolinguística se assemelham, também, por fornecerem elementos para melhoria e modificação de livros e materiais didáticos, uma vez que a escola, muitas vezes, ainda é considerada uma das fomentadoras dos preconceitos linguísticos existentes na sociedade. Por exemplo, para as formas pronominais analisadas nesta tese, há diversos vieses de explicação, alguns de cunho gramatical/tradicional e, por vezes, preconceituosos sobre o fenômeno. Sobretudo atribuem aos indígenas a fala *para/pra mim fazer*, como pode ser visto na Figura 10.

¹¹ “[...] Todos los dialectos son tanto espaciales como sociales, puesto que todos los hablantes tienen un entorno social igual que una localización espacial...”

Figura 10 – Representação da forma “para eu/mim fazer” na mídia digital

Para mim / Para eu	
Veja se tem algum erro para mim corrigir.	Veja se tem algum erro para eu corrigir.
 <p>Explicação: Eu é um pronome pessoal reto, devendo ser utilizado quando assume a função de sujeito. Assim, para eu deve ser usado sempre que se referir ao sujeito da frase e for seguido de um verbo no infinitivo que indique uma ação.</p> <p>Mim é um pronome pessoal oblíquo tônico, sendo utilizado quando assume a função de objeto indireto, devendo estar sempre precedido por uma preposição. Dessa forma, para mim deve ser usado quando for complemento de um verbo transitivo indireto.</p> <p>Exemplo: Você trouxe a roupa para mim? / Pensei que esse embrulho tivesse chegado para mim.</p>	

Fonte: Só Escola, 2017 ¹².

Em alguns ambientes sociais, principalmente na escola, além da forma *para/prá + eu/mim + infinitivo*, ainda, nota-se que:

Outros estereótipos referem-se ao ‘bem falar’. Ouvimos dizer em todos os países que há um lugar onde a língua nacional é pura (diz-se, no caso da França, que seria a província de Anjou; no caso do Brasil, que seria São Luís do Maranhão), que existem sotaques desagradáveis e outros harmoniosos etc. Por trás desses estereótipos se perfila a noção de bon usage (‘uso certo’), a ideia segundo a qual há modos de bem falar a língua e outros que, em comparação, são condenáveis. Encontramos assim em todos os falantes uma espécie de norma espontânea que os leva a decidir que forma deve ser proscrita, que outra deve ser admirada: não se fala assim, se fala assado etc (CALVET, 2002, p. 60).

Com a divulgação dos trabalhos dialetológico e sociolinguístico, alguns pensamentos do senso comum são desfeitos, pois as pesquisas sobre a língua oral têm evidenciado que muitos desses mitos amparam-se em questões socioeconômicas. Nesse sentido, vale ressaltar que

O tema é rico para aprofundarmos nossa busca de alternativas pedagógicas que permitam pôr a escola na vanguarda, sensibilizando as crianças e os jovens para a variação e para seus sentidos sociais e culturais; contribuindo para uma reconstrução do nosso imaginário nacional sobre a realidade linguística e, acima de tudo, combatendo a violência simbólica que ainda atravessa nossas relações sociais (FARACO, 2008, p. 182).

Votre (2007, p. 53) comenta que “a escola controla, evita e pune, com veemência o uso de formas com supressão e/ou troca de líquidas, como *framengo* e *pobrema* e os fenômenos sintáticos com resquícios dos casos latinos nos pronomes, como *dá para mim sair*”. Se as formas linguísticas diversas são aceitas pelo ambiente escolar, não taxadas como

¹² Disponível em: <<https://www.soescola.com/2017/10/para-mim-fazer-ou-para-eu-fazer.html>>. Acesso em: 15 abr. 2019.

“feias, pobres e erradas”, os alunos poderão se sentir pertencentes ao ambiente escolar, pois entenderão que a variante linguística da sua comunidade é mais uma forma de falar, evitando a evasão. Assim,

Nessa discussão, é importante destacar que o reconhecimento de variedades do PB não é uma motivação para se pensar em uma substituição de uma por outra, mas na validação de todas elas, inclusive daquelas marginalizadas. À escola, inegavelmente, cabe o ensino da modalidade que o estudante não domina, isto é, as normas cultas e planejadas, que permitem a inserção desse sujeito na sociedade de prestígio, uma vez que a língua vernacular já lhe é culturalmente conhecida (BORTONIRICARDO, 2004; COELHO et al, 2018). Nessa perspectiva, não é uma particularização dialetal no ensino que se propõe, da mesma forma que uma unidade também não. O que se destaca, de fato, é a conscientização acerca da plurivariabilidade da língua e seus condicionamentos (LIMA; COSTA, 2022, p. 719).

Atualmente, de forma ainda tímida, alguns materiais de ensino de português já trazem aspectos que refletem as pesquisas sobre a língua falada, aproximando academia – escola – sociedade. A *Base Nacional Comum Curricular – BNCC –* (BRASIL, 2017), documento oficial que norteia a educação básica no Brasil, de forma assertiva e com base nos *Parâmetros Curriculares Nacionais – PCN –* (BRASIL, 1997), contempla aspectos que direcionam para um ensino de língua mais libertário e convergente com a diversidade de estudantes, normas, dialetos e falares espalhados por todo território brasileiro, haja vista que

[...] essa nova pedagogia conscientizadora linguística e de respeito de fato à diversidade dialetal brasileira, respeito no sentido de, no processo de ensino da língua materna, trabalhar-se com as variantes conviventes, em nível de igualdade e de interesse, sabendo explicá-las, sem escamotear para os alunos as avaliações sociais para cada uma delas, é um caminho fundamental [...] é uma das formas de começar a transformar a sociedade discriminadora em que vivemos, uma vez que sabemos que, nelas, as variantes linguísticas desempenham papel de profunda e complexa significação (MATTOS E SILVA, 1995, p. 51).

Após apresentar os aportes teóricos das duas áreas que lidam com a variação sobre a língua falada, Dialetologia e Sociolinguística, às quais esse estudo se vincula, nota-se que, como se pontuou nesta subseção, epistemologicamente, há aspectos que apontam ora para a convergência ora para divergência entre elas.

Este trabalho situa-se entre as duas perspectivas teóricas, vinculado ao que apregoam as duas áreas de estudo da variação. Os pronomes *eu* e *mim* terão a sua distribuição areal descrita prioritariamente, a partir do que espelham os dados das capitais brasileiras, apoiando-se no método cartográfico dialetal, a fim de melhor esboçar esse tipo de variação. Também, os condicionamentos da estrutura linguística e da estrutura social serão controlados e, de forma sistemática, analisados à luz do suporte variacionista. Tal junção é considerada importante, haja vista o objetivo de traçar um panorama do fenômeno no Brasil, no que tange à língua falada e aos *corpora* analisados.

2.4 ROTEIROS POSSÍVEIS PARA OS CASOS DE “NÃO DADO” E POUCOS DADOS

Ao observar a história das duas principais áreas responsáveis pelos estudos sobre a variação linguística, Dialectologia e Sociolinguística, bem como aos seus métodos, nota-se que há algumas divergências no modo de tratar os dados linguísticos recolhidos nas pesquisas *in loco*, para além das abordagens prioritárias, estas últimas, vistas na subseção anterior. Com isso, quando uma pesquisa pretende unir os seus respectivos aportes teóricos, alguns questionamentos surgem, haja vista as notáveis diferenças no que concerne à exegese.

Dentre as inúmeras perguntas cabíveis, pretende-se, aqui, destacar três, com o fito de promover e fomentar outras discussões sobre as temáticas, bem como instaurar, quiçá, um novo olhar e/ou perspectiva para o modo de tratamento do não dado e dos poucos dados nas pesquisas sobre a língua falada.

Após esse preâmbulo, passa-se para as indagações: a) Como tem sido trabalhado o não dado nas pesquisas sobre a variação linguística?; b) Há diferenças simbólicas para a interpretação do não dado?; e, por sua vez, c) As pesquisas com poucos dados podem atestar a existência/inexistência dos fenômenos, no que tange à língua falada?

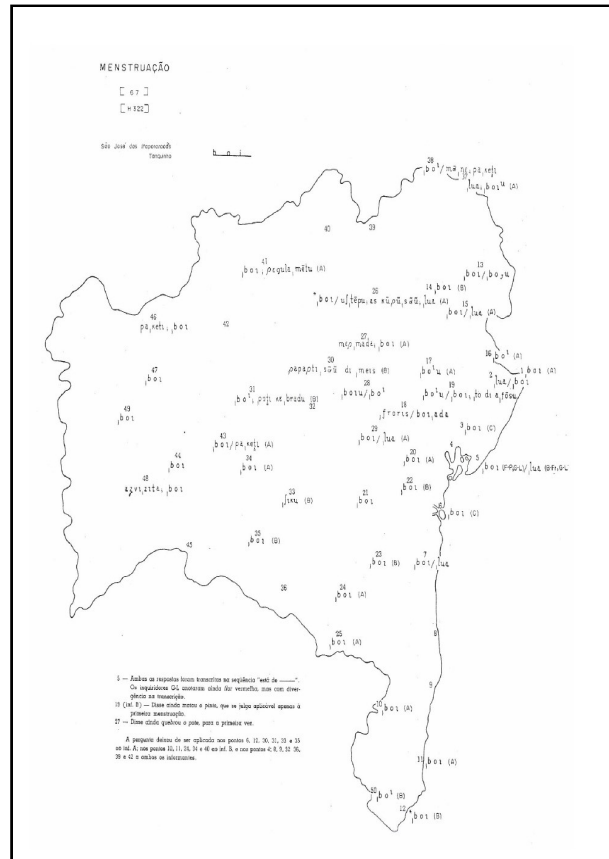
Espera-se que, com o esboçar das respostas, no discorrer desta subseção, haja um novo caminho para o entendimento e para a compreensão de dialetólogos, sociolinguistas e cartógrafos sobre esses aspectos que, até então, vêm recebendo atenção ínfima. A discussão inicia-se, então, retomando o primeiro questionamento:

a) Como tem sido trabalhado o não dado nas pesquisas sobre a variação linguística?

Para respondê-lo, faz-se necessária uma regressão no tempo cronológico, com o intuito de comprovar, com dados dos atlas linguísticos regionais/estadual, o local que vem sendo, historicamente, atribuído à não obtenção dos dados.

Vale ressaltar que a pesquisa possui um caráter ilustrativo, embora não tenha sido exaustiva. Por exemplo, no APFB, pioneiro fruto geolinguístico brasileiro/baiano, conforme subseção 2.1.2, as cartas apresentam os itens catalogados, nos respectivos pontos, representados a partir de transcrições fonéticas. No entanto, ao observar as cartas, nos locais em que nenhuma resposta foi computada, os pontos aparecem sem nenhuma sinalização, deixando a informação lacunar. Tal assertiva pode ser comprovada, por exemplo, na Carta 87 Menstruação, Figura 11.

Figura 11 – Carta 87 Menstruação do APFB

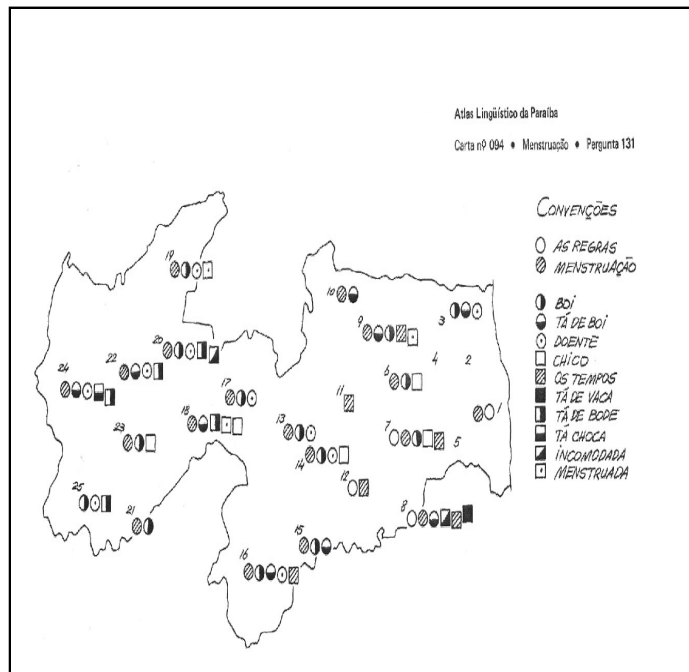


Fonte: Rossi (1963).

Nota-se que, ao observar a Figura 11, para alguns pontos sem nenhuma resposta cartografada, há uma nota explicativa na própria carta do atlas, tais como nos pontos 4, 8, 9, 32, 36, 39 e 42. Em um ponto – 45 – não há informação sobre o que ocorreu na respectiva localidade, deixando, desse modo, diversas interrogações para quem tem acesso à carta.

Nessa mesma linha, seguiu a cartografia do ALPB (ARAGÃO; MENESES, 1984), quando se observa a Carta 094-Menstruação, Figura 12.

Figura 12 – Carta 094 Menstruação do ALPB

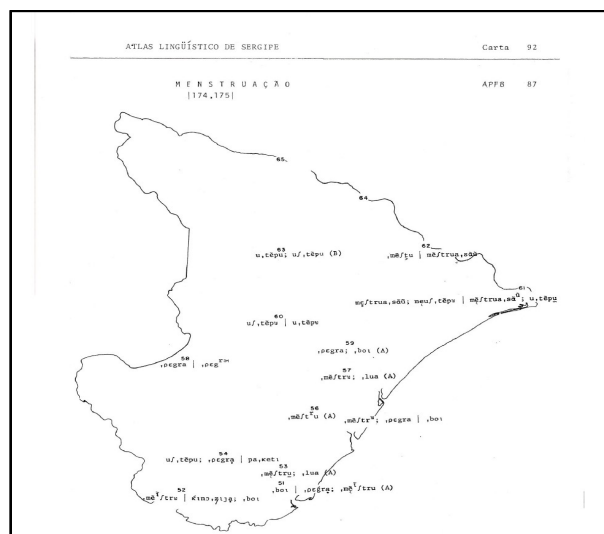


Fonte: Aragão; Menezes (1984).

Na referida carta do ALPB, Figura 12, os pontos 2, 4 e 5 não possuem informação sobre o item pesquisado, muito menos notas explicativas e/ou remissivas. Novamente, assim como o APFB, a cartografia linguística deixa a cargo do leitor fazer inferências possíveis pela falta da informação.

Ainda, em perspectiva alusiva/histórica, recorre-se a outro exemplo, ao ALS (FERREIRA et al., 1987), a partir da Carta 92-Menstruação.

Figura 13 – Carta 92 Menstruação do ALS



Fonte: Ferreira et al., (1987).

Por sua vez, a lacuna da informação linguística, no que tange à cartografia, foi notada na Figura 13, mais especificamente nos pontos 64 e 65. Mais uma vez, outro importante atlas para os estudos dialetais brasileiros reproduz um modo de cartografar que evidencia o tratamento dado para a não obtenção de dados, o que, conseqüentemente, pode não favorecer e facilitar que a leitura das cartas seja feita plenamente, levando ao leitor a inferências de diversas ordens.

A intenção não é rechaçar os critérios eleitos pelos estudos dialetológicos mais antigos, mas apontar que, embora a ausência possa ser representada pela não marcação nas cartas, tal decisão metodológica dificulta a leitura e, conseqüentemente, o entendimento das cartas linguísticas, uma vez que as informações adicionais sobre o fato precisam estar em outro local, no texto escrito do pesquisador e/ou nas notas das cartas. No entanto, defende-se que as cartas linguísticas devam ser vistas como textos independentes, ou seja, sejam textos autônomos e de fácil acesso aos seus leitores e consulentes.

A tradição dialetológica brasileira, representada aqui por três dos cinco primeiros atlas publicados – APFB; ALPB e ALS – não costumava representar o *não dado* nas cartas linguísticas. Esse modo cartográfico ultrapassou gerações e produtos geolinguísticos, ao observar, em perspectiva histórica, às propostas de periodização dessa área da Linguística, bem como os demais atlas publicados no Brasil até os dias atuais. Às vezes, tais informações aparecem em notas, apontando diversos motivos, como por exemplo: questões de ordens técnicas como falhas no equipamento gravador e/ou nas fitas cassete, descuido e/ou desatenção por parte do inquiridor, cansaço, inibição, esquecimento e/ou desconhecimento do referente pesquisado, estas últimas, por parte do informante.

Nas pesquisas em Sociolinguística, como se observou na seção 2.2, adjetivada, desde os primórdios, como Quantitativa, a busca exaustiva pelo dado linguístico leva ao pesquisador a não ter um olhar mais apurado para as ocorrências em que a informação buscada não teve êxito. Também, ao conferir os frutos sociolinguísticos, oriundos dessa vertente de estudos sobre a língua falada, constata-se a pouca representatividade da não obtenção do dado almejado.

No Brasil, nota-se que tal pensamento vem sendo, aos poucos, renovado e atualizado, permitindo novos horizontes para o não dado. Nesse sentido, Aguilera e Yida (2008), por meio dos dados do Projeto ALiB, analisaram as respostas fornecidas pelos informantes das capitais, para as 159 questões do Questionário Fonético-Fonológico – QFF – dando ênfase para as não respostas e para as respostas inadequadas, haja vista que

Dentre as dezenas de trabalhos que já foram desenvolvidos a partir dos dados coletados para o ALiB, nenhum ainda se voltou para o problema da ausência de respostas ocorrida durante a recolha. Esta questão, no entanto, é crucial no momento da cartografia dos dados pelas lacunas que deixa nos espaços destinados ao preenchimento com as variantes buscadas (AGUILERA; YIDA, 2008, p. 17).

Sendo assim, confere-se a este trabalho ser o pioneiro a abordar a temática no âmbito do Projeto ALiB. Refuta-se, apenas, ao que nomearam de “problema da ausência de respostas...”, haja vista que tal ausência pode significar aspectos outros, para além da expectativa do pesquisador, ao tentar documentar o fato linguístico. Ainda conforme as perspectivas de Aguilera e Yida (2008), após o levantamento e o cotejo dos dados, sob a ótica diatópica e diastrática, as conclusões apontam para alguns fatos, a saber:

(i) Os informantes de baixa escolaridade tiveram um índice de não-respostas bem maior do que o índice dos informantes de nível superior, o que revela que a escolarização contribui na ampliação do universo cognitivo dos entrevistados; (ii) A postura do entrevistador diz respeito à maior ou menor facilidade de reformular a questão para chegar à resposta desejada; (iii) A maior parte das questões levou a respostas válidas do ponto de vista do preenchimento da pergunta, mas inadequadas para o objetivo do QFF; [...] (V) Os informantes da região Nordeste apresentaram o maior índice de não-respostas, seguidos dos falantes das regiões Norte e Sudeste (informantes de baixa escolaridade) e Centro-Oeste e Sudeste (informantes do nível Superior) [...]; (vii) As mulheres, ratificando pesquisas anteriores (Silva-Corvalan: 1989 e Paiva: 2003), foram mais produtivas, com menor número de abstenções. (viii) Quanto ao fator diageracional, os idosos demonstraram maior conhecimento de mundo com um índice menor de não-respostas quando comparados com a produtividade dos mais jovens (AGUILERA; YIDA, 2008, p. 28-29).

Ao retomar essa discussão, anos depois, em 2014, a partir de um espectro mais amplo, dados de todos os questionários do Projeto ALiB, após findar as pesquisas de campo, Aguilera (2014), no volume 1 do atlas nacional, apresenta reflexões sobre a temática:

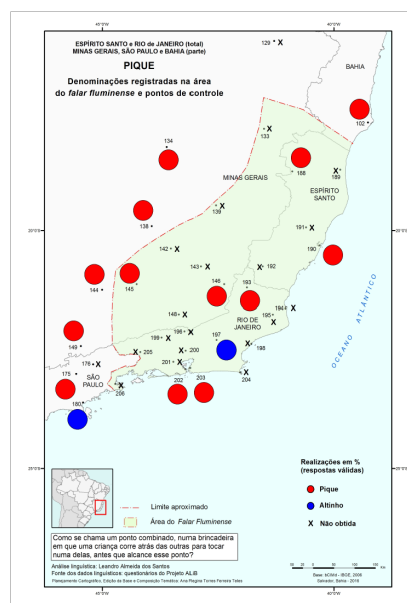
[...] Salientaram-se neste estudo as dificuldades que algumas questões e campos semânticos apresentaram: (i) ao informante no momento de fornecer o registro adequado e (ii) ao entrevistador para reformular a questão e torná-la acessível à compreensão do entrevistado; foram muitas as lições que aquele (entrevistador) aprendeu com este (informante), tais como compreender melhor o mundo do outro, suas limitações e seus medos diante de uma situação diferente, pouco frequente em seu cotidiano. A combinação de vários elementos envolvidos na pesquisa de campo – questionários pertinentes, bem elaborados e testados, informantes bem selecionados e inquiridores devidamente preparados – leva a resultados satisfatórios. No entanto, em se tratando de informantes e inquiridores, está-se dialogando com seres humanos complexos, com reações nem sempre previsíveis, fazendo de cada entrevista uma realidade única, irrecorrível (AGUILERA, 2014, p. 111).

Santos L. (2016), também, através dos dados do Projeto ALiB, mas restrito ao universo do campo temático dos jogos e das diversões infantis, propõe, para a Dialetologia e Cartografia brasileiras, a representação das não-respostas, sugerindo que o caminho trilhado seja, de fato, acrescido nos demais estudos hodiernos. Desse modo,

[...] Vale destacar a inovação em cartografar os dados NS/NL/NO, agrupados na forma *não obtida*, tal feito torna-se importante para os estudos dialetológicos, por se tratar de uma realidade encontrada nas pesquisas dessa natureza, e que, geralmente, ficava fora do registro cartográfico, além de poder colaborar para o entendimento dos fenômenos estudados e representados no espaço geográfico, deixando de forma explícita os dados encontrados nas localidades, extinguindo possíveis inferências dos leitores das cartas (SANTOS, L., 2016, p. 97).

Nesse sentido, as Figuras 14 e 15 são exemplos que ilustram a representação sugerida pelo referido autor:

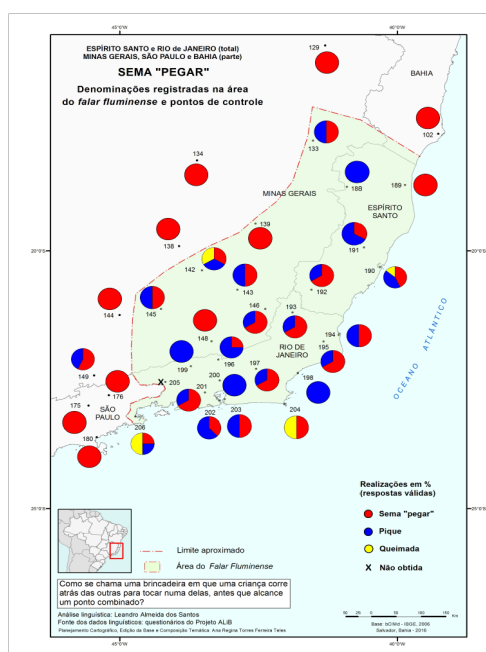
Figura 14 – Carta Pique



Fonte: Santos L. (2016, p. 159).

Nota-se que a informação linguística lacunar, antes deixada nos pontos em que nenhuma resposta foi catalogada pela tradição dialetológica, foi preenchida, deixando a leitura das cartas mais acessível. Tal fato pode ser observado também na Figura 15, em um único ponto, 205.

Figura 15 – Carta Sema Pegar



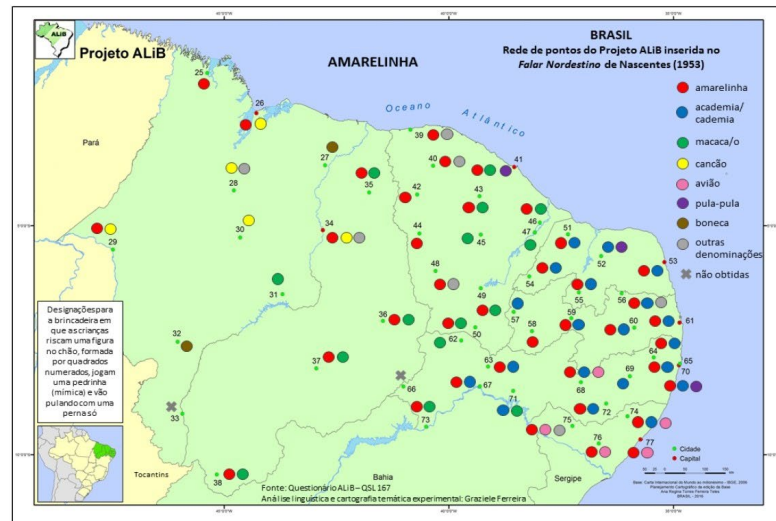
Fonte: Santos L. (2016, p. 153).

Santos L. (2020), ao aprimorar a proposta anterior, de 2016, coteja dados das pesquisas sobre áreas dialetais do Brasil – Ribeiro (2012); Portilho (2013); Santos L. (2016); e Santos G. (2018) –, por meio do campo temático jogos e diversões infantis do Projeto ALiB, reforçando a relevância das não respostas para as pesquisas sobre a língua falada, comparando os resultados obtidos para a questão 164 do Questionário Semântico-Lexical. Ao constatar o elevado índice de não obtenções nos referidos estudos, Santos (2020) tece algumas considerações:

O alto grau de importância do não dado para as investigações dialetais e sociolinguísticas. As NS/NL/NO devem, também, ser o foco de observação dessas pesquisas e, além disso, devem ser cartografadas, conforme proposição de Santos (2016). Com isso, o intuito é fazer um melhor uso dos dados coletados, ou seja, não deve haver frustração por parte do pesquisador, em relação aos dados recolhidos. Sugere-se, então, usufruir de todo o material que a pesquisa empírica forneceu, seguindo os critérios de tratamento do material coletado. As considerações aqui apresentadas não são definitivas e visam a apontar possibilidades de mais estudos sobre áreas dialetais brasileiras, bem como um tratamento mais adequado para a exegese dos dados sociodialetais, sobretudo, às não respostas (SANTOS, 2020, p. 17).

Vale mencionar, nesse sentido, dois estudos recentes, após a proposta sugerida por Santos (2016), que já trazem em suas respectivas cartas linguísticas o não dado representado, Santos G. (2018) e Souza (2021), Figuras 16 e 17, o que permitem vislumbrar um tratamento mais criterioso às não respostas, bem como um novo cenário, em que outros estudos e/ou outras cartas linguísticas sejam aprimorados, facilitando a leitura e o acesso aos consulentes.

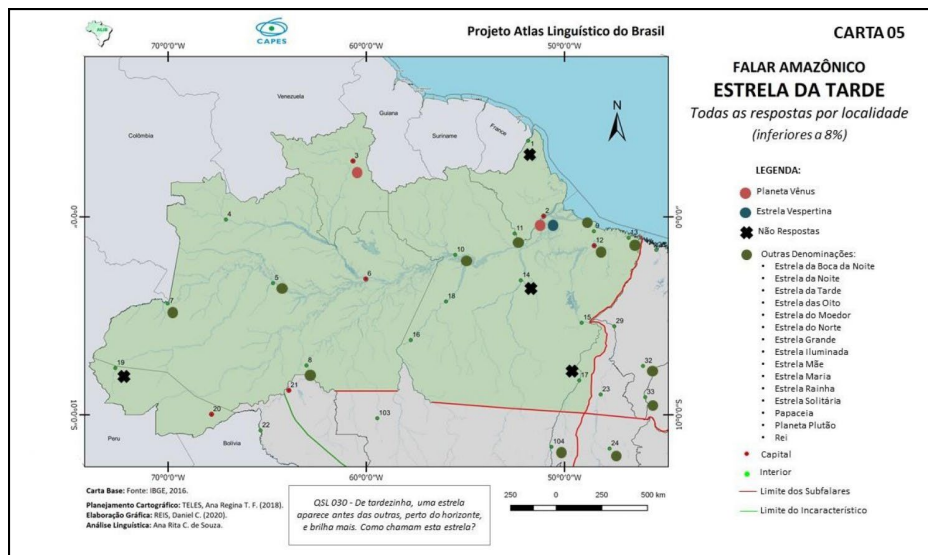
Figura 16 – Carta Amarelinha



Fonte: Santos G. (2018, p. 196).

Ao observar a Figura 16, Carta Amarelinha, no estudo de Santos G. (2018), visualiza-se que há variantes linguísticas cartografadas, mas também há pontos em que não houve dado catalogado, expresso pela simbologia do “não obtidas”, pontos 33 e 66.

Figura 17 – Carta Estrela da Tarde



Fonte: Souza (2021, p. 172).

Nota-se, na observação da carta Estrela da Tarde, Figura 17, a preocupação da pesquisa em fornecer melhores caminhos para a leitura da carta, tendo em vista que, em alguns pontos, 1, 14, 17 e 19, não houve respostas catalogadas inferiores a 8%, conforme critério estabelecido no estudo.

Ao comparar os frutos geolinguísticos do passado e do presente, no que concerne ao tratamento e à cartografia das respostas não obtidas, há, atualmente, um notável aprimoramento no modo de representação. Isso é perceptível nos trabalhos anteriormente mencionados: Santos L. (2016); Santos G. (2018); e Souza (2021). Além desses, nota-se também o trabalho de Silva (2020). Destaca-se que, além das diferenças estéticas, se a intenção é facilitar a leitura das cartas, a representação cartográfica do não dado pode colaborar e pode promover maior acesso a elas.

A inexistência do dado linguístico, expressa pela cartografia, leva o leitor a inferências de várias hipóteses. Nas pesquisas de campo, tais ausências, em determinados pontos, são comuns nas pesquisas sobre a língua falada. Porém, no que tange à exegese e à cartografia, acredita-se que, ao tratar adequadamente o não dado, indicando, por meio de símbolo cartográfico, que ele não foi obtido, o linguista e o cartógrafo cumprirão os seus papéis de facilitadores, ajudando ao público em geral na leitura e na interpretação da carta, do cartograma, do mapa e, conseqüentemente, dos atlas linguísticos, tão úteis e necessários para o entendimento da diversidade linguística, quer seja pela existência, quer seja ausência do dado.

Para além disso, o uso de tais informações pode indicar caminhos interpretativos de cunho não somente linguístico, mas também antropológico, sociológico e histórico sobre as localidades. No entanto, existem outros aspectos que devem ser discutidos na exegese e na interpretação desses dados, que serão alvos de considerações propositivas na tentativa de responder à segunda pergunta.

b) Há diferenças simbólicas para a interpretação do não dado?

Uma resposta afirmativa simples daria conta de elucidar tal questionamento, mas espera-se, para a discussão proposta, um pouco mais de detalhamento. Para tal, recorrem-se às várias possibilidades que possam representar o Não Sabe (NS); Não Lembra (NL) e o Não Obtida (NO), nas palavras de Ziamandanis (1999):

Indica falta de resposta quando o informante quer responder mas não consegue; ele fica irritado porque tem a palavra na ponta da língua, lembra que sempre fala que é um pai maior, mas não consegue acessar. Em vez disso, o pesquisador sinaliza ignorância quando o informante desiste; Ele dá de ombros e seu rosto reflete uma expressão vazia. A ignorância é o desconhecimento de um conceito, do uso de um objeto e revela uma perda definitiva de uma parte da expressão (ZIAMANDANIS, 1999, p. 657-658)¹³.

¹³ Apunta falta de respuesta cuando el informante quiere responder pero no puede; está molesto porque tiene el vocablo en la punta de la lengua, o recuerda que lo decía siempre un pariente mayor, pero no logra accederse a él. En cambio, el investigador, apunta desconocimiento cuando el informante se da por vencido; se encoge de

Logo, nota-se que, apesar de, geralmente, serem agrupados, tais dados representam informações diferentes, possuindo simbologias próprias e indicando possibilidades de leitura de diversas ordens.

Para as questões em que o informante responde desconhecer o referente pesquisado (NS), mesmo após as reformulações feitas pelo inquiridor, o valor dessa informação vem logo expresso, uma vez que o referente buscado não pertence ao universo de saber do inquirido. Muitas vezes, há formas que caíram em desuso, logo, não são mais conhecidas pelos informantes, como são os casos dos elevados índices de não obtenções para alguns brinquedos e brincadeiras infantis, tais como o *chicotinho-queimado*, por exemplo. Nesse sentido, o desconhecimento do referente buscado evidencia as mudanças ocorridas socialmente, que podem estar relacionadas com as novas formas de brincar na contemporaneidade (SANTOS, 2020).

Para as perguntas em que o informante diz não lembrar (NL) do referente perguntado, há uma valoração que pode estar diretamente ligada com o momento da entrevista, haja vista que as entrevistas são gravadas e, geralmente, não existe relação de proximidade entre os envolvidos no evento. Outra hipótese que pode ser aludida aqui, ainda nesse âmbito do momento situacional, diz respeito à inibição, pois, em alguns casos, podem existir *tabus* linguísticos, alimentados socialmente, sobre determinados campos temáticos e/ou questões. Citam-se, por exemplo, a menstruação e suas variantes que podem não figurar na fala dos informantes homens, quando, geralmente, interrogados por inquiridoras. Vale lembrar que essas hipóteses também ilustram os casos de NS, dentre outras hipóteses que podem ser aventadas.

Por sua vez, para os casos das respostas que não foram obtidas, as (NO), são vários os fatores que podem ser elencados, tais como: por parte do inquiridor, esquecimento em realizar a pergunta, ter pronunciado o sema da questão, ao fazê-la; ou por questões de ordens técnicas, falhas nos equipamentos de gravação, pelo ambiente, muitas vezes, não ser apropriado, o aparelho não ter captado de forma audível parte do inquirido, são alguns dos exemplos que podem ser lidos e interpretados para, posteriormente, ser inclusos nesse subgrupo de tratamento das respostas não obtidas. Dessa forma,

As autorreflexões e autoanálises, portanto, são constantes na pesquisa sociodialetal, sobretudo no momento da exegese, quando decisões precisam ser tomadas, a partir do material coletado. Advoga-se, nesse sentido, um melhor tratamento, discussão e

aproveitamento das não respostas. Para além de fazer a quantificação, por meio dos critérios sociais estabelecidos pela pesquisa, também levantar hipóteses de interpretação dessa informação de cunho social, histórico e cultural (SANTOS, 2020, p. 15-16).

Então, sugere-se que haja cautela, por parte dos pesquisadores, no tratamento das diferentes simbologias que o não dado fornece. Não se rechaça a proposta dos agrupamentos, entre dados linguísticos validados e dados em que, por diversos motivos, o objetivo desenhado e esperado pelo pesquisador não foi alcançado. No entanto, caso seja de interesse da pesquisa, um olhar diferenciado e interpretativo permitirá o levantamento de hipóteses para os diversos caminhos que o NS, NL e NO podem fornecer.

Nesta tese, desenvolveu-se a percepção para, por exemplo, os casos das NO, visto que, em muitas comunidades, o fenômeno ora analisado é cercado por monitoramentos linguísticos, devido ao estigma que sofrem os informantes que têm como usos os *pra mim fazer; pra mim ler; pra mim ir* etc. Por tal motivo, às vezes, livros, materiais didáticos e até o meio social propagam as noções de que essas construções são “erros gramaticais”, inibindo que tais ocorrências sejam documentadas, cerceando, dessa maneira, a manifestação linguística vernacular.

c) As pesquisas com poucos dados podem atestar a existência/inexistência dos fenômenos, no que tange à língua falada?

Na tentativa de responder ao terceiro questionamento feito no início desta subseção, faz-se outra pergunta: Quantos dados devem ser exaustivamente catalogados para que sejam representativos nas pesquisas sobre a variação linguística?

Buscou-se, na vasta literatura das duas áreas, Dialetoлогия e Sociolinguística, informações que pudessem elucidar esses questionamentos. Nos estudos dialetológicos, poucas ocorrências e/ou até uma única ocorrência são de enorme relevância para a pesquisa. Muitas vezes, esse fato é desconsiderado pelos pesquisadores, mas, ainda assim, a validação e o interesse são diferenciados, quando se coteja com os estudos sociolinguísticos. Para o estudo dialetológico, a partir do que apregoa Rossi (1967):

[...] Convirá, porém, nunca esquecer que a dialectologia é essencialmente contextual: fato apurado num ponto geográfico ou numa área geográfica só ganha luz, força e sentido documentais na medida em que se preste ao confronto com o fato correspondente – ainda que por ausência – em outro ponto ou outra área (ROSSI, 1967, p. 104).

Diferentemente, nas palavras de Medina Morales (2005), encontram-se respostas sugestivas que fornecem material para pensar, mais uma vez, a validação dos fenômenos linguísticos encontrados nas pesquisas *in loco*, em especial, àquelas que tomam por base as orientações e os fundamentos sociolinguísticos, a saber:

Uma questão importante é a densidade das variantes no texto, ou seja, que estas sejam numericamente relevantes para não confundir um fato sociolinguístico, ou seja, realizado por um grupo de falantes, com uma manifestação individual. Nesse sentido, Labov estabeleceu para a sociolinguística de campo que dez ou vinte dados linguísticos de uma determinada variável são suficientes para representar uma matriz de variação. No entanto, estes números podem ser válidos para alguns casos e não para outros, embora ele parta do princípio da homogeneidade do comportamento linguístico (MEDINA MORALES, 2005, p. 125)¹⁴.

Nesse sentido, quando se observa a tradição dos estudos em Sociolinguística, há uma primazia pelas pesquisas que apresentam dados quantitativos expressivos, com números elevados de ocorrências, sendo estes os critérios mais importantes para a validação desses estudos. Sendo assim, historicamente, para esta área de estudos sobre a língua falada, a quantidade é posta em um lugar de privilégio, os dados numéricos/percentuais possuem grau de cientificidade, após submissão da informação linguística aos programas computacionais, matemáticos e estatísticos. Ressalta-se que tais procedimentos são necessários.

Para ilustrar, por exemplo, os percentuais apresentados pelas variantes binárias objetos desta tese, os pronomes *eu* e *mim*, nas estruturas e contextos analisados, sobretudo o uso do *mim*, não documentados com alta frequência, mas também não são poucos os casos, inserem-se no limiar e são amplamente reconhecidos pelos falantes da língua, podendo significar estigma linguístico, quanto à utilização do *mim* em certos contextos e em dadas comunidades. Tal aspecto, no que se refere à obtenção de formas sintáticas, já foi ressaltado por Labov (2008 [1972], p. 223), quando aponta para *A raridade das formas sintáticas*.

Mesmo sabendo das diferentes orientações teóricas e metodológicas das duas áreas da variação, que devem ser respeitadas, entende-se que as pesquisas que documentaram uma quantidade não expressiva de dados da língua possam ser igualmente representativas, uma vez que a inexistência do fenômeno na comunidade pesquisada pode indicar, isoladamente e/ou conjuntamente, fatores intralinguísticos e extralinguísticos de relevância ímpar para a atual pesquisa e/ou para futuras pesquisas.

¹⁴ Una cuestión importante es la densidad de las variantes en el texto, a saber, que éstas sean numéricamente relevantes para no confundir un hecho sociolingüístico, esto es, realizado por un grupo de hablantes con una manifestación individual. A este respecto, Labov estableció para la sociolingüística de campo que diez o veinte datos lingüísticos de una variable dada son suficientes para representar una matriz de variación. No obstante, estas cifras pueden ser válidas para unos casos y no para otros, aunque él parte del principio de homogeneidad de la conducta lingüística.

Nesta tese, por exemplo, ao unir os aportes teóricos da Dialetologia e da Sociolinguística Quantitativa, se situando na interface entre os dois ramos, defende-se que os dados coletados em sincronias distintas, como propõem essas áreas, sejam validados, de modo a espelhar as diferenças diatópicas e/ou sociais, no que tange ao fenômeno em questão, na fala dos brasileiros, oriundos das capitais.

Nessa perspectiva, de fato, constata-se que as ocorrências que envolvem *para/prá + eu/mim + infinitivo*, sejam elas com os usos do pronome previsto pela norma canônica, *eu*, ou o pronome que surge como forma “inovadora”, *mim*, portanto, alvo de estigmas, podem fornecer explicações de ordem linguística, vindas do latim vulgar, por exemplo, mas também de cunho social, haja vista que, em algumas comunidades, a forma inovadora é estereotipada, ao passo que, em outras, é a norma utilizada, até mesmo por falantes que possuem acesso aos bens culturais e têm bastante tempo de escolarização.

Entende-se que, tanto no ramo da variação espacial quanto no ramo da variação social, o não dado seja utilizado de forma produtiva, bem como as pesquisas que não obtiveram expressivos dados quantitativos sejam, também, validadas, tendo em vista o universo teórico e metodológico das áreas e os propósitos do pesquisador. As alternâncias pronominais entre o *eu/mim* ora examinadas, embora não constate um número significativo de dados, pelas porcentagens documentadas, não devem ser vistas como sem significação, sobretudo pela orientação teórica basilar escolhida para observação dos dados, pois, para as pesquisas que têm os pressupostos teóricos dialetológicos, como nesta tese, vale destacar que:

Dois aspectos fundamentais estão, pois, na gênese da dialetologia, qualquer que seja o princípio metodológico seguido: A) o reconhecimento das diferenças ou das igualdades que a língua reflete e B) o estabelecimento das relações entre as diversas manifestações linguísticas documentadas ou entre elas e a ausência de dados registrados, circunscritos a espaços e realidades prefixados (CARDOSO, 2010, p. 25).

Com isso, o anseio é que as provocações esboçadas nesta tese, em especial, nessa subseção, sirvam para fomentar mais discussões e para a produção de material investigativo, a fim de que se possa, cada vez mais, aprimorar as ferramentas e os métodos das áreas da variação, inclusive, neste trabalho.

2.5 SÍNTESE DA SEÇÃO

Esta seção foi construída com o objetivo de apresentar os pressupostos teóricos que dão suporte a essa tese. Na primeira subseção, 2.1, são apresentados os fundamentos

dialetológicos, desde o seu surgimento, contexto interno e externo, e os avanços desse ramo da Linguística, bem como o estágio atual, no que se refere ao Brasil. Além disso, foram apresentados fatos salustares da Dialetoлогия brasileira, tais como as propostas de divisões dialetais, os marcos de periodização em diversas perspectivas e uma tentativa de dirimir um dissenso terminológico quanto ao qualificador pluridimensional, ora usado para a Dialetoлогия ora para a Geolinguística, o método.

Nota-se, a partir da exposição e da trajetória desse ramo dos estudos linguísticos, um fazer dialetoológico progressivo, tanto em relação aos métodos empregados quanto ao número de trabalhos realizados, sobretudo, no Brasil, em forma de atlas, quer sejam em extensões territoriais macro e micro. Considerando essas diversas produções, percebe-se que, estando na 5ª fase de periodização, a Dialetoлогия brasileira possui atlas com dimensões diversas, a saber: monodimensional, bidimensional e pluridimensional. Logo, defende-se que tais qualificadores sejam associados ao método, à Geolinguística.

Na segunda subseção, 2.2, são apresentados os fundamentos sociolinguísticos e os principais conceitos advindos desse ramo dos estudos linguísticos, a fim de que se vislumbrem, pelo viés social, os possíveis caminhos de entendimento da configuração do fenômeno.

Nesta subseção, descrevem-se os caminhos trilhados pela Sociolinguística Variacionista, com o fito de evidenciar como esse ramo dos estudos linguísticos imprime uma nova perspectiva de abordagem sobre a língua falada, com ressonâncias em outros ramos, a exemplo, na Dialetoлогия.

Nas duas últimas subseções, 2.3 e 2.4, respectivamente, são apresentados dois aspectos: i. um cotejo entre os ramos responsáveis pelos estudos e pesquisas sobre a variação e mudança linguísticas, destacando os pontos convergentes e divergentes entre eles; ii. empreende-se uma discussão, com base em alguns fatos históricos, tais como cartas linguísticas dos atlas estaduais, sobre o lugar atribuído ao não dado (ausência de respostas) tanto nas pesquisas dialetoológicas quanto nas pesquisas sociolinguísticas, por meio de três questionamentos.

Em virtude do surgimento dos estudos sociolinguísticos, têm-se, na Linguística, dois ramos que se debruçam sobre a língua falada, Dialetoлогия, priorizando o aspecto espacial, e Sociolinguística, priorizando, o aspecto social. No entanto, há aspectos convergentes entre eles, sobretudo, quando focalizam, por exemplo, o combate ao preconceito linguístico.

Ao remontar a história desses estudos, dialetoológicos e sociolinguísticos, nota-se a ínfima relevância atribuída ao não dado/ a ausência de respostas, principalmente na exegese

e na cartografia. Todavia, conforme Santos (2016/2020), tais dados apontam para caminhos qualitativos e interpretativos, permitindo, dessa maneira, ao pesquisador elaborar ilações em perspectivas diversas.

Após essa exposição teórica que embasam a pesquisa, na seção 3, apresentam-se alguns aspectos históricos e sociais sobre os pronomes, de modo geral, do latim ao português brasileiro.

3 AS VIAS DOS PRONOMES

A aparência indivisa, de monobloco inquebrável ou de estrutura uniforme e não diversificada, que a muitos de nós pode a língua oferecer, é inteiramente falsa. A rigidez que se atribui à língua ou que nela se espera encontrar, por uma visão despistada da sua estrutura, inexistente agora, como antes, no português como em qualquer outra língua natural...

(CARDOSO, 1994, p. 229).

Nesta seção, são descritos, de forma breve, os aspectos históricos e sociais que moldaram a língua portuguesa falada no Brasil, dando ênfase aos fatores que desencadearam a formação do quadro pronominal que, com o passar dos anos, vem sofrendo diversas modificações. Vale mencionar que, antes de adentrar nas peculiaridades pronominais, os caminhos sociohistóricos e sociolinguísticos serão evidenciados, haja vista que são elementos fulcrais para se entender o atual quadro de variação e mudança do PB. Além disso, os aspectos referentes ao dissenso sobre algumas hipóteses para a formação do PB, bem como a categoria dos pronomes, no que tange à real natureza deles, serão enfatizados.

3.1 NAS ROTAS DA HISTÓRIA DA LÍNGUA PORTUGUESA

Nesta subseção, busca-se traçar um percurso histórico-cronológico das origens da língua portuguesa, com ênfase no latim, considerada como a mãe das línguas românicas. Nesse sentido, quando se refere à língua, os aspectos inerentes à história devem ser revisitados, posto que:

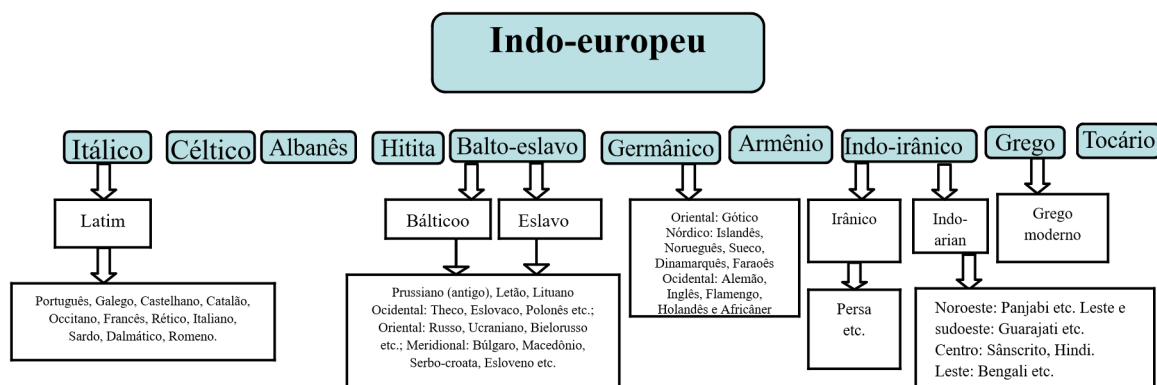
todo ato de fala, sendo ao mesmo tempo histórico e livre, tem uma extremidade ancorada em sua necessidade histórica, em sua condição historicamente necessária – que é a língua –, e outra extremidade que aponta para uma finalidade significativa inédita e que, portanto, vai mais além da língua já estabelecida (COSERIU, 1988, p. 47-48).

Para se compreender o presente, ou seja, o estágio sincrônico da língua e, conseqüentemente, dos fenômenos observados em dada sincronia, é necessário, por vezes, recorrer aos veios históricos, a fim de poder traçar possibilidades de entendimento e explicação para o comportamento da língua, principalmente dos fenômenos encontrados na língua falada.

Aqui, informa-se que não há intenção de fazer uma historicização longa e exaustiva, mas, apenas, apresentar aspectos salutares para auxiliar na compreensão do PB, como também fornecer meios para delinear o fenômeno ora estudado na tese. Em meio ao já citado

entendimento linguístico atual, através da ótica diacrônica, as origens do português remontam ao latim – literário, eclesiástico e vulgar – que, por sua vez, se oriunda do itálico, ramo do indo-europeu ocidental, como se observa na Figura 18.

Figura 18 – Genealogia da língua portuguesa



Fonte: Adaptado de Ilari e Basso (2006, p. 16).

Conforme evidenciado por Ilari e Basso (2006), o português – e as outras línguas românicas – têm sua origem a partir da variedade vulgar do latim, pois tal variedade passou a ser falada em grande parte nos lugares conquistados pelo Império de Roma.

Posto isso, passa-se, a partir de então, a situar os veios históricos da língua portuguesa, no que concerne aos processos de formação do povo e do PB¹⁵, com o entendimento de que as possíveis diferenças linguísticas encontradas tanto em relação às variedades geográficas quanto às variedades sociais, em um país com dimensões continentais, como o Brasil, se espelham, dentre outros aspectos, nos fatores desde o processo de ocupação do território, nos mais diferentes ciclos econômicos, nos processos de migração e imigração e na participação de matrizes étnico-linguísticas diversas.

Tais fatores são compreendidos à luz de cenários heterogêneos e com alto nível de complexidade, dadas às situações que envolvem esses fatos, que são salutares para contribuir para a agenda que visa a elucidar o PB falado, como o que se pretende nesta tese.

Com as navegações do Estado lusitano, a língua portuguesa aporta no Brasil no século XVI, o que, de certo modo, favoreceu e desencadeou muitas ações para a formação linguística do PB. Diante disso, inicialmente, destaca-se o contato entre o PE, trazido pelas embarcações portuguesas, e as línguas indígenas, os povos autóctones, de dois troncos

¹⁵ Maiores detalhes ler a vasta literatura sobre o assunto, para ilustrar, sugerem-se: Mattos e Silva (2004); Cardoso; Mota; Mattos e Silva (2006) e Faraco (2016).

linguísticos¹⁶ – Tupi e Macro-Jê. Houve, então, diversos choques – cultural e intencional –, uma vez que uma parte buscava colonizar – dominação seguida de exploração – enquanto outra parte buscava a catequização – convertimento por meio da religião. No entanto, os intentos não se efetuaram de forma plena, porque os povos autóctones apresentavam uma diversidade cultural e linguística, o que dificultou os projetos de expansão dos portugueses na nova colônia. Segundo Faraco (2016, p. 121):

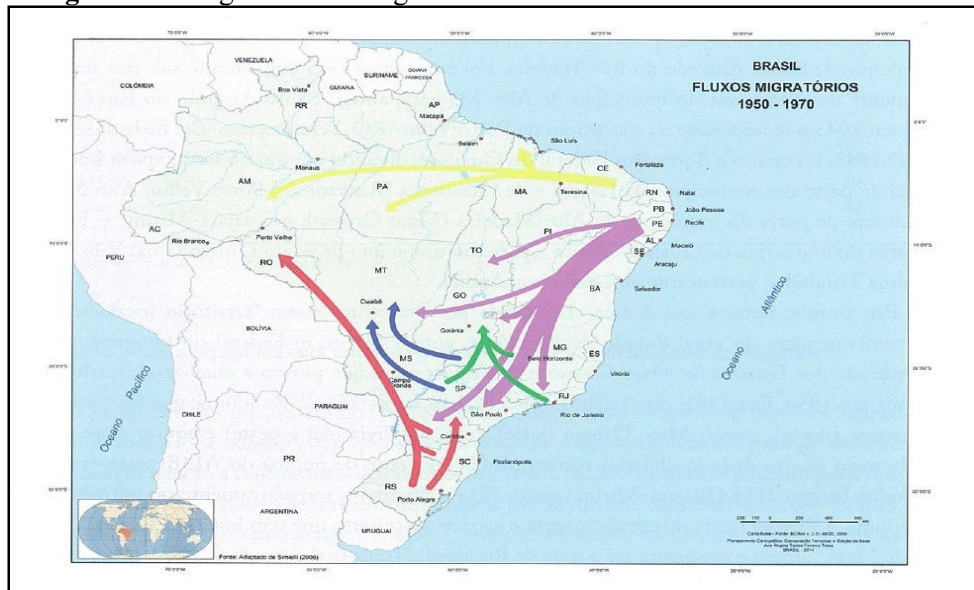
Essa desestruturação das sociedades originárias e o assentamento das bases de uma nova formação socioeconômica – a sociedade colonial – configuraram um novo quadro de relações sociointeracionais que afetou profundamente as línguas nelas envolvidas, desde seu completo desaparecimento (com o extermínio de seus falantes ou com a sua diluição da nova formação socioeconômica) até direções evolutivas acionadas pelas inusitadas situações de contato, aloglossia e bilinguismo que aqui se criaram, fazendo, de um lado, emergir as chamadas línguas gerais (paulista e amazônica) e, de outro, traçando as primeiras grandes linhas que resultaram no modo polarizado pelo qual se deu a disseminação da língua portuguesa do Brasil (FARACO, 2016, p. 121).

Nas três décadas iniciais da colonização, a língua portuguesa esteve restrita ao litoral, estabelecendo contatos com as línguas indígenas do tronco Tupi. Com o passar do tempo, por meio das Entradas e Bandeiras, avança-se para o interior e um novo contato foi estabelecido, dessa vez com povos pertencentes ao tronco Macro-Jê. Esses acontecimentos foram os propiciadores para o surgimento da língua geral paulista – oriunda do tronco Tupi, levada pelos bandeirantes pelas áreas sul e oeste de São Paulo, estruturada nas primeiras décadas do século XVII – e a língua geral amazônica – língua dos mestiços amazonenses, de base tupinambá, estruturada paulatinamente em tempo da ocupação de terras que seriam denominadas como o Estado do Grão-Pará e do Maranhão (RODRIGUES, 2010, p. 37-39).

Nas figuras 19 e 20, notam-se alguns fluxos migratórios ocorridos no Brasil em épocas distintas, 1950-1970 e 1970-1990, respectivamente. De certa maneira, esses movimentos do homem pelo espaço geográfico podem também ser levados em consideração para o entendimento do modo como se delinearam os perfis sociais e linguísticos no Brasil.

¹⁶ Em Rodrigues (1996), encontram-se maiores detalhamentos sobre os povos e línguas indígenas existentes na época da colonização.

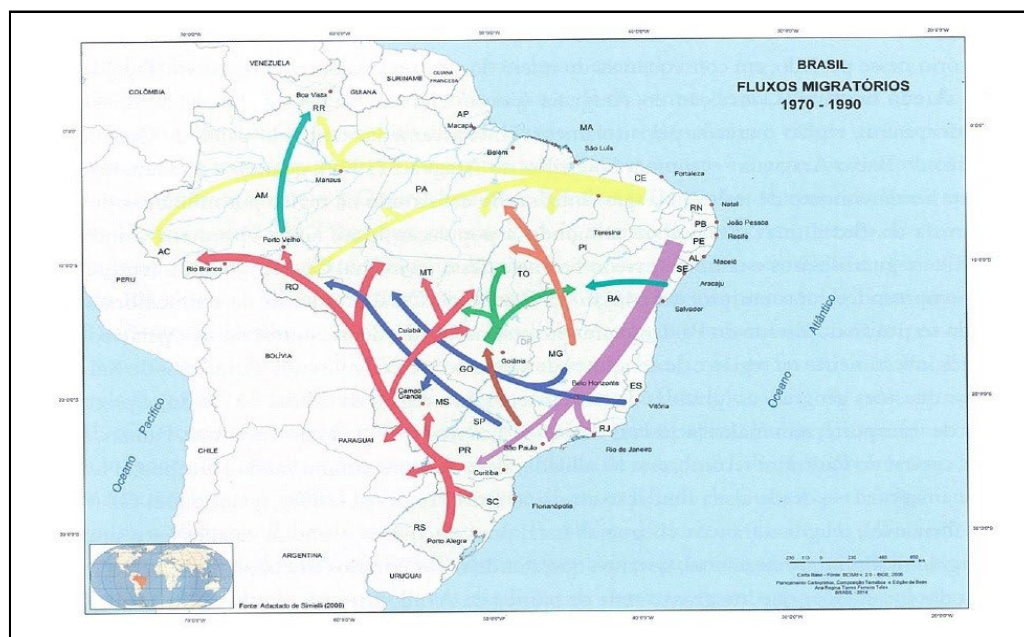
Figura 19 – Alguns fluxos migratórios ocorridos no Brasil entre 1950 a 1970



Fonte: Isquierdo e Teles (2014, p. 50).

Segundo Castilho (2016, p. 174), oito focos irradiadores foram responsáveis pelo povoamento e implantação da língua portuguesa no litoral do Brasil: i) inicialmente, no século XVI (1532-1554), São Vicente/São Paulo (1535), Olinda/Recife (1549), Salvador (1557), Rio de Janeiro (1557); ii) no século XVII, São Luís do Maranhão (1612) e Belém (1616); iii) no século XVIII, Florianópolis (1738) e Porto Alegre (1752). Baseada em ciclos econômicos, na época, a economia da colônia volta-se, inicialmente, para o cultivo e a produção de cana de açúcar, no Nordeste, propiciados pela mão de obra dos povos africanos, atores importantes no contexto linguístico brasileiro.

Figura 20 – Alguns fluxos migratórios ocorridos no Brasil entre 1970 a 1990



Fonte: Isquierdo e Teles (2014, p. 51).

Nos séculos XVI e XVIII, com a expansão e solidificação da economia açucareira, da mesma maneira que o tráfico de pessoas, grandes ciclos de importação de africanos chegaram ao território brasileiro, sendo a língua falada pelos povos africanos escravizados o principal meio difusor da língua portuguesa em terras brasileiras. É preponderante mencionar que o crescimento e direcionamento das levas de escravizados aumentavam a partir dos ciclos econômicos promissores da época, haja vista que

Associadas a esses ciclos estão as razões econômicas: cultivo de cana-de-açúcar e de tabaco, nos séculos XVI e XVII; exploração das minas de ouro e de diamantes, como também o cultivo do algodão, do arroz e da colheita de especiarias, no século XVIII, cultivo do café, no século XIX. Deve-se assinalar que o final de um ciclo não significa a interrupção da chegada de negros da região precedente; sendo assim, a separação deve ser considerada válida em suas grandes linhas, mesmo porque a política de Portugal foi sempre a de promover a mistura das diferentes etnias africanas, para impedir a concentração de africanos de mesma origem numa mesma capitania (PETER, 2006, p. 124).

Conforme a história social brasileira, afirma-se que outros fatos são importantes para a difusão da língua portuguesa em solo brasileiro, dentre eles destaca-se o ciclo do ouro, na transição entre os séculos XVII e XVIII, na região de Minas Gerais. Nesse período, ocorre a mudança da capital do Brasil, a sede deixa de ser Salvador e passa a ser o Rio de Janeiro. Tais acontecimentos foram propiciadores para a chegada de mais portugueses e, sobretudo, mais levas de africanos escravizados. Vale ressaltar que os mais diversos ciclos que marcaram a economia também foram responsáveis pelas mobilidades geográficas dentro do

território, pois, geralmente, havia deslocamentos migratórios internos e externos para a área central do ciclo econômico em apogeu. O Quadro 5 demonstra alguns dos ciclos econômicos mais rentáveis e o (s) século (s) em que foi/foram a mola mestra da economia brasileira.

Quadro 5 – Ciclos econômicos marcantes para a formação do povo brasileiro

SÉCULO (S)	CICLO
XVI	Pau-Brasil
XVI e XVII	Cana-de-Açúcar
XVI e XVII	Tabaco/Fumo
XVII e XVIII	Ouro/Mineração
XVII e XVIII	Pecuária/Gado
XVII	Drogas do Sertão
XVIII	Algodão
XIX	Cacau
XIX	Erva-Mate
XIX e XX	Borracha
XIX e XX	Café
A partir do XX	Soja

Fonte: Elaborado pelo autor, conforme Isquierdo e Teles (2014, p. 61).

Ainda sobre momentos históricos, no que tange ao cenário econômico e aos movimentos migratórios:

Nesse contexto, não pode ser desconsiderada a importância de cada um desses ciclos econômicos para o povoamento das diferentes regiões e, por extensa extensão, para a formação da sociedade brasileira, à medida que geraram inúmeras levas migratórias que favoreceram contatos interétnicos e linguísticos, sejam em decorrência do convívio entre brasileiros das diversas regiões, sejam entre a população brasileira e imigrantes que aqui fizeram e contribuíram para a configuração do povo brasileiro (ISQUERDO, 2014, p. 67).

Além dos ciclos econômicos, dentre os marcos históricos para a formação sociolinguística e social do povo brasileiro, destacam-se: o decreto do Marquês de Pombal, a partir de 1758 – tornando o uso obrigatório da língua portuguesa nas colônias de Portugal –, e a chegada da Família Real portuguesa – que acarretou um acelerado crescimento para a capital, Rio de Janeiro, com a criação de banco, imprensa e faculdades (LUCCHESI, 2009).

Sabe-se que as características culturais e linguísticas existentes no território brasileiro foram as responsáveis para o cenário de multilinguismo generalizado (MATTOS E SILVA,

2008). No entanto, diversos glotocídios e homicídios ocasionaram a dizimação de milhares de línguas tanto indígenas quanto africanas que, à época, compunham a torre de babel à brasileira. Logo, conforme aponta Faraco (2016, p. 142):

A hegemonia da língua portuguesa na sociedade brasileira não se deu, portanto, do dia para a noite, nem estava dada já no século XVI. Foi antes o resultado de um processo de longa duração intrinsecamente interligado com fatores socioeconômicos e não como pretenderam, algo ingenuamente, alguns filólogos que atribuíram a vitória do português à sua (suposta) superioridade... (FARACO, 2016, p. 142).

Para o quadro sociodialetal e sociolinguístico brasileiro, também, foram atraídos imigrantes, tais como alemães, italianos e japoneses, estimulados pelo governo brasileiro – que visava, com isso, ao branqueamento da raça –, mas essas línguas e culturas ficaram restritas aos pequenos grupos e em determinadas áreas geográficas. A partir do panorama apresentado, dessa forma, vale trazer à tona os atores principais para o quadro multidialetal brasileiro, a saber:

a. o português europeu na sua dialeção diatópica, diastrática e diacrônica, que teria ao longo do período colonial um contingente médio de 30% da população brasileira; seria esse português europeu base histórica do português culto brasileiro que começaria a elaborar-se a partir da segunda metade do século XVIII; b. as línguas gerais indígenas, que, plurais e dialetalizadas, poderiam até confundir-se com o português geral brasileiro nas áreas geográficas delimitáveis em que se difundiram; c. o português geral brasileiro, antecedente histórico do português popular brasileiro que, adquirido na oralidade e em situações de aquisição imperfeita, é difundido pelo geral do Brasil sobretudo pela maciça presença da população africana e dos afrodescendentes que perfizeram todo o período colonial (MATTOS E SILVA, 2008, p. 90).

A história social do português brasileiro pode ser sintetizada assim, na visão de Castilho (2010), conforme Quadro 6:

Quadro 6 – Panorama da história sociolinguística do PB

ANO	ACONTECIMENTO(S)
1500	Descoberta do Brasil, por Pedro Álvares Cabral.
1532	Começo do povoamento, em São Vicente (SP).
1534	Organização das capitanias hereditárias, doadas a pessoas de “pequena nobreza”. Apenas as de São Vicente e Pernambuco dão certo.
1535	Fundação de Olinda e Recife. Tem início a chegada de escravos africanos. O tráfico só terminaria em 1855. Teriam sido trazidos 18 milhões de escravos.
1549	Fundação de Salvador e organização do Governo Geral do Brasil, como sede em Salvador, Bahia.
1550	Povoadores minhotos se instalam em Pernambuco.
1554	Fundação de São Paulo.
1612	Fundação de São Luís do Maranhão.
1616	Fundação de Belém.
1654	O território passa a ser governado diretamente de Lisboa, como duas colônias independentes: o Estado do Brasil, com sede em Salvador, e o Estado do Grão-Pará e Maranhão, com sede em São Luís.
1565	Fundação do Rio de Janeiro.
1738	Fundação de Florianópolis.
1751	O Estado do Grão-Pará e Maranhão passa a ter sua sede em Belém. O irmão de Pombal, Francisco Xavier de Mendonça Furtado, impõe a língua portuguesa sobre a língua geral do Norte. A lei não deu certo, e até hoje se ouve falar nheengatu na Amazônia.
1752	Fundação de Porto Alegre.
+/-1790	A língua portuguesa domina sobre a língua geral paulista.
1808	Chegada da família real ao Rio de Janeiro; 16 mil portugueses saem dos navios.
1822	Independência do Brasil.
1832	Fundação das faculdades de Direito de São Paulo e Recife.
1870	Intensifica-se a migração europeia, sobretudo para o Sudeste e o Sul do país.
1889	Proclamação da República.
1922	Movimento modernista.
1930	Estado Novo.
1934	Fundação da Universidade de São Paulo.
1945	Redemocratização do Brasil.
1970	A população urbana excede a população rural.
1980	Expande-se a fronteira rural, surgindo novos centros em Rondônia e Roraima.

Fonte: Adaptado de Castilho (2016, p. 176).

O Quadro 6 demonstra alguns fatos que podem alimentar as ideias refutativas, portanto, sobre certa homogeneidade linguística no Brasil, pois o que se observa é um PB, nos mais diversos níveis, diferenciado do PE. Tal entendimento se faz à luz dos ramos de estudos sobre a variação, uma vez que apregoam que língua e fatores sociais não podem ser dissociados. Além disso, o PB apresenta um quadro linguístico peculiar, como afirmam Leite e Callou (2010):

A diversidade que existe em qualquer ponto espelha uma pluralidade cultural e não se pode presumir para a expansão do português no Brasil de uma forma linguística única, pois a época em que se deu a colonização, a origem dos colonizadores e as consequências linguísticas de um contato heterogêneo, devem ser considerados (LEITE; CALLOU, 2010, p. 12).

Nesta tese, então, assumem-se os pressupostos da Dialectologia, amparando-se, como suporte, na Sociolinguística Variacionista, em que ambas concebem a língua como heterogênea e diversificada, que pode variar no tempo e no espaço, sob condicionamentos internos e externos. Logo, vislumbra-se um cenário linguístico brasileiro diverso, não somente pela pluralidade de normas, mas também por outras línguas que coexistem em seu amplo e vasto território, uma vez que:

Para compreendermos a questão é preciso trazer alguns dados: no Brasil de hoje são falados por volta de 215 idiomas. As nações indígenas do país falam cerca de 180 línguas (chamadas de autóctones), e as comunidades de descendentes de imigrantes outras 30 línguas (chamadas de línguas alóctones). Some-se a estas ainda as línguas de sinais, com destaque para LIBRAS, língua brasileira de sinais, e para línguas afro-brasileiras ainda usadas nos quase mil quilombos oficialmente reconhecidos no Brasil. Somos, portanto, um país de muitas línguas, plurilíngue (OLIVEIRA, 2009, p. 20).

Desse modo, ao olhar para a história linguística do PB, pode-se, certamente, encontrar pelos veios históricos explicações para tamanha pluralidade e variedades de normas, sendo, em muitos casos, espelhada pela convivência de diversas normas no mesmo espaço geográfico.

Essa diversidade de normas, de certa maneira, é reflexo dos processos de colonização, dos diversos movimentos de migração e imigração – internos e externos – do surgimento das primeiras cidades, bem como, a partir da chegada da família real, a criação de bancos e faculdades. Nesse sentido, observam-se que fatores geográficos – século e data de fundação – e sociais – surgimento do ensino universitário –, no que tange aos modos de utilização da língua, podem ser diferenciados, por meio desses dois aspectos elementares.

Assim, as capitais brasileiras¹⁷, quanto ao aspecto século e data de fundação, ao observar o eixo horizontal, podem ser divididas em três blocos, conforme ilustra o Quadro 7:

Quadro 7 – Capitais brasileiras quanto ao século e a data de fundação

SÉCULO (S)	CAPITAIS
XVI	Recife (12/03/1537); Salvador (29/03/1549); Vitória (08/09/1551); São Paulo (25/01/1554); Rio de Janeiro (01/03/1565); João Pessoa (05/08/1585); Natal (25/12/1599).
XVII	São Luís (08/09/1612); Belém (12/01/1616); Manaus (24/10/1669); Florianópolis (23/03/1673); Curitiba (29/03/1693).
XVIII	Cuiabá (08/04/1719); Fortaleza (13/04/1726); Porto Alegre (26/03/1772).
XIX	Maceió (05/12/1815); Teresina (16/08/1852); Aracaju (17/03/1855); Campo Grande (21/06/1872); Rio Branco (28/12/1882); Boa Vista (09/07/1890); Belo Horizonte (12/12/1897).
XX	Porto Velho (02/10/1907); Goiânia (24/10/1933); Macapá (04/02/1958).

Fonte: Adaptado de Isquierdo (2014b, p. 11).

Quanto ao eixo vertical, em especial, ao surgimento de faculdades/universidades, as capitais brasileiras podem ser divididas em dois blocos, conforme Quadro 8.

Quadro 8 – Capitais brasileiras quanto ao surgimento do ensino universitário

SÉCULO	CAPITAIS
XIX	Salvador (1808); Rio de Janeiro (1808); Recife (1827); São Paulo (1827); Porto Alegre (1895); Belo Horizonte (1898) ¹⁸ .
XX	Belém (1902); Fortaleza (1903); Manaus (1909); Boa Vista (1909); Curitiba (1913); Vitória (1930); Campo Grande (1930); Teresina (1931); Maceió (1931); João Pessoa (1934); Cuiabá (1934); Natal (1947); Goiânia (1948); Aracaju (1948); Florianópolis (1960); Rio Branco (1964); São Luís (1966); Macapá (1970); Porto Velho (1971).

Fonte: Elaborado pelo autor (2023).

Nesta tese, a fim de elucidar os dados sobre o PB falado nas décadas de 1970 e 1990 (Projeto NURC) e nos anos 2000 (Projeto ALiB), tais dados serão observados à luz desses fatos, haja vista o entendimento de que a história social, a história política e a história linguística, bem como as construções fomentadas pelo imaginário coletivo, de determinado

¹⁷ Informações mais detalhadas sobre a sociohistória das capitais encontram-se em Isquierdo (2014b).

¹⁸ Faculdade de Direito que foi transferida da cidade de Ouro Preto-MG.

povo, no que tange à língua, são elementos que fornecem pistas para elucidar aspectos referentes às escolhas e usos de formas em detrimento de outras.

Nessa direção, aspectos da sociohistória, tais como o século e a data de fundação das capitais – Brasil colônia e Brasil república –, os fluxos migratórios, os ciclos econômicos e o século e a data de fundação do ensino universitário, podem dar indícios de caminhos interpretativos e explicativos para a diversidade de falares que caracteriza o PB e, certamente, moldam as configurações de determinados fenômenos.

3.1.1 Trajetos sobre as hipóteses da formação do Português Brasileiro

Assim, são facilmente percebidas as faces da heterogeneidade sociocultural da variedade do português que se consolidou em solo brasileiro. Tal vertente heterogênea e diversificada que caracteriza o PB encontra, de certa forma, três hipóteses interpretativas que tentam explicar a configuração dessa variedade, a saber: a Transmissão Linguística Irregular (TLI) (LUCCHESI, 2001); a criouliização prévia (GAY, 2005); e, por fim, a deriva secular e a confluência de motivos (NARO; SCHERRE, 2007)¹⁹.

A importância das três propostas, no entanto, para uma situação linguística bastante complexa e polarizada, como a do Brasil, necessita de outros exemplos e mais confirmações das hipóteses elencadas, a fim de elucidar a configuração multilíngue do PB. Neste sentido, Lins (2009) afirma existirem lacunas consideráveis nas propostas interpretativas e sugere que:

o trabalho do pesquisador, pelo visto, deve ser o de perscrutação afinada com o requinte orquestral, de maneira a sair do módulo de estudos cromáticos para o de estudos diatônicos, ou seja, sair de uma escala semitonizada de espaço e tempo entre fatos lingüísticos registrados em períodos distintos da sócio-história brasileira, para montar uma escala capaz de reconstruir/esboçar numa ordem natural, portanto numa seqüência ordenada, os fatores lingüísticos e sócio-históricos do português brasileiro (LINS, 2009, p. 293).

Entende-se que, dos três modelos apresentados, embora se considerem as lacunas apontadas por Lins (2009), há de se levar em consideração as ideias apresentadas e defendidas por Lucchesi (2001), da Transmissão Linguística Irregular (TLI). No entanto, a TLI não será utilizada para interpretação dos dados devido à limitação do recorte observado, apenas capitais. Em trabalhos futuros, com a ampliação do olhar sobre o banco de dados do

¹⁹ Um estudo aprofundado sobre as hipóteses interpretativas pode ser encontrado em Lins (2009).

Projeto ALiB, por exemplo, atentando-se ao que foi catalogado nas localidades do interior, talvez seja possível notar a polarização sociolinguística.

Nesse contexto, a partir do cenário sociodialetal brasileiro, há a hipótese de polarização sociolinguística, proposta por Lucchesi (2001, 2002, 2004, 2015). O referido autor sistematiza a hipótese afirmando que há um conjunto de normas²⁰, que podem ser divididas em dois polos: a) a norma popular, caracterizada e definida pela maioria da população, geralmente, sem acesso aos bens culturais, pouco e nenhum grau de instrução escolar e com precárias condições socioeconômicas; e b) a norma culta é caracterizada pela pequena parcela da população, constituída pela fala de indivíduos que possuem formação escolar, fácil acesso aos bens culturais e sociais, além de boas condições socioeconômicas (LUCCHESI, 2002).

Ainda, na perspectiva da caracterização desse viés sociodialetal do PB, cabe trazer à tona que o cenário atual brasileiro deve ser observado a partir de três *continua*: “[...] o rural-urbano, o de oralidade-letramento e o de monitoração estilística.” (BORTONI-RICARDO, 2005, p. 40). Neste sentido, o *continuum* de urbanização é revelador de aspectos sociohistóricos da realidade linguística brasileira, visto que o país tinha sua economia rural até o século XX, além de permitir a análise dos aspectos socioecológicos do povo.

Destaca-se, também, o *continuum* de oralidade/letramento, possuidor de interfaces com o *continuum* anteriormente citado e com o de monitoração de estilos de usos da língua, haja vista que, ao ser inserido em uma cultura predominantemente oral ou letrada, os recursos comunicativos do falante serão, de certo modo, modificados, ou seja, as práticas sociais em que o indivíduo está inserido podem ser analisadas.

Por sua vez, no que tange ao *continuum* de monitoração estilística, observam-se o grau de atenção e de planejamento da fala no contexto da interação, decorrentes de elementos como: a acomodação do falante a seu interlocutor; o apoio no contexto, na produção do enunciado; o grau de complexidade cognitiva que envolve a produção linguística, bem como a familiaridade do falante com o evento comunicativo que está sendo desenvolvido por ele no momento da elocução (BORTONI-RICARDO, 2005).

Lucchesi (2015), a partir da observação das três *continua* – o rural-urbano, o de oralidade-letramento e o de monitoração estilística – associa o modelo proposto por Bortoni-Ricardo (2005) com o quadro plural e polarizado e conclui que:

²⁰ Embora a discussão sobre as normas e os adjetivos culta, popular, padrão, etc seja pertinente, entende-se que maiores detalhamentos podem ser encontrados na ampla literatura sobre o PB, tais como em: Bagno (2002); Lucchesi (2002); Faraco (2008) e Mattos e Silva (1995).

[...] têm-se, no ponto extremo da norma popular, as comunidades rurais mais isoladas do interior do país, cujos falantes, imersos na oralidade, exibiram um espectro de variação estilística bastante reduzido, já que é escasso também o seu contato com as demais variedades do português. No outro extremo da norma culta, estariam os falantes dos grandes centros urbanos, com alto grau de escolaridade, cujo extenso repertório linguístico amplia bastante o espectro de sua variação estilística. Entre um extremo e outro, vislumbra-se um *continuum* de níveis intermediários. À medida que se passa do polo extremo da norma popular para a norma culta, cresce progressivamente o grau de urbanização e letramento, bem como o espectro da variação estilística (LUCCHESI, 2015, p. 22).

A partir disso, aqui, sugere-se que o cenário sociodialetal brasileiro seja analisado com base nas perspectivas que levem em consideração um *continuum* de urbanização, atentando-se para os fatores que envolvam toda a sociohistória das localidades, os contextos de usos linguísticos, a história do falante e os aspectos sociais imbricados na manifestação linguística concreta dele.

Descortinar essa situação sociolinguística não é tarefa simples, mas, sim, complexa, pois há o envolvimento dos aspectos de diversas ordens, sobretudo as violências cultural e linguística – glotocídios – ocorridas no Brasil. Também, as diferentes maneiras em que se deu a ocupação do território, as diversas etnias e povos envolvidos, o processo de urbanização e, posteriormente, o desenvolvimento econômico, devem ser considerados, a fim de fornecer caminhos interpretativos para a língua portuguesa falada no Brasil, uma vez que:

Em função de uma realidade lingüística que abarca uma pluralidade de normas vernáculas, de normas cultas e – dadas as contradições de autor para autor – de normas subjetivas das gramáticas escolares, constatam-se vários pólos cultos e populares sobrepostos, em vários níveis, e, por isso mesmo, às vezes, de difícil reunião em um mesmo extremo. De fato, a complexidade sociolinguística do Brasil aproxima-se bem mais da imagem de uma rosa-dos-ventos que da de pólos (CALLOU et al., 2006, p. 262).

Ademais, acredita-se que, para a composição de um panorama mais próximo da realidade, independentemente das noções de polo e os pontos cardeais, as diferenças nos eixos vertical e horizontal não podem ser desconsideradas, haja vista as distintas formas linguísticas que são evidenciadas, quando observados esses pontos.

Tendo em vista que os fatos sociohistóricos e culturais desempenham importante relevância para o entendimento sobre os aspectos linguísticos, no que concerne aos pronomes, cabe recorrer à história ao latim, a fim de verificar possíveis correlações com o quadro pronominal do PB.

3.2 OS ATALHOS DOS PRONOMES

Como se constata, no latim clássico, os nomes possuíam seis casos – o *nominativo*, o *acusativo*, o *dativo*, o *genitivo*, o *vocativo* e o *ablativo* – a cada caso estava relacionada uma ou algumas funções sintáticas, tais como: a) *nominativo*: funções sintáticas subjetivas, como o sujeito e o predicativo do sujeito; b) *acusativo*: complementos verbais (objeto direto); c) *dativo*: representação do objeto indireto; d) *genitivo*: adjunto adnominal restritivo; e) *vocativo*: caracteriza uma evocação, apelo, chamado; f) *ablativo*: advérbios, que trazem noções de tempo, causa, modo, lugar... No entanto, algumas alterações ocorreram do latim clássico para o latim vulgar, como, por exemplo, a redução no número de declinações. O latim clássico, no que tange à estrutura, manteve-se inalterado praticamente, ao passo que o latim vulgar apresentou apenas três casos para declinações, quais sejam: o nominativo, o dativo e o acusativo.

A simplificação de casos, no latim, pode ser vista como o resultado da ação de diversos fatores. Tal língua, desde o início, apresentou uma tendência a reduzir o número dos casos, como se observa no singular. O *instrumental* e, posteriormente, o *locativo* confundiram-se com outros casos – *ablativo* e *genitivo* – e a parte final do *vocativo* – nos nomes masc. em *-us* – tendia a ser absorvida pelo *nominativo*; no plural, uma única forma servia para o *dativo* e o *ablativo*, para o *instrumental* e o *locativo* (VÄÄNÄNEN, 1968, p. 180).

Sendo assim, da variedade clássica para a variedade vulgar, ficou concretizada a redução da flexão casual latina de seis para três casos: o *nominativo*, o *acusativo* e o *dativo*. As funções sintáticas foram redistribuídas, de acordo com esses três casos: a) *nominativo*: Sujeito e Vocativo; b) *acusativo*: Objeto direto, Predicativo do Objeto, Complementos de duração, medida, tempo etc. e, por fim; c) *dativo*: Complemento de atribuição, interesse e fim; Complemento de posse. Porém, há autores que defendem que a redução foi bastante drástica, restando apenas dois casos – *nominativo* e *acusativo* – conforme Coutinho (1974 [1938], p. 226), por exemplo, para a 1ª declinação:

O processo de redução dos casos latinos originou-se na classe dos nomes, teve reflexo direto em outra classe gramatical, a dos pronomes. Portanto, pode ser visto que, a partir do exposto, os oito casos – no indo-europeu – passaram para seis – no latim clássico – e que, por sua vez, por processos de simplificação dos paradigmas de flexão, sobretudo neutralizações, restaram dois casos – latim vulgar. Neste sentido, a função sintática, antes

vista na forma composicional da palavra, por meio de morfemas que carregavam a função de atribuição de casos, passa a requerer outros meios para ser identificada.

Observa-se que, na variedade clássica da língua latina, a classe dos pronomes se delineava, apenas, em torno de duas pessoas do discurso com a função de sujeito, a que fala (*ego – nos*) e a que ouve (*tu – vos*). Conforme Almeida (1995, p. 13), esses pronomes distribuíam-se entre os seis casos latinos, a saber:

i. casos retos – nominativo (*ego, tu, nos e vos*); – vocativo (*tu e vos*); ii. casos oblíquos – genitivo (*mei, tui, nostrum / nostri, vestrum / vostri*); – dativo (*mihi, tibi, nobis e vobis*); – ablativo (*me, te, nobis e vobis*); – acusativo (*me, te, nos e vos*).

Desse modo, portanto, o quadro pronominal da variedade vulgar se transformou em relação à clássica, tanto pela diminuição da flexão casual quanto pelo aspecto, decorrente do aparecimento da terceira pessoa do discurso, tendo surgido a partir das formas demonstrativas – *ipse e ille*. Apenas *ille* consagra-se como forma de representação da terceira pessoa²¹. Segundo Maurer Jr. (1959, p. 106), a configuração do paradigma de pronomes do latim vulgar delineava-se da seguinte forma:

Quadro 9 – Paradigma dos pronomes do latim vulgar.

	1ª pessoa	
	Singular	Plural
Nominativo	<i>eo</i>	<i>nos</i>
Dativo	<i>mi (e mihi)</i>	<i>nos (e nobis)</i>
Acusativo	<i>me</i>	<i>nos</i>
	2ª pessoa	
	Singular	Plural
Nominativo	<i>tu</i>	<i>vos</i>
Dativo	<i>ti (e tibi)</i>	<i>vos (e vobis)</i>
Acusativo	<i>te</i>	<i>vos</i>

Fonte: Adaptado de Maurer Jr. (1959, p. 106).

Vale salientar que as formas pronominais latinas *ego > eo > eu*, diacronicamente, sofreram modificações oriundas dos processos fonológicos. Assim como as formas *mihi > mi > mim*. Ainda, nessa direção, Almeida (2011, p. 217) salienta que existem os dativos de referência – *para mim, para eles* – que, em português são utilizados para indicar a

²¹ Em Maurer Jr. (1959), é possível encontrar as características mais detalhadas sobre os pronomes de 3ª pessoa.

perspectiva da pessoa sob a qual a afirmação é feita, tal como “Vergílius óptimus poeta míhi est” – “Virgílio é para mim o melhor poeta”.

A partir dessas discussões, nota-se que a redução na flexão de casos da língua latina originou-se nos nomes e, posteriormente, a classe dos pronomes foi afetada, no latim vulgar. Tais modificações no quadro pronominal da variedade vulgar latina, de certo modo, estenderam-se aos demais romances.

3.2.1 Noções de pessoa e de caso: a natureza dos pronomes

Antes de adentrar nos aspectos peculiares aos pronomes de primeira pessoa do singular, vale mencionar que há uma grande discussão sobre a real natureza dos pronomes, incluindo a noção de pessoa, de caso e sobre um quadro pronominal que demonstre como se delinham os pronomes pessoais do PB. Nesse sentido, reflexões devem ser feitas, haja vista que:

[...] já algum tempo deixamos de viver no país do eu, tu, ele, nós, vós, eles, mas ainda é com estes trajes que as pessoas do discurso se apresentam aos desavisados. Que a norma gramatical dos manuais escolares não serve de espelho para “a língua como ela é” nossas crianças percebem sempre [...] Todos, enfim, falam um português mal-aprendido. Isso são coisas mais que sabidas, não chovamos no molhado. A questão é: que foi feito dos pronomes pessoais? (LOPES; CUNHA, 1994).

No que tange à conceituação, conforme a etimologia latina, pronome é uma palavra variável designada para substituir o nome e/ou acompanhá-lo. Então, antes de apresentar o mosaico pronominal latino, é importante destacar as noções referentes aos nomes nessa língua. Desse modo, observa-se que o latim é considerado uma língua concisa e sintética, não possuía artigo, tinha três gêneros – masculino, feminino e neutro – e, em relação ao número, apresentava dois – singular e plural.

Ao observar os nomes e pronomes, na perspectiva diacrônica, nota-se que, por exemplo, esta língua indo-europeia ocidental era considerada uma língua detentora de uma variedade flexional, uma vez que possuía os oito casos – nominativo, vocativo, genitivo, dativo, ablativo e acusativo – além do locativo e o instrumental (MARTINET, 1997, p. 237).

Ravizza (1940) define o pronome como a palavra que faz o papel de um nome, além de concordar com ele em gênero e número; apresenta seis espécies de pronomes: pessoal, possessivo, demonstrativo, relativo, interrogativo e indefinido, também, para os pessoais, o autor apresenta cinco casos; e destaca que a forma *mihi*, o dativo da primeira pessoa, na poesia, é encontrada como *mi*, a forma contracta.

Em Grimal et al., (1986), não há definição para os pronomes, a seção destinada para eles é iniciada com a apresentação das 1ª e 2ª pessoas, no singular e plural, divididos em cinco casos latinos. Posteriormente, há reflexões sobre a terceira pessoa. O que chamou a atenção durante a consulta à obra foi o item “Particularidades”, pois há uma observação sobre o pronome sujeito que, conforme os referidos autores, “não é expresso, salvo para insistir sobre a pessoa. Tu rides, ego fleo. Tu ris, eu choro. Ego laboro. Sou eu que trabalho” (GRIMAL et al., 1986, p. 32). Com isso, infere-se que os autores demonstram a função contrastiva dos pronomes.

Almeida (1997) afirma que o pronome se caracteriza por ser a palavra que substitui ou pode substituir o substantivo, além de apresentar diversas espécies, e cita os pessoais – ao mesmo tempo, substitui o nome de um ser, põe esse nome em relação com a pessoa gramatical – os do caso reto exercem função de sujeito. O referido autor apresenta um quadro com os pronomes pessoais latinos, tal qual os outros autores, conforme Quadro 10, nessa seção. Sobre o sujeito, encontra-se a definição de pessoa ou coisa que pratica a ação, ou seja, agente da ação verbal.

Quadro 10 – Paradigma de flexão de casos dos nomes da 1ª declinação do latim

Latim Clássico			Latim Vulgar		
	Singular	Plural		Singular	Plural
Nominativo	<i>luna</i>	<i>lunae</i>	Nominativo	<i>luna</i>	<i>lune</i>
Vocativo	<i>luna</i>	<i>lunae</i>	–	–	–
Genitivo	<i>lunae</i>	<i>lunarum</i>	–	–	–
Dativo	<i>lunae</i>	<i>lunis</i>	–	–	–
Ablativo	<i>luna</i>	<i>lunis</i>	–	–	–
Acusativo	<i>lunam</i>	<i>lunas</i>	Acusativo	<i>luna</i>	<i>lunas</i>

Fonte: Adaptado de Coutinho (1974 [1938], p. 226).

Historicamente, no que se refere à noção de “pessoa”, a organização pronominal gira em torno de três pessoas, a partir da variedade do latim vulgar até as línguas românicas. As formas que representam a primeira pessoa no PB, *eu* e *nós*, a pessoa que fala, as segundas, *tu* e *vós*, representam a pessoa com quem se fala, e as terceiras, *ele/ela* – *eles/elas*, de quem se fala (BECHARA, 2010).

Mattos e Silva (2001) apresenta um quadro pronominal do português arcaico subdividido em três pessoas, os primeiros, os pronomes de primeira e segunda pessoas fazem referência ao emissor e ao receptor; por sua vez, os pronomes de terceira pessoa são

anafóricos, pois exercem a função de substituir sintagmas nominais já referidos. Por fim, os pronomes pessoais são sintetizados no Quadro 11, com as alomorfas decorrentes do período arcaico.

Quadro 11 – Distribuição dos pronomes pessoais do português arcaico

Referência	Distribuição ²²	Gênero	Número	
			Singular	Plural
+E (1ª pessoa)	1	-	eu	nós*
	2	-	me ~ mi m' mh'	nos
	3	-	(Prep) min comigo ~migo	(Prep) nós conosco ~conosco
+ R (2ª pessoa)	1	-	tu	vós
	2	-	te ~ti t' ch'	vos
	3	-	(Prep) ti contigo ~tigo	(Prep) vós convosco ~vosco
-E-R (3ª pessoa)	1	masc. fem.	ele ~el ela	eles elas
			(h) omen ome, ome	
	2	masc. fem.	o ~ uos ~ us lolos nonos a ~laas ~ las nanas lhe ~hilhes ~lhis	
			se ~si s' xi' x'	
	3	masc. fem.	(Prep) ele ~el ela	eles elas (Prep) si consigo

Fonte: Mattos e Silva (2001, p. 99).

Ao observar os pronomes do período arcaico, percebe-se que algumas formas sofreram alteração, como os átonos de segunda pessoa.

Câmara Jr. (1978) afirma que os pronomes se definem pela noção gramatical de pessoa em que eles consistem. O referido autor sistematiza os pronomes portugueses, alertando para o fato de o quadro pertencer a um campo teórico, uma vez que “[...] em

²² Os números 1, 2 e 3, conforme proposição da autora, referem-se às pessoas do discurso.

nenhuma região da língua portuguesa ele se realiza exatamente”. Tal quadro é sumarizado assim: “P1: eu; me; mim; comigo; P2: tu; te; ti; contigo; P4: nós.....; conosco; P3: ele (a); o (a) lhe; P6: +/S/; + /S/.” (CÂMARA JR., 1978, p. 119).

Para ele, a P1 e a P4 são as formas *eu* e *nós*, singular e plural, respectivamente, responsáveis por quem fala. As P2 e P5 são as formas *tu* e *vós*, são os ouvintes, com quem se fala. E tudo o que fica fora do eixo falante e ouvinte são as formas P3 e P6, *ele/ela*, *eles/elas*. As referidas formas pronominais do caso reto podem ser utilizadas em frase isolada ou como sujeito de um verbo. Por sua vez, as formas oblíquas podem ser do tipo adverbial: “[...] me, nos, te, vos, o, a, ou lhe; os, as, ou lhes...”, são formas dependentes e podem ser utilizadas junto ao verbo, com função de complemento. E as formas oblíquas também apresentam outra série, são partículas tônicas que funcionam sob regência de uma preposição e estão restritas às formas P1 e P2²³.

No que diz respeito às noções de caso, Câmara Jr. (1978, p. 96) afirma que a distribuição dos pronomes pessoais difere muito da distribuição de casos latinos em pelo menos três dimensões: na formal, na funcional e na semântica. Além disso, conforme o autor, é “[...] uma afirmação inexata e confunde sistemas de formas inteiramente diversos...”, atribuir às formas pronominais da língua portuguesa a incumbência de salvaguardar mais fielmente os resquícios dos casos pronominais do latim (CÂMARA JR., 1978, p. 96).

Em Monteiro (1994), pode-se observar, a partir de um estudo com dados da norma culta, que o sistema de caso do português brasileiro encontra-se em processo de desestruturação e rearranjos, com indícios de ser extinto, no que se refere aos pronomes pessoais. Logo, “[...] é importante insistir na introdução de novos pronomes, cujas formas tanto servem para expressar o sujeito como qualquer outra função.” (MONTEIRO, 1994, p. 72).

Benveniste (2005), quando se propõe a analisar a classe dos pronomes, afirma que eles possuem naturezas distintas, havendo duas pessoas para o discurso e uma referência zero, que fica fora da relação das duas primeiras pessoas, para ele: “[...] os pronomes não constituem uma classe unitária, mas espécies diferentes segundo o modo de linguagem do qual são os signos...” (BENVENISTE, 2005, p. 281). Nessa perspectiva, a instância comunicativa é realizada por meio do eixo falante-ouvinte, em que as primeiras pessoas se dirigem às segundas pessoas, e a terceira pessoa, sempre, é utilizada para se referir a elementos que estão de fora da instância comunicativa. Com isso, os postulados propostos

²³ O autor não considera a forma “para si”.

pelo referido autor transformaram as abordagens teóricas sobre os aspectos pronominais, originando diversos trabalhos nas mais variadas línguas.

Perini (2005), na obra *Gramática Descritiva do português*, contrapondo a definição de pronome encontrada na gramática tradicional, afirma que “A definição de ‘pronome’ dada nas gramáticas, mesmo quando aplicável, nunca consegue delimitar exatamente o grupo de itens pretendido” (PERINI, 2005, p. 330). Também informa que há obscuridade nas palavras “substitui” e “acompanha”, afirmando que a classe dos pronomes não apresenta coerência sintática e que ela precisa desaparecer. Com base nesses argumentos, o referido autor propõe uma nova classificação, a saber:

Em primeiro lugar, há um grupo de itens que se encaixam na classe dos **substantivos tipo 1** [...] se definem pelos traços [+CP, -Mod, -Pv, -T, SN, +NSN, -Int—, -PN, -.ssimo]; ou, para abreviar, são itens que só podem ocorrer como complementos do predicado ou núcleos de um SN: [+CP, +NSN]. Os itens que compõem essa subclasse dos substantivos são todos chamados ‘pronomes’ na gramática tradicional: eu, ele, nós, eles, etc [...] Depois, há ‘pronomes’ que veem ser classificados como **substantivos tipo 2**, ao lado de muitos itens tradicionalmente chamados “substantivos”. Trata-se dos itens alguém, algo, tudo, que [...] diferem dos itens precedentes por poderem ocorrer no SN acompanhados por outro termo, como em esse alguém, um alguém, alguém carinhoso; tudo de bom; algo misterioso... (PERINI, 2005, p. 331).

Por fim, conforme Perini (2005), há uma síntese da nova configuração dos pronomes. Além dos anteriormente citados, substantivos dos tipos 1 e 2, existem os relativos e interrogativos: que, qual e quem; [+PDet]: os que exercem função de pré-determinante – todos, ambos; [+Det]: os determinantes – o, um, este, esse, aquele, alguns etc.; [+Poss]: os que expressam ideia de posse – meu, seu, nosso etc.; [+Qf]: os quantificadores, por exemplo: muitos, vários, único, terceiro e outros; [+Num]: aqueles que podem ser numerados, tais como: outro, dois. Em relação aos aspectos morfossintáticos, o autor define o sujeito como o termo da oração que estabelece uma relação de concordância com o NdP e atesta existirem em português dois tipos de orações reduzidas, as de infinitivo e gerúndio, cujo NdP não se flexiona em número e pessoa, já as do particípio são sintagmas adjetivos muito regulares.

Na *Novíssima Gramática da Língua Portuguesa*, Cegalla (2008) apresenta os pronomes como palavras que são utilizadas indicando a pessoa do discurso, para substituir ou determinar os substantivos. Subdivide-os em seis, pessoais, possessivos, indefinidos, demonstrativos, relativos e interrogativos. Os pessoais são os substitutos dos substantivos e representam as pessoas gramaticais, estão divididos em retos – que funcionam, em regra,

como sujeito da oração – e oblíquos – funcionam como objetos ou complementos, que, por sua vez, estão agrupados quanto à acentuação em tônicos e átonos.

Em Bechara (2009), há direcionamentos para o entendimento das pessoas do discurso, as determinadas, 1ª e 2ª pessoas, e a indeterminada, a 3ª pessoa. No que se referem à semântica, os pronomes vão indicar dêixis. Desse modo: “Os pronomes pessoais designam as duas pessoas do discurso e a não pessoa (não *eu*, não *tu*), considerada, pela tradição, a 3ª pessoa...” (BECHARA, 2009, p. 139). Dentre os pronomes pessoais, existem as formas retas, que funcionam como sujeito, e as formas oblíquas, tônicas e átonas, que exercem função de complemento.

Bagno (2011) recusa a noção de classe dada aos pronomes e afirma que pronome designa função. O referido autor adere às postulações feitas por Benveniste (1902-1976²⁴), e afirma que:

Apesar de tradicionalmente chamarmos as palavras *eu/ tu/ você/ nós/ a gente/ vocês* de ‘pronomes’, elas não se encaixam naquela definição, também tradicional em nossa pedagogia de língua, de ‘palavra que substitui o nome’ – uma definição, aliás, que seria bem melhor descartar para sempre. Afinal, se substituirmos a palavra *eu* pelo nome da pessoa que ela indicia, o enunciado deixa de ser de primeira pessoa e se torna imediatamente de não-pessoa (BAGNO, 2011, p. 457).

Para os pronomes pessoais, Bagno (2011) utiliza índices de pessoas, asseverando que esses pronomes são, de forma eminente, dêiticos. No entanto, ainda conforme tais postulados, somente as formas ditas de 1º e 2º pessoas devem ser vistas como elementos dêiticos, pois são responsáveis pelas indicações das pessoas do discurso. As formas de 3º pessoa são anáforas e, por isso, devem ser inclusas em outra classe de palavras.

Observam-se, após esse breve cotejo teórico, as divergências sobre a classificação da classe/função dos pronomes. Com isso, notam-se algumas convergências, quanto à real natureza e à definição, sobretudo entre as proposições de Benveniste (1902-1976) e Bagno (2011). Além dessa, percebe-se que as ideias propostas por Monteiro (1994) dialogam com as prescritas por Câmara Jr. (1978), uma vez que ambos não discordam das noções de caso atribuídas às formas pronominais pessoais do PB, embora haja ressalvas quanto à preservação dos casos latinos, com a tendência de extinção na variedade falada no Brasil.

²⁴ Ressalta-se que essas observações foram retiradas da obra de Bagno (2011).

3.3 SÍNTESE DA SEÇÃO

Esta seção foi elaborada com vistas a apresentar fatos sociais e históricos que podem, certamente, exercer influência sobre os fenômenos linguísticos. Para tal, na subseção 3.1, buscou-se traçar um panorama do latim ao português, destacando diversos fatos, tais como as reduções na morfologia flexional latinas, inicialmente, na classe dos nomes e, posteriormente, nos pronomes; os elementos de constituição e formação do povo brasileiro, tais como as navegações lusitanas, o contato com os indígenas, o tráfico de negros escravizados, devido aos diversos ciclos econômicos, propiciando muitos movimentos de deslocamentos internos e externos. Para além disso, ainda sob o viés histórico, o século, a data de fundação e surgimento do ensino universitário nas capitais podem ser favorecedores para o entendimento das diferenças e semelhanças linguísticas em solo brasileiro. Buscou-se, também, trazer à tona uma breve discussão sobre as hipóteses interpretativas sobre a formação do PB, dando ênfase para a TLI.

Ao caminhar pelo viés da história linguística e social no PB, notam-se aspectos que são importantes para o entendimento sincrônico de determinados fenômenos, sobretudo, espelhados pela morfossintaxe, com resquícios da variedade vulgar latina, bem como do multilinguismo generalizado ocorrido no Brasil. Fatores que merecem destaque, nesse contexto colonial, também são os diversos ciclos econômicos, os fluxos intensos das migrações, o tráfico de negros escravizados e as diversas ações implementadas com a chegada da família real, além do decreto pombalino.

Na subseção seguinte, 3.2, são descritos aspectos sobre a natureza dos pronomes, de forma geral, e como a classe dos pronomes tem sido exposta em algumas gramáticas, permitindo, desse modo, vislumbrar aspectos não consensuais sobre conceito, classificação e especificação dos pronomes.

Observa-se que há muitas direções sobre a classe dos pronomes. A diversidade de aspectos envolvidos e modos de abordar o assunto é importante, principalmente por se tratar de uma classe aberta e totalmente flexível.

Em seguida, após essa discussão, na seção 4, debruça-se a atenção sobre as diversas perspectivas teóricas sobre o fenômeno ora estudado na tese, com fins de apresentar uma ampla e atual revisão de literatura.

4 SENDAS SOBRE A ESTRUTURA “PARA/PRA + EU/MIM/ + VERBO INFINITIVO”

E a despeito da repugnância dos nossos letrados em considerar dialeto o português falado no Brasil, a despeito da classificação de simples “brasileirismos” dada aos nossos modismos dialetais, a língua portuguesa tal como é falada pelo povo, e tal como também é falada na intimidade, despreocupadamente, pelas pessoas cultas, vai impondo suas formas especiais, já forçou as portas da literatura, e está obrigando os gramáticos a tomarem conhecimento de sua presença. Já conquistou, enfim, o seu lugar ao sol.

(MARROQUIM, 1934, p. 18).

De antemão, embora haja nesta subseção uma revisão ampla, sob diversas abordagens teóricas sobre o fenômeno ora estudado e analisado nesta tese, vale frisar que esta pesquisa aborda dados orais, e que o tratamento e as análises seguirão as orientações teóricas da Dialetoлогия, com um suporte da Sociolinguística Quantitativa.

Nessa seção, portanto, pretende-se apresentar, em forma de revisão de literatura, estudos que abordaram os pronomes *eu* e *mim*, em posição de sujeito. Intenta-se construir um panorama sobre os estudos que investigaram as particularidades da primeira pessoa do singular, sobretudo, a do caso reto e a do caso oblíquo tônico. Desse modo, há um levantamento das pesquisas que antecederam a proposta aqui descrita. Tais pesquisas, de certa maneira, ajudaram a observação mais apurada dos dados coletados nos dois *corpora*, a partir das amostras de falas do Projeto NURC e do Projeto ALiB, alvos de análise dessa tese.

4.1 ESTUDOS SOBRE O PORTUGUÊS BRASILEIRO

As pesquisas linguísticas sobre o PB, principalmente as que observam os aspectos linguísticos sob uma perspectiva de um tripé elementar, a saber, língua – cultura – sociedade, assinalam que, na língua falada, há uma necessidade de repensar a organização pronominal, no que tange à primeira pessoa do singular, pois observam-se certas construções que são feitas indistintamente *para/prá + eu/mim + infinitivo* – exemplos (3) e (4) – em contextos de fala e escrita espontâneas e mesmo na fala culta.

(3)Esse livro é **para eu ler**.

(4)Esse livro é **para mim ler**.

De modo geral, a construção (04) recebe uma avaliação negativa, muitas vezes, é considerada desprestigiada e sofre estigmas sociais em algumas variedades do PB. Assim

sendo, cabe trazer à tona estudos para o entendimento das peculiaridades das formas pronominais do PB e outras variedades do português, sobretudo os que se detiveram a analisar o quadro dos pronomes pessoais na primeira pessoa do singular, em especial, a estrutura foco de análise nesta tese.

4.1.1 Perspectiva das Gramáticas

Compreende-se, a esse respeito, que o que se encontra nas gramáticas são representações de modelos de língua. Em outras palavras, as concepções de linguagem estão subjacentes às formas de abordagem vistas nas gramáticas. Nesse sentido, acredita-se que as gramáticas refletem as concepções de língua e linguagem da orientação teórica seguida pelo autor da obra, por isso as diferentes gramáticas e suas tipologias vão espelhar os fenômenos linguísticos de maneiras diversas.

Santos (2017) analisou os pronomes – conceito, classificação, tipologia, topologia e uso – em sete gramáticas brasileiras do século XX. O estudo também contemplou a Historiografia da Linguística brasileira, além de uma biografia cuidadosa sobre a vida dos autores das gramáticas examinadas. Ao concluir as análises empreendidas no estudo, a autora afirma que, em relação aos pronomes pessoais, seis compêndios gramaticais tratam da combinação dos pronomes átonos. Ao abordar as postulações feitas por Cunha (2001), Santos (2017) esclarece que:

O gramático esclarece que, do cruzamento de duas construções corretas. ‘Isto não é trabalho *para eu fazer*’ e ‘Isto não é trabalho *para mim*’, surgiu uma terceira que, aparentemente, desconhecida em Portugal, mas generalizada na língua familiar no português do Brasil e combatida pelos gramáticos e professores de língua materna: ‘Isto não é trabalho *para mim fazer*’ (2001, p. 298). Como se pode observar, nessa nova construção, o pronome oblíquo é sujeito do verbo no infinitivo (SANTOS, 2017, p. 170).

Além disso, no entanto, cita Pereira (1918, p. 201): “Há casos, porém, em que o sujeito do verbo no infinitivo é representado por um pronome oblíquo. Isso ocorre com verbos fazer, deixar, ver, ouvir e mandar [...]”.

Nesta seção, os aspectos pertinentes à representação do fenômeno estudado na tese serão evidenciados por meio de nove gramáticas, tais como: a gramática histórica (três), a gramática normativa/tradicional de língua portuguesa (uma), a gramática descritiva (duas), e, por fim, a pedagógica/escolar (três). Foram consultadas gramáticas com tipologias diversas, publicadas nos séculos XX e XXI, conforme o Quadro 12.

Quadro 12 – Gramáticas de língua portuguesa dos séculos XX e XXI

Obra	Autor (es)	Ano de publicação	Século	Tipo
Compêndio de Gramática Histórica Portuguesa	Dr. José Joaquim Nunes	1945	XX	Histórica
Gramática Histórica da Língua Portuguesa	M. Said Ali	1964	XX	Histórica
Gramática Histórica	Ismael de Lima Coutinho	1976	XX	Histórica
Gramática Normativa da Língua Portuguesa	Rocha Lima	2018	XXI	Normativa/tradicional
Gramática de Usos do Português	M ^a Helena de Moura Neves	2000	XX	Descritiva
Nova Gramática do Português Brasileiro	Ataliba T. de Castilho	2016	XXI	Descritiva
Gramática Escolar da Língua Portuguesa	Evanildo Bechara	2010	XXI	Pedagógica/escolar
Gramática Pedagógica do Português Brasileiro	Marcos Bagno	2011	XXI	Pedagógica/escolar
Gramática Reflexiva, texto, semântica e interação	William Cereja; Thereza Cochar	2013	XXI	Pedagógica/escolar

Fonte: Elaborado pelo autor (2023).

É necessário, no entanto, apontar que não há intenção de classificá-las ou adjetivá-las, mas demonstrar as diversas abordagens desde as funções desempenhadas pelos pronomes pessoais até a estrutura morfossintática *para/prá + eu/mim + infinitivo*, a fim de melhor compreender as alternâncias entre os pronomes na posição de sujeito, quando precedidos pela preposição *para*.

No *Compêndio de Gramática Histórica Portuguesa*, de Nunes (1945), encontram-se distinções sobre as formas átonas e tônicas. Os pronomes *eu, nós, tu, vós, ele e eles* correspondem às funções nominativas de sujeito, ou de complemento, se acompanhadas de

preposição. Por sua vez, os pronomes correspondentes ao caso latino *acusativo*, *me*, *nos*, *te* e *vos*, desempenham a função sintática de complementos diretos e indiretos. Os complementos indiretos, *dativos*, são representados pelas formas pronominais: *me*, *te*, *mim*, *ti* e *lhe*. No entanto, tais formas, para além de complementos indiretos:

[...] podem ainda exprimir outras relações **as formas tônicas mim e ti, quando acompanhadas por preposição que não seja com, facto que se observa desde os mais antigos monumentos escritos, e até mesmo desempenhar função de sujeito, mas este último emprego parece quase exclusivo da língua popular**, a ajuizar da sua ocorrência em Gil Vicente em expressões como estas: *mas casemos eu e ti*, vol. I, pág. 33: *ora vamos eu e ti*, id. pág. 56, e no povo, que ainda diz, por exemplo, *eu sou mais velho ca ti*; *tu és mais novo ca mim* (NUNES, 1945, p. 240, grifo nosso).

Desse modo, os pronomes *dativos* designados para a função de complemento, em textos antigos, podem ser encontrados com outras funções, sobretudo com a função canonicamente caracterizada pelos pronomes *nominativos*.

Said Ali (1964) reflete sobre a falta de definição específica para designar o pronome, mas afirma que todo pronome é, em sua essência, um substantivo e pode ser chamado de pronome-substantivo, ou um adjetivo, pronome-adjetivo. Em relação aos pronomes pessoais, são utilizados como pronomes absolutos, têm singular e plural. Os pronomes de 1ª pessoa caracterizam a pessoa que fala, os de 2ª pessoa são aqueles representativos para a pessoa a quem era dirigido o discurso, e existem os de 3ª. Sobre as formas oblíquas, existem duas formas: átonas – não podem ser regidas de preposição – tais como, *me*, *te*, *nos*, *vos*, *lhe*, *lhes*, *o*, *a*, *os*, *as*, *se*; por seu turno, a tônica – são sempre dependentes de preposição – por exemplo, *mim* (outrora *mi*), *ti*, *nós*, *vós*, *éle*, *éla*, *éles*, *élas* e o *si*. Anteposto pela preposição *a*, essas últimas formas podem ser utilizadas para expressar o objetivo direto enfático. Ademais, nota-se uma mobilidade ainda maior dos pronomes tônicos, pois:

O emprego curioso das formas acusativas em lugar do nominativo repugna a linguagem culta de hoje. Já não era assim no falar antigo em certas frases comparativas [...] **Em lugar de como eu dizem ainda hoje na Beira como mim**. Muito vulgar é em Portugal a frase: “Se eu a ti fôsse” ou “Se eu fôsse a ti”. Em Gil Vicente 3, 325, lê-se A grandeza de misericórdia e largueza que tu es e ella he ti...’ (SAID ALI, 1964, p. 95) [grifo nosso].

Em Coutinho (1976), há apontamentos que dividem os pronomes em: pessoais, demonstrativos, possessivos, interrogativos e indefinidos. No que se refere aos pronomes pessoais, o referido autor afirma que são eles que mais guardam os vestígios do latim, são os mais utilizados na variedade vulgar que na clássica e são separados em tônicos – acentuados – *eu*, *mim*, *tu*, *ti*, *nós*, *vós*, e os átonos – sem acento – *me*, *te*, *se*, *lhe*, *o*, *a*, *nos*, *vos*, *lhes*, *os*, *as*.

Para os pronomes *eu* e *mim*, conforme Coutinho (1976, p. 253), tem-se:

...eo (nom.) por ego > eu. Eo aparece em textos latinos do século VI, segundo Meyer-Lübke. Mi (dat.) por mihi > mim. A nasalização do -i foi provocada pela presença da nasal inicial. Apesar de ser *mim* do fim do século XV, em Camões ainda se encontra *mi*: “Ouve os danos de *mi*” (Lusíadas).

Em Lima (2018), encontra-se um capítulo nomeado como “Emprego do pronome”, e na seção das formas retas ou subjetivas – *eu, tu /você, ele/ ela, nós, vós/vocês, eles/ elas* – há referência ao emprego para designar o sujeito. Na seção das formas oblíquas, observa-se a bipartição entre as formas objetivas diretas – *me, te, o, a, nos, você, o, a, os, as, vos, se* – e as formas objetivas indiretas, que são átonas – *me, te, lhe (a ele, a ela), nos lhe (a você) lhes (a eles, a elas), vos, se, lhes*; ou são tônicas – sempre regidas de preposição – *mim, ti/ você, nós, vós/vocês*. Há, também, uma observação: “Na linguagem arcaica, era frequente o uso das formas tônicas do objetivo indireto em função objetiva direta [...] Na evolução da língua foi-se abandonando o emprego das formas tônicas como objeto direto...” (LIMA, 2018, p. 388). Sobre o emprego enfático, existem exemplos nos quinhentistas e seiscentistas dos pronomes tônicos assumindo função de pronomes retos, conforme assinala Lima (2018).

Neves (2000) define os pronomes pessoais como elementos que possuem uma natureza fórica – com duas funções basilares: interacional e textual – pois esses têm uma capacidade de fazer referência pessoal, e, em especial, os do caso reto fazem referências às pessoas gramaticais no singular e no plural, são tônicas e podem ocorrer como sujeito de verbos em forma finita. A partir da sua natureza referenciadora, exercem outra função na oração de explicitar a natureza temática do referente.

Além disso, há considerações sobre a existência de formas átonas e tônicas e que as formas pronominais oblíquas tônicas *mim* e *ti*, regidas de preposição, são restritas às funções completivas. A norma gramatical estabelece que as formas pronominais *eu* e *tu* não podem ser regidas por preposição e só podem exercer a função de sujeito da oração. Entretanto, é perceptível na linguagem popular, literária e jornalística, construções com esses pronomes precedidos da preposição *entre*, em algumas orações (NEVES, 2000, p. 456).

A *Nova Gramática do Português Brasileiro*, de Ataliba T. de Castilho, publicada em 2016, em sua primeira edição, traz reflexões sobre o estatuto categorial dos pronomes. Segundo proposta do referido autor, faz-se necessária uma abordagem que leve em consideração as propriedades discursivas, semânticas, gramaticais e de gramaticalização, pois:

Do ponto de vista semântico-discursivo, os pronomes (1) representam as pessoas do discurso, pelo caminho da dêixis, (2) permitem a retomada ou antecipação de

participantes, pelo caminho da foricidade (anáfora e catáfora). Do ponto de vista gramatical, essa classe exibe as propriedades morfológicas de (i) caso; (ii) pessoa e número; (iii) gênero. Morfemas afixais e lexemas distintos expressam essas propriedades. Quanto ao caso, embora o português seja uma língua de caso abstrato, uma subclasse dos pronomes, a dos pessoais, preservou a distinção de casos herdada do latim vulgar, marcando-a através de seus lexemas. É o caso dos itens nominativos (eu, tu, ele, nós), o acusativo (o), os acusativos-dativos (me, te, se, nos) e o dativo (mim, ti, si, lhe). Essa marcação de caso particulariza os pronomes pessoais dentre os demais... (CASTILHO, 2016, p. 474-475).

Nesse sentido, na seção de descrição dos núcleos pronominais, o referido autor afirma que os pronomes pessoais são bem suscetíveis a mudanças, principalmente no que tange à língua falada, e apresenta um quadro ilustrativo da configuração dos pronomes pessoais no PB atual, a saber:

Quadro 13 – Os pronomes pessoais do PB propostos por Castilho (2016)

Pronomes pessoais no português brasileiro				
	PB FORMAL		PB INFORMAL	
	Sujeito	Complemento	Sujeito	Complemento
1sg	eu	me, mim, comigo	eu, a gente	eu, me, mim, prep + eu, mim
2sg	tu, você, o (a) senhor (a)	te, ti, contigo, prep. / o (a) senhor (a)	você /ocê/tu	você/ocê/te, ti, prep. + você/ocê (= doce, cocê)
3sg	ele/ela	o/a/lhe, se, si, consigo	ele/ela	ele, ela, lhe, prep + ele, ela
1pl	nós/ a gente	nos/nós/ a gente	nos/nós/ a gente	conosco/nós/ a gente
2pl	vocês	vocês	vocês	vocês
3pl	eles/ elas	eles/elas/se	eles/elas/lhe	eles/elas

Fonte: Castilho (2016, p. 477).

Segundo o referido autor,

Os pronomes pessoais são bastante suscetíveis a mudanças. Estudos recentes têm apontado para sua reorganização no PB, sobretudo na modalidade falada, com fortes consequências na estrutura sintática da língua. A centralidade desses pronomes no sistema das línguas explica por que a reorganização do quadro dos pronomes repercute nos demais pronomes, na morfologia verbal, na concordância verbal e na estrutura funcional da sentença... (CASTILHO, 2016, p. 477).

Com isso, conforme o Quadro 13, há formas que coexistem tanto no PB formal quanto no informal desempenhando a mesma função sintática. No entanto, constata-se também a

existência de formas peculiares do PB informal, como: para a primeira pessoa do singular, a forma nominativa *eu* aparece na função de complemento, quando precedida pela preposição *para*.

Bechara (2010), referindo-se ao aspecto semântico, afirma que os pronomes são considerados como dêixis – anáfora e catáfora – gestos verbais, possuem duas pessoas determinadas no discurso, *eu* e *tu*, e a terceira pessoa, indeterminada, faz remissão a outra coisa ou pessoa que está fora da relação de interação entre os participantes. Passa-se, então, à análise da configuração do quadro pronominal do PB. Na perspectiva do referido autor, as formas pronominais do caso reto são responsáveis pelas funções de sujeito; por sua vez, as formas pronominais oblíquas respondem como adjuntos e argumentos verbais. As últimas dividem-se em átonas – complementam o sentido do verbo sem auxílio de preposição – e as tônicas – por meio de preposição são ligadas ao verbo, como pode ser visto no Quadro 14.

Quadro 14 – Os pronomes pessoais, conforme a visão da gramática tradicional

Pronomes pessoais:		Retos	Oblíquos átonos (sem prep.)	Oblíquos tônicos (com prep.)
Singular	1ª pessoa	eu	me	mim
	2ª pessoa	tu	te	ti
	3ª pessoa	ele, ela	lhe, o, a, se	ele, ela, si
Plural	1ª pessoa	nós	nos	nós
	2ª pessoa	vós	nos	vós
	3ª pessoa	eles, elas	lhes, os, as, se	eles, elas, si

Fonte: Bechara (2010, p. 132).

Importante ressaltar que o PB atual, segundo observado no Quadro 14, distribui-se em dois casos, reto e oblíquo, no que tange às formas pronominais pessoais. Nota-se um equívoco, haja vista a reorganização pronominal feita no PB, sobretudo quanto à língua falada. Por exemplo, tal quadro ignora o desuso da segunda pessoa do plural, *vós*, a competição entre as formas *nós* e *a gente*, além da inserção do pronome *ocê* na segunda pessoa do singular, em detrimento de *tu*, em grande parte do território brasileiro²⁵.

Na *Gramática Pedagógica do Português Brasileiro*, de Bagno (2011), existem hipóteses para a utilização do pronome *mim* em posição de sujeito, “para mim + infinitivo”.

²⁵ Nas perspectivas da Dialetoлогия e da Sociolinguística, vários pesquisadores brasileiros têm ratificado o predomínio da forma *ocê*, dentre eles, destacam-se: Deus (2009); Alves (2010); Cardoso *et al.* (2014b), entre outros.

Para Bagno (2011), essa forma recebe avaliação negativa por parte da comunidade, embora venha ganhando, cada vez mais, amplitude entre os falantes dos grandes centros urbanos e que possuem escolarização, por exemplo, em São Paulo, na fala de professores, advogados, médicos, engenheiros e empresários. Também, há direcionamentos que indicam que a forma ocorre devido ao cruzamento de duas construções que se fundem em uma única, sintetizando a sintaxe, a semântica e a pragmática: a) *O ilustrador trouxe uns desenhos pra mim.* b) *O ilustrador trouxe uns desenhos para eu examinar.* c) a) + b) = *O ilustrador trouxe uns desenhos para mim examinar.* (BAGNO, 2011, p. 730) Tal ilustração apresentada pelo referido autor se aproxima bastante da abordagem explicativa dada por Cunha (2001), anteriormente mencionada.

Segundo Bagno (2011, p. 730) “A construção resultante do cruzamento permite dizer duas coisas de uma vez só, representando um ganho funcional, uma comunicação mais eficiente...”. Outra forma de explicar o fenômeno é por meio da braquilogia, figura de linguagem que remete ao termo grego *brakhys*, que significa breve, curto, reduzido, ou seja, formas que sofreram braquilogia gramaticalizam-se em estruturas fixas, com isso, uma ou as duas formas que se cruzaram tendem a desaparecer da linguagem corrente. Por fim, Bagno (2011) apresenta outra hipótese explicativa sobre o fenômeno, afirmando que o favorecimento do *mim* é por causa da preposição *para*, uma vez que as formas oblíquas são regidas por preposições (BAGNO, 2011).

Cereja e Cochar (2013), em *Gramática Reflexiva: texto, semântica e interação*, no capítulo 12, seção destinada para tratar dos pronomes, apontam para a existência dos pronomes pessoais – aqueles que indicam as três pessoas do discurso – e são classificados em retos – função sintática de sujeito ou predicativo do sujeito – oblíquos – são os complementos, subdividem-se em átonos – podem exercer funções de objeto direto, indireto – e tônicos – sempre precedidos de preposição. No entanto, há uma observação feita pelos autores: “[...] os pronomes oblíquos às vezes podem fazer o papel de adjunto adnominal, de complemento nominal e até de sujeito” (CEREJA; COCHAR, 2013, p. 280).

Figura 21 – Representação do fenômeno em estudo em uma gramática pedagógica

Spider: ‘O que der para eu fazer para massagear meu ego, vou fazer’

Campeão dos médios do UFC, Anderson Silva se diz um homem vaidoso com a aparência: ‘Só não vou ao cabeleireiro porque não tenho cabelo’.

(Disponível em: <http://sportv.globo.com/site/eventos/combate/noticia/2012/03/spider-o-que-der-para-eu-fazer-para-massagear-meu-ego-vou-fazer.html>. Acesso em: 20/4/2012.)



Anderson Silva.

Fonte: Cereja e Cochar (2013, p. 145).

Ao citarem essa reportagem, conforme Figura 21, os autores trazem à discussão os aspectos sobre a utilização do *eu* e *mim* em posição de sujeito, quando precedidos pela preposição *para*. Desse modo, indicam que os pronomes pessoais, *eu* e *tu*, exercem a função de sujeito, ao passo que os pronomes *mim* e *ti*, oblíquos tônicos, desempenham outras funções, e concluem:

De acordo com a norma-padrão, na fala do lutador Anderson Silva não seria adequado empregar *mim* no lugar de *eu*. Isso porque o pronome reto *eu* funciona como sujeito do verbo *fazer*, que se encontra no infinitivo [...]. Na norma culta, os pronomes pessoais retos *eu* e *tu* (e somente eles) não podem ser regidos de preposição e desempenhar função de complemento. Por isso, na função de complemento, a preposição obriga o uso dos pronomes oblíquos tônicos *mim* e *ti*... (CEREJA; COCHAR, 2013, p. 146).

Além disso, encontram-se considerações, tais como: a menção feita pelos autores pela extinção do pronome *vós* nas situações de comunicação cotidiana; apresentam o quadro tradicional dos pronomes pessoais e uma subseção chamada de “Contraponto”, em que os autores apresentam um quadro novo do sistema de pronomes pessoais do PB, a partir da perspectiva adotada por Castilho (2010).

Dos quatro tipos de gramáticas analisadas, histórica, normativa/tradicional, descritiva e pedagógica/escolar, notam-se, em algumas delas, aspectos semelhantes, tais como: a) as gramáticas históricas apontam para o fato de ocorrências de dativos assumindo a posição de nominativos, como já espelham os textos antigos, sobretudo textos literários; b) a gramática normativa/tradicional prescreve os usos, afirmando que os pronomes do caso reto devem ocupar a posição de sujeito, ao passo que os pronomes do caso oblíquo ocupam a posição de objeto e complemento, acenando para “erro” gramatical, quanto ao fenômeno ora estudado; c) as gramáticas descritivas trazem aspectos mais contextualizados sobre as ocorrências de *mim* em posição de sujeito; por sua vez, d) as gramáticas pedagógicas/escolares divergem, pelas perspectivas teóricas adotadas pelos diferentes autores: Bechara (2010) segue a linha da prescrição, ao passo que Bagno (2011) e Cereja e Cochar (2013) trazem noções mais

relacionadas ao contexto de usos das formas *eu* e *mim* e suas respectivas atribuições sintáticas canônicas, bem como os rearranjos feitos nessas atribuições.

Vale ressaltar aqui o respeito a todas as abordagens feitas pelos diversos tipos de gramáticas elencadas e mencionadas nessa subseção, uma vez que o foco de observação depende da orientação teórica eleita pelo autor da obra. No entanto, após o cotejo estabelecido entre as perspectivas, quando possível, as abordagens dos três últimos tipos – histórica; descritiva e pedagógica/escolar – serão retomadas, haja vista apresentarem noções que se assemelham com os dados apresentados nessa tese. Não desprezando e/ou rechaçando a abordagem da gramática de orientação tradicional, mas a atual configuração pronominal brasileira já evidencia que o quadro dos pronomes apresentado por tal abordagem carece de atualização.

4.1.2 Perspectiva Gerativa

Pereira e Roncarati (1993), por meio de dados do dialeto-padrão carioca – baseando-se na intuição, enquanto falantes, em casos como “1a. Isto é para eu fazer” – e, para o dialeto não-padrão, utilizaram seis entrevistas – segmentadas, respectivamente, em três do *Competências Básicas do Português* (LEMLE; NARO, 1997) e três da *Amostra Censo* (UFRJ, 1980) – em casos como “1b. Isto é para mim fazer”, analisaram estruturas sintáticas em que o fenômeno ora estudado na tese está presente. Segundo as referidas autoras,

Verificamos, portanto, que há justificadas razões para se afirmar que existe a reanálise de *para* como núcleo de COMP atribuindo caso oblíquo ao sujeito de infinitivo nas construções do tipo 1b, e também para se afirmar que esta reanálise está mais avançada nos dialetos não-padrão, tendo em vista as posições sintáticas mais variadas onde se constata a presença de *para* e o sujeito oblíquo de infinitivo. No entanto, as instâncias de 1b no dialeto padrão nos dão também evidências de que este dialeto não está imune à reanálise. A ausência de sujeitos oblíquos precedidos de preposição diferente de *para* é mais uma evidência desta reanálise (PEREIRA; RONCARATI, 1993, p. 25).

Para as autoras, a alternância entre os pronomes *eu*, caso reto, e *mim*, caso oblíquo, expressos pelas ocorrências 1a e 1b, demonstra uma diferenciação morfológica, supondo que tal diferenciação possa refletir também em uma diferença de caso abstrato. Portanto, as ocorrências de *mim* em posição de sujeito, quando introduzido pela preposição *para*, em orações infinitivas, se dão nos dois dialetos, com predominância no dialeto não-padrão, pois ocorre a marcação excepcional de caso, oriunda da reanálise do *para*, enquanto complementizador, que, ainda assim, preserva a sua propriedade de atribuição de caso.

Silva (1996), ao analisar a estrutura da frase no português brasileiro, mais especificamente o infinitivo flexionado, menciona que a preposição *pra*, em certos contextos frasais, permite um contraste em estrutura livre que pode informar qual é o Caso do sintagma nominal que a segue. Conforme a autora,

(82) a. Ele deu o livro pra eu levar. b. Ele deu o livro pra mim levar. (83) a. ? Ele deu o livro pra amanhã eu levar. b. * Ele deu o livro pra amanhã mim levar. O que os exemplos parecem mostrar é que quando o Caso é realmente atribuído pela preposição - o Caso Oblíquo, realizado como **mim** em (82b) e (83b), a anteposição de um elemento adverbial torna agramatical a frase, enquanto o Nominativo, realizado como **eu** em (82a) e (83a), não parece sensível à intervenção do advérbio (SILVA, 1996, p. 176).

Nessa direção, a referida autora afirma que a concordância infinitiva possui um Caso nominativo autônomo, quando atribuído ao sujeito infinitival pela relação especificador/núcleo. Em contrapartida, o Caso do sujeito também pode vir da parte externa da frase infinitiva, o que permite a preposição receber papéis de regente e atribuidor de Caso no PB, quando funcionado como complementador infinitival (SILVA, 1996).

Mioto, Silva e Lopes (2007) recorrem ao mesmo postulado das autoras supramencionadas para explicar os casos do pronome *mim* ocupando a posição de sujeito, quando precedidos pela preposição *para*, em orações infinitivas. No entanto, eles não assumem que o *para* é reinterpretado como núcleo de CP (complementador), mas que, nos casos com *mim*, não há CP (nem mesmo nulo) entre o *para* e a sentença infinitiva. Para os autores, existem a marcação canônica/regular de caso e a marcação excepcional de caso. A marcação canônica/regular pode ser entendida, além de outras situações, quando os pronomes refletem os casos que possuem os *Determinant Phrases* (DPs), ou seja:

O Caso NOMINATIVO é atribuído à esquerda, ao contrário do OBLÍQUO e do ACUSATIVO que são atribuídos à direita; note que, além disso, a atribuição de Caso é diferente uma vez que o NOMINATIVO é atribuído na relação Spec-núcleo, enquanto o ACUSATIVO e o OBLÍQUO são atribuídos na relação núcleo-Compl (MIOTO; SILVA; LOPES, 2007, p. 179).

A marcação excepcional de caso, *Exceptional Case Marking* (ECM), por sua vez, pode ser entendida como a que se diferencia da marcação canônica/regular por possuir um núcleo que atribuirá Caso aos argumentos de outros núcleos. Por sua vez, a marcação canônica/regular de caso ocorre quando um núcleo atribui caso a um elemento que funciona como seu argumento, seja na posição de seu complemento ou de seu especificador. Para ilustrar a diferença entre as duas citadas marcações de caso, os referidos autores analisam duas sentenças, a saber: a) Maria fez palhaçada para eu rir; b) Maria fez palhaçada para mim rir. Para a primeira sentença, os autores argumentam que “[...] não traz desconforto porque

evidencia uma marcação canônica de Caso: o núcleo Agr presente no contexto do infinitivo pessoal marca casualmente o pronome cuja forma deixa ver o Caso NOMINATIVO” (MIOTO; SILVA; LOPES, 2007, p. 186). Enquanto que, para a sentença b, eles afirmam que

[...] temos em mãos uma marcação casual que não é canônica, no sentido de que ela não acontece na configuração núcleo-complemento, como fazem normalmente as preposições e os verbos. Observe que não é exatamente o complemento que está recebendo o Caso da preposição em (23b); o complemento da preposição não é o pronome *mim* mas uma sentença infinitiva, que em princípio não necessita de marcação casual. Na verdade, é o pronome *mim*, argumento externo de *rir*, que está sendo marcado por Caso pela preposição. Esta é a razão pela qual dizemos que a marcação é excepcional: não é o complemento que recebe o Caso, mas o 'filho' do complemento, digamos assim (MIOTO; SILVA; LOPES, 2007, p. 186).

Ainda, nesse sentido, tais autores asseveram que a sentença b é gramatical, mesmo sendo policiada e combatida pelos gramáticos, que defendem o não uso de tal sentença, por parte dos falantes da língua. Segundo Miotto, Silva e Lopes (2007), tais gramáticos se sentem incomodados, pelo fato de “[...] uma atribuição de Caso que envolve a preposição como núcleo atribuidor e o DP preenchido pelo pronome como receptor. O pronome não consegue disfarçar que se trata do Caso OBLÍQUO e que se trata de ECM” (MIOTO; SILVA; LOPES, 2007, p. 186).

Carvalho (2008), ao estudar a estrutura interna dos pronomes pessoais do PB, a partir das perspectivas léxica e semântica, com base nas teorias da Gramática Gerativa, utiliza a geometria de traços para analisar o comportamento dos pronomes pessoais, afirmando ser a alternância pronominal entre o *eu* e o *mim* em posição de sujeito o resultado de uma simplificação morfológica ocasionada no quadro pronominal do PB.

Para o autor, a categoria de caso do PB obedece à geometria de traços: [C[OBL[GEN][ABL]]]²⁶. Também, há um sincretismo das formas pronominais, sobretudo no que tange a primeira pessoa do singular. Assim, tal sincretismo “[...] é justificado pela subespecificação dos traços formativos destes pronomes: quanto menos traços em sua estrutura, menos restrições o pronome encontra em sua distribuição.” (CARVALHO, 2008, p. 102). Esse sincretismo atingiu, de certo modo, as formas pronominais da primeira pessoa do singular, haja vista que nas outras pessoas o que se evidencia é uma estabilização desse fenômeno, pois:

Aparentemente, uma outra sorte de variação atinge as formas pronominais de primeira pessoa do singular, o sincretismo para as diferentes formas Casuais. A tradicional forma nominativa do pronome de primeira pessoa (*eu*) pode ser encontrada em todas as posições argumentais em algumas variedades do PB [...]. Ainda, a forma oblíqua (*mim*) pode ser encontrada como sujeito encaixado em orações infinitivas posicionadas [...] O sincretismo nas outras pessoas

²⁶ C = complementizador; OBL = oblíquo; GEN = genitivo; ABL = ablativo.

pronominais é um fenômeno aparentemente mais estabilizado em PB. Nesta língua, por exemplo, as formas nominativas dos pronomes pessoais restantes (você, ele, ela, nós, a gente, vocês, eles, elas) são também os representantes dativos oblíquos, sem nenhuma marginalidade em sua aceitação (CARVALHO, 2008, p. 17).

O referido autor continua:

A observação de uma reorganização nos padrões de concordância para todas as pessoas pronominais, por um lado, e o sincretismo pronominal encontrado em PB, por outro, parecem de alguma forma fenômenos conectados. O primeiro fato é provavelmente causado pela substituição de alguns pronomes e a consequente redução do paradigma verbal [...] O segundo fato, apesar de não trazer uma alteração aparente nos padrões de concordância, resulta de uma similar reorganização no paradigma, uma vez que não há correspondência unívoca entre uma forma pronominal e uma única função Casual, como esperado (CARVALHO, 2008, p. 18).

A partir da reorganização pronominal ocorrida no PB, o autor critica as generalizações que não consideram os rearranjos feitos por algumas variedades e propõe:

Quadro 15 – Os pronomes pessoais no PB, conforme Carvalho (2008)

Tipologia dos pronomes de base fonológica		
	clíticos	não-clíticos
1sg	me	eu/mim
2sg	te/lhe	você/tu
3sg	se/lhe	ele/ela
1pl	nos	nós / a gente
2pl	–	vocês
3pl	–	eles/elas

Fonte: Carvalho (2008, p. 33).

Para além disso, o referido autor analisa as proposições anteriores, utilizando os pressupostos da gramática Gerativa, mais especificamente, os estudos de Pereira e Roncarati (1993) e o de Miotto, Silva e Lopes (2007), afirmando que tais postulados não se sustentam por dois principais motivos, a saber: a) presença de um elemento que possui a função de atribuir o caso – no infinitivo pessoal – quando a sentença expressa o *eu*, caso reto, na posição de sujeito, e isso não ocorre quando o *mim* figura na posição de sujeito, sendo estas duas construções semelhantes; b) em algumas sentenças, a forma nominativa é precedida por pronomes sem que um verbo seja posto após ele. No entanto, há algumas sentenças com “para eu fazer” que têm o “eu” como nominativo, mas em outras sentenças têm “para eu fazer”, mas o “eu” é realização do caso oblíquo, assim como em “trouxe aquilo para eu”.

Rocha, em 2010, por meio da morfossintaxe de casos nos pronomes pessoais do português brasileiro falado nos dias atuais, em Minas Gerais, analisa a alternância pronominal entre *eu* e *me*, e apresenta um modelo para Caso, a saber: [KP[DP[NP]]]²⁷. Para Rocha (2010), a estrutura sintática é composta por: um sintagma nominal SN como complemento de um sintagma determinante DP e o autor deve estar propondo a existência de uma categoria funcional “K”, que projeta um sintagma *Kase Phrase*, Sintagma Casual, que toma o DP como complemento. Então, nessa análise, Caso não é apenas uma propriedade, um traço do SN/NP, mas uma categoria gramatical por inteiro. Tal modelo desencadeia implicações específicas para a forma e o comportamento sintático pronominal, pois esses podem ser realizados fonologicamente DP ou KP, correspondentes às estruturas de derivações maiores que D ou N.

Ao analisar o quadro proposto por Carvalho (2008, p. 33), Quadro 15, e as considerações feitas pelo referido autor sobre a argumentatividade e a distribuição das formas pronominais, ao propor para o PB a seguinte divisão: em clíticos – os que ocupam a posição de objeto – tais como: *me*, *te/lhe*, *se/lhe* e *nos*; e não-clíticos – os que ocupam tanto a posição de sujeito quanto a de objeto – tais como: *eu/mim*, *você/tu*, *ele/ela*, *nós/ a gente*, *vocês*, *eles/elas*, Rocha (2010) afirma existir inconsistência nessa proposição, uma vez que o pronome *mim* é uma exceção, quando posto junto das formas pronominais de distribuição irrestrita. Desse modo, “[...] as formas não oblíquas seriam formas *default* em relação a Caso morfológico, enquanto as formas oblíquas seriam aquelas que realizam Caso morfológicamente” (ROCHA, 2010, p. 19). E, para o autor, as formas:

mim e *migo* são KPs com spell-out do tipo K-DP, ou seja, *prep+mim* e *comigo*. Dentro da proposta de Bittner & Hale (1996), isso equivale a dizer que *mim* e *migo* possuem Caso inerente, pois sempre apresentam K (preposição) realizado como uma instância de Seleção, o que determina a presença do pronome e da preposição acompanhante desde a Numeração (ROCHA, 2010, p. 100).

Ademais, sumariza a discussão e elabora um quadro em que apresenta as formas pronominais separadas em formas restritas e formas de distribuição livre:

²⁷ KP = Sintagma Casual; DP = Sintagma Determinante; NP = Sintagma Nominal.

Quadro 16 – Formas pronominais, conforme Rocha (2010)

	Oblíquos	Não-oblíquos
1sg	me/mim	eu
2sg	te/lhe	você/tu
3sg	se/lhe	ele/ela
1pl	nos	nós/ a gente
2pl	–	vocês
3pl	–	eles/ elas

Fonte: Rocha (2010, p. 18).

Na perspectiva da Gramática Gerativa, portanto, observa-se que das quatro abordagens sobre o fenômeno analisado nesta tese, três convergem (PEREIRA; RONCARATI, 1993; SILVA, 1996; MIOTO; SILVA; LOPES, 2007), pois atribuem a ocorrência do fenômeno ora analisado nesta tese por atribuição de Caso, sejam pela marcação canônica e/ou pela marcação excepcional. A terceira abordagem, portanto, a última, proposta por Carvalho (2008), diverge das anteriores, postulando que os pronomes devem refletir o resultado da combinação de estruturas das categorias pessoa, número e gênero, a partir de noções de subespecificação, animacidade, participante e falante, uma vez que não existe relação par a par entre os pronomes e os Casos. Em Rocha (2010), mesmo analisando as alternâncias pronominais entre o pronome átono, *me*, e o pronome pessoal, *eu*, encontram-se considerações que, conforme Quadro 16, levam a aludir à validade dessa proposição, sobretudo, por conta da colocação do pronome *mim* entre as formas restritas.

Embora se considerem as contribuições trazidas por esses estudos, no que tange às análises e às elucidações sobre as estruturas de *para/prá + eu/mim + infinitivo*, com base nos aspectos formais e teóricos da Gramática Gerativa, nesta tese, a perspectiva de abordagem do fenômeno se dará por meio dos suportes teóricos postulados pela Dialetologia, e da Sociolinguística Laboviana, como suporte adicional.

4.1.3 Perspectiva Sociocognitiva

Torrent (2005), através dos pressupostos teóricos do modelo sociocognitivista e também do conceito de Construção, de proposição de Goldberg (1995), estuda as construções de Dativo com Infinitivo (DCI), em específico, as construções de *para mim + infinitivo*, através de dados de fala disponíveis no banco de dados do Programa de Estudos sobre o uso da Língua – PEUL/RJ – e dados de escrita retirados de sites da internet.

Para o autor, o esquema da sintaxe do *para x infinitivo* se originado no Latim, por meio de duas construções que resultam da cláusula de finalidade, bem como de Dativo Agente, dando origem a uma nova e única construção. Também, afirma que o DCI é produto de uma Herança por Mesclagem entre a fonte 1, construções Transitivas Transferenciais – são aquelas constituídas por um elemento que é Agente atua sobre um Paciente fazendo com que o mesmo se transfira para um Beneficiário, ou Destinatário –, e a fonte 2, construções Transitivas Básicas – são aquelas que um Agente atua sobre um Paciente –, havendo, dessa maneira, uma fusão de dois papéis temáticos distintos em um elemento, o produto da Mescla. (TORRENT, 2005).

Torrent (2005), após buscar caminhos históricos para explicação do *para + mim + infinitivo* no latim e nas gramáticas tradicional, gerativa e esclarecida, propõe que a estrutura DCI apresenta algumas características, a saber:

- (i) O DCI é uma Construção, cujo esquema sintático é *para x infinitivo*; (ii) A posição de *x* pode ser ocupada por vários tipos de Nominais, sejam eles Pronomes Dativos, Pronomes Nominativos, Nomes ou Sintagmas; (iii) A ocorrência de Pronomes Dativos na posição de *x* é apenas uma forma de reforçar morfologicamente a atribuição de caso Dativo a este elemento, a qual é feita sintaticamente através da preposição *para* e denunciada, a princípio, pela ausência de concordância deste elemento com o Infinitivo, conforme ocorria em Latim com o Gerundivo[...] (TORRENT, 2005, p. 72-73).

Além disso, conforme postulados finais do autor, as construções do DCI podem se mesclar com outras construções; o DCI pode ou não possuir ligação com um Verbo Transferencial Finito; e, por fim, a construção do DCI diferencia de outras construções da língua, uma vez que elas possuem características semânticas próprias, sejam os papéis de Beneficiário, Agente ou Experienciador. Por fim, o autor sugere que pesquisas quantitativas possam investigar o fenômeno, bem como um estudo sobre aspectos do Português Arcaico, a fim de traçar os caminhos do fenômeno até o PB.

Até então, pelo que se tem notícia, o trabalho de Torrent (2005) é o precursor, a partir dessa vertente de observação teórica, a analisar o fenômeno *para mim + infinitivo* no PB. As postulações feitas pelo referido autor têm relevância significativa e podem servir de base para, juntamente com os olhares teóricos da Dialetologia e da Sociolinguística, ajudar a interpretar os possíveis caminhos do fenômeno em diferentes sincronias do PB.

4.1.4 Perspectiva Sociolinguística

Em 1991, Monteiro, utilizando amostras de fala – 60 inquéritos (12 de cada capital: Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre) – do Projeto de Estudo da Norma Linguística Urbana Culta, o Projeto NURC²⁸, observa a natureza dos pronomes, a variação e a reestruturação do quadro pronominal do PB, bem como as funções dêitica e anafórica dessa classe. Posteriormente, em 1994, o estudo apresentado e defendido como tese de doutoramento foi publicado em formato de livro impresso, a obra é intitulada *Pronomes pessoais: subsídios para uma gramática do português*.

Nesta obra, Monteiro (1994) assevera que os pronomes pessoais possuem uma complexidade, e que seu entendimento perpassa as perspectivas sintática, semântica e pragmática, sobre uma abordagem morfossintática. Além disso, tal autor afirma que a primeira pessoa do singular preservou a sua flexão de caso, uma vez que, apesar dos rearranjos pronominais ocorridos no PB, a primeira pessoa não foi impactada.

Desse modo, segundo o referido autor, o pronome do caso reto *eu* continua como a única forma representativa do sujeito, ao passo que as outras formas oblíquas, *me*, *mim*, *comigo*, exercem as funções de complementos verbais e adverbiais, para o paradigma de primeira pessoa do singular, se observados os aspectos atinentes às variedades cultas. No entanto, sobre as construções com infinitivo, nota-se a seguinte observação:

E, não raro, o sujeito de orações infinitivas, por influência de uma preposição, assume a forma oblíqua tônica: (114) *tem tanta (propaganda) que para mim lembrar é difícil* (121.POA.DID.F.F.); (115) *isto aí é um dado do problema para ti chegar no outro* (206.poa.ef.f.i.). O fenômeno se explica em virtude do sincretismo de funções e conseqüente possibilidade de dupla interpretação. Assim, em 114, podemos ter como leituras: a) “*tem tanta propaganda que, para mim, lembrar, é difícil*”; b) “*tem tanta propaganda que pra (eu) lembrar é difícil*” (MONTEIRO, 1994, p. 85-86).

Monteiro (1994), ao citar Álvarez (1979), lembra que o pronome *ti* exerce a função de pronome sujeito na segunda pessoa do singular, sendo um equivalente ao pronome *tu* no galego, como resultado de uma transferência do dativo, haja vista que transformações dessa natureza acontecem, de forma recorrente, nas línguas românicas (MONTEIRO, 1994, p. 85-86) em sentenças infinitivas. Para ilustrar a assertiva, o mencionado autor utiliza o contexto retirado do *corpus* do NURC, a saber: “*isto aí é um dado problema pra ti chegar no outro* (206.POA.EF.F.I.)”. Em alguns dialetos brasileiros, observa-se a alternância *tu~ti* em infinitivas finais, sobretudo no Rio Grande do Sul, região com grande influência das línguas

²⁸ Mais detalhes sobre as características do Projeto NURC serão evidenciados na seção 5 desta tese.

de imigração, bem como da língua espanhola, esta última, devido à localização fronteiriça com Argentina e Uruguai.

Figueiredo (2007) analisa construção *para* + *IPS* + *infinitivo*, investigando se as formas pronominais *eu* e *mim* estão em processo de variação ou mudança linguística em orações infinitivas iniciadas pela preposição *para*. Para o empreendimento, utiliza-se de dois *corpora*, com informantes do Rio de Janeiro, a partir de entrevistas sob os moldes da Sociolinguística Variacionista, com dados do PEUL/UFRJ, amostras de épocas distintas, 1980 e 2000, dos tipos: Amostra Censo – 64 informantes – Amostra Tendência – 32 informantes.

Nas duas amostras, os informantes estão distribuídos sistematicamente em gênero/sexo, idade – 7 a 14 anos; 15 a 24 anos; 25 a 49 anos; e maiores de 50 anos; e escolaridade – primário, ginásio e colegial²⁹. Foram utilizadas como variáveis linguísticas a função sintática da oração infinitiva – orações subjetivas, orações objetivas, orações finais e as completivas nominais –, o paralelismo, o tipo de texto, a modalidade, a correferência de sujeitos e a natureza semântica do verbo da oração principal. As variáveis sociais consideradas foram idade, idade e escolaridade e escolaridade e gênero/sexo.

Nas análises iniciais, o referido autor constatou que:

A variante *eu*, forma prescrita pelas gramáticas normativas, não se mostrou expressiva em nossos dados. Ao contrário, ocorre com quase nula frequência, 4% nas amostras checadas. A variante *mim*, forma discriminada, apresenta percentagem mais elevada, 20% e a variante zero, que não é mencionada na disputa, destaca-se, com 75% dos casos. Considerando a tendência do português brasileiro a não apagar o pronome sujeito na fala (Paredes Silva, 1988) este comportamento se mostra ainda mais interessante (FIGUEIREDO, 2007, p. 15).

Figueiredo (2007), ao analisar a variação entre o pronome *mim* e a variante zero, uma vez que o pronome *eu* não foi considerado, devido à baixa frequência (4%), concluiu que há uma “[...] sistematicidade da distribuição entre *mim* e *zero* ao mesmo tempo em que identificamos os contextos linguísticos e sociais para seu uso. Validamos assim, implicitamente, as variantes estudadas”. A função sintática da oração infinitiva foi a variável mais relevante, mas também a escolaridade apresentou índice de relevância notável. Ao passo que, também, “o pronome explícito *mim* funciona inequivocamente como sujeito de infinitiva, independentemente da função que eventualmente exerça na oração principal” (FIGUEIREDO, 2007, p. 107).

²⁹ Atualmente, correspondem à Ensino Fundamental I, Ensino Fundamental II e Ensino Médio, respectivamente.

Carvalho e Almeida (2013), por meio da fala dos informantes do português popular de Feira de Santana-Ba, observaram as formas pronominais *eu* e *mim* em estruturas “preposição *para* + *eu* e preposição *para* + *mim*” nas posições de sujeito e de objeto indireto, descrevendo os usos, além de correlacioná-los aos fatores linguísticos: forma da preposição – *para*; *pra* ou *pa* – e aos fatores extralinguísticos: faixa etária – faixa 1: jovens até 35 anos; faixa 2: pessoas até 50 anos; e faixa 3: idosos a partir de 60 anos; gênero/sexo: masculino e feminino; escolaridade: alfabetizados ou analfabetos; localidade: zona rural e zona urbana. Para o empreendimento, o *corpus* é constituído por 24 entrevistas nos moldes sociolinguísticos, estratificadas em 12 informantes da zona rural e 12 informantes da zona urbana.

No cômputo geral dos dados, no que se refere ao fenômeno ora analisado, as autoras afirmam que, das 34 ocorrências, 26,5% correspondem à forma *para mim*, totalizando nove ocorrências. Para o único fator linguístico observado, notou-se que a forma mais utilizada é a forma *pra*, uma vez que as formas *para/pa* em frases na posição de sujeito não foram encontradas. Conforme as autoras: “[...] percebemos uma quase total ocorrência da forma *pra*. Não houve nenhuma realização da forma *para* prescrita pelas gramáticas e houve ainda duas ocorrências da forma *pa*, apenas na posição de objeto indireto” (CARVALHO; ALMEIDA, 2013, s/d).

Quanto aos resultados dos fatores extralinguísticos, para a faixa etária, os informantes da faixa 2, seguidos pelos informantes da faixa 3 são os que mais utilizam forma *para mim*, com 50% e 40%, respectivamente, e os informantes da faixa 1 apresentaram 11,1%. Na variável sexo/gênero, a forma considerada inovadora e estigmatizada, *para mim*, foi encontrada 43,8% na fala das mulheres e 11,1% na fala dos homens. Para o fator localidade, nota-se que o contexto urbano favorece a utilização da forma *para mim*, uma vez que apresentou 50% do número de ocorrências, ao passo que, no contexto rural, 10%. A escolaridade apresentou percentuais de 29% para os informantes alfabetizados e 0% para os informantes analfabetos. No tocante ao *para mim*, na fala de informantes analfabetos, trata-se 0% de em um total de apenas três dados. O que não quer indicar, uma tendência de uso da forma padrão entre os analfabetos, mas sim um registro com poucos dados.

A partir da fala do português popular da Bahia, Mendes (2016) investigou a flexão casual dos pronomes pessoais nas diversas funções sintáticas, no *continuum* de urbanização do PB, utilizando-se de 160 entrevistas em quatro *corpora*: o do português afro-brasileiro (PAF) – Helvécia, Cinzento, Rio de Contas e Sapé –; a do português popular de cidades do interior de pequeno porte (PPI) – Santo Antônio de Jesus e Poções –; a do português popular

da cidade do interior de médio porte (PPF) – Feira de Santana –; a do português popular urbano (PPS) – Salvador; esses dados estão diretamente vinculados a dois grandes projetos, respectivamente, *Projeto Vertentes do Português Popular do Estado da Bahia*, os três primeiros *corpora*, e o Projeto *A língua portuguesa no semiárido baiano*, o *corpus* sobre Feira de Santana.

Sob a ótica de uma realidade linguística brasileira que se configura por normas, divididas em dois polos: as normas cultas urbanas e as normas populares (LUCCHESI, 2015), a autora sugere um quadro para os pronomes encontrados nas duas normas:

Quadro 17 – Os pronomes pessoais do PB, segundo Mendes (2016).

Normas cultas urbanas		Normas populares (e semi-cultas) urbanas	
Sujeito	Complementos verbais e adverbiais	Sujeito	Complementos verbais e adverbiais
eu	me, mim, comigo	eu	me, mim, comigo
eu, você	você, te, lhe, ti, contigo	tu, você	você, te, lhe, ti, contigo
ele (as)	ele (a) o, a, lhe	ele/ela	ele, ela
vocês	vocês	vocês	vocês
nós, a gente	a gente, nos, conosco	nós, a gente	nós, a gente, nos, conosco
ele (as)	ele (as), os, as, lhes	ele (as)	ele (as)

Fonte: Mendes (2016, p. 191).

Mendes (2016) afirma que “[...] a flexão de caso dos pronomes pessoais de primeira pessoa do singular é um aspecto que contrapõe variedades populares mais urbanizadas e variedades populares menos urbanizadas...”, haja vista que os informantes dos grandes e médios centros urbanos – Salvador e Feira de Santana – têm sido resistentes quanto ao uso dos pronomes flexionados de primeira pessoa do singular, no que tange às posições de adjuntos adverbiais e complementos verbais, a saber:

Quadro 18 – Pronomes pessoais de primeira pessoa do singular, conforme Mendes (2016)

A flexão de caso dos pronomes de primeira pessoa do singular no português popular da Bahia			
	Sujeito	Complementos Verbais	Adjuntos Adverbiais
PAF	eu	eu ~ me/mim/comigo	eu ~ mim/comigo
PPI	eu	eu ~ me/mim/comigo	eu ~ mim/comigo
PPF	eu	me/mim/comigo	eu ~ mim/comigo
PPS	eu	me/mim/comigo	eu ~ mim/comigo

Fonte: Mendes (2016, p. 201).

A referida resistência do PPF e do PPS não ocorre, ao observar as falas do PAF e do PPI (MENDES, 2016). O fator urbanização foi elencado como um importante favorecedor para a aplicação da regra da flexão dos pronomes pessoais de primeira pessoa, haja vista que existe uma clivagem, no português popular baiano, constituída por variedades do PAF e PPI, em um grupo, e, no outro grupo, PPF e PPS. A referida autora não apresenta cálculos específicos para os pronomes *eu* e *mim*, haja vista ter aspirações e objetivos diferentes.

Maia (2016) investigou o uso variável dos pronomes pessoais de 1ª pessoa do singular *eu/mim*, em orações infinitivas iniciadas por *para*, e *mim/me*, em posição pré-verbal, utilizando-se de narrativas pessoais – textos escritos – construídas por estudantes do Ensino Fundamental de duas turmas, 6º e 9º anos, de um colégio estadual localizado em Ilhéus-Ba, utilizando como variáveis sociais para delimitação dos usos a escolaridade e o gênero.

Para os pronomes *eu/mim*, no cômputo geral, a referida autora afirma que, além da competição desses dois pronomes, *eu* (9,7%) e *mim* (21%), o zero anafórico/catafórico (69,3%) foi, portanto, bastante encontrado nas produções dos estudantes. No que tange aos fatores sociais:

[...] a diferença mais acentuada ocorre com o uso da variante estigmatizada, para *mim*, mais usada pelas alunas (84,6%) do que pelos alunos (15,4%). São elas também que usam mais o apagamento (53,5%), mas com uma diferença não muito grande em relação aos alunos (46,5%). Quanto a essa regra do apagamento, os dados evidenciam que tanto alunos quanto alunas a reconhecem como uma estratégia natural, talvez por conta da dificuldade que eles têm para identificar as regras para o uso das formas expressas. No que se refere à variante padrão, para *eu*, numa leitura horizontal, não percebemos diferença de uso, ou seja, não há variação; a variação é registrada numa leitura vertical: os alunos usam mais a variante prescrita e as alunas, a variante estigmatizada (MAIA, 2016, p. 41).

Maia (2016), na conclusão do estudo, propõe alternativas didáticas para que os professores trabalhem os pronomes pessoais nas aulas de Língua Portuguesa, por meio de um caderno de orientações, indicando que tais alternativas visam o combate ao preconceito

linguístico e uma maior aceitação dos fenômenos que surgem na língua e que, por vezes, contrariam o que preconiza o padrão normativo.

Gomes (2019) analisou a fala de 152 informantes de sete localidades interior de São Paulo – São José do Rio Preto; Mirassol; Guapiaçu; Bady Bassit; Onda Verde; Cedral; e Ipiguá – utilizando o banco de dados do IBORUNA³⁰, Amostra Censo. Os informantes estão distribuídos em faixa etária – 7 a 15 anos, 16 a 25 anos, 26 a 35 anos, 36 a 55 anos, e mais de 50 anos –; escolaridade – Ensino Fundamental I, Ensino Fundamental II, Ensino Médio, e Ensino Superior –; gênero – masculino e feminino –; e renda – até 5 salários mínimos; de 6 a 10 salários mínimos; de 11 a 24 salários mínimos; mais de 25 salários mínimos.

No estudo, foram encontrados 418 dados, desses, 28 ocorrências do pronome *eu*, 82 ocorrências do *mim* e, por sua vez, 308 ocorrências do zero anafórico. Segundo a referida autora,

Os resultados mais relevantes da análise ternária revelaram que a variante *zero* é a mais utilizada por todos os perfis sociais. Em relação às variantes *mim* e *eu*, nota-se distribuição social com maior predomínio da variante *mim* na fala de informantes do sexo masculino, com baixa escolaridade, pertencentes às faixas etárias mais jovens e às classes socioeconômicas inferiores – encaixamento social indicativo de uma mudança em curso em direção à forma não padrão (GOMES, 2019, p. 138).

Nessa perspectiva, ao concluir o estudo, Gomes (2019) indica que, na variedade escolhida para o estudo, os resultados apontam que os usos dos pronomes *mim* e *eu* estão condicionados por fatores tais como gênero, escolaridade e renda. Também, destaca que o perfil que mais utilizava a variante estigmatizada, *mim*, é o homem, com escolaridade até o Ensino Médio e que possui baixa renda.

Segundo a referida autora, “[...] Com esses resultados, conclui-se que não há variação estável entre as duas formas na variedade investigada, mas uma mudança em curso em direção à variante *mim*, apenas refreada pelo alto grau de escolaridade do falante” (GOMES, 2019, p. 140). Por fim, ademais, sugere que outras pesquisas sejam feitas sobre o fenômeno, tendo como foco sincronias passadas, observando o português europeu e o português africano.

Silva e Nogueira (2018) observaram a alternância dos pronomes pessoais *eu* e *mim* na posição de sujeito das orações, por meio de entrevistas sociolinguísticas, com dados orais, da cidade de Jacobina, interior da Bahia. O estudo contempla uma amostra de fala com 12 informantes, estratificados em seis informantes de cada sexo – homem e mulher – duas

³⁰ Maiores informações podem ser obtidas no site do ALIP: Disponível em: <<https://alip.ibilce.unesp.br/bancosde-dados/banco-de-dados-iboruna>>. Acesso em: 30 ago. 2022.

faixas etárias – 14 a 28 anos e de 29 a 75 anos – com níveis de escolaridade distintos – Ensino Fundamental II, Ensino Médio e Universitário. Segundo o estudo, foram encontradas 34 ocorrências válidas para o cômputo dos dados, 9% são da utilização das duas formas, *eu* e *mim*; 6% são da forma pronominal *eu*, e, por fim, 85% são da forma pronominal *mim*.

Conforme dados da pesquisa, o pronome *mim*, embora seja uma forma estigmatizada, quando utilizada em posição de sujeito, foi a mais produtiva na fala dos jacobinenses. No que tange aos fatores extralinguísticos, os relevantes para a realização do fenômeno, quando observada a variação diagenérica, os homens são os que mais utilizam a forma oblíqua tônica *mim*. Tal forma também é mais frequente na fala dos mais jovens. Além disso, por meio do estudo, os autores afirmam que:

Os resultados encontrados revelam que quanto menor o nível de escolaridade, maior é a utilização da variante estigmatizada *mim* desempenhando a função de sujeito na fala dos entrevistados, e menor o uso da forma considerada de prestígio *eu na função de sujeito* corroborando a ideia de que a escola atua como preservadora das formas de prestígio (SILVA; NOGUEIRA, 2018, p. 52).

Por fim, o estudo atesta que os pronomes pessoais de primeira pessoa do singular estão em variação, na variedade analisada, e que podem ser encontrados na fala de qualquer informante, independente da sua escolaridade.

Novais (2018) analisou o fenômeno dos usos dos pronomes *eu* e *mim* em posição de sujeito, por meio dos dados do Projeto LUSA³¹ – amostra constituída por 96 informantes – e das percepções dos estudantes universitários do Sertão do Estado de Alagoas também. Para a coleta de dados do último *corpus*, a referida autora utilizou-se de uma entrevista-questionário – teste de reação subjetiva – a amostra recolhida consta de 70 participantes, mas somente 56 desses foram validados para o cômputo geral dos dados.

No que tange aos resultados, após o tratamento dos dados do Projeto LUSA, foi constatado um baixo índice de ocorrência dos pronomes *eu* e *mim* no contexto esperado e analisado, 12 ocorrências no total geral, sendo essas cinco ocorrências de *pra eu* e sete de *pra mim*, todas mediante a preposição em forma contracta, *pra*. Nas entrevistas, as percepções dos universitários:

[...] obtivemos um resultado significativo com um total de 224 respostas sobre a preferência dos falantes em relação ao uso dos pronomes *eu* e *mim*, sendo que 157 são referentes ao uso do eu como sujeito de frases infinitivas em oposição de 67 para o uso de *mim* na posição de sujeito precedido pela preposição *pra*. Esses resultados mostram um percentual de 70% das respostas para a variante *eu* e 30% para a variante *mim*... (NOVAIS, 2018, p. 50).

³¹ Mais detalhes podem ser obtidos em Vitório (2017).

Ainda, nesse contexto, a autora reflete sobre as percepções dos entrevistados, analisando sobre a ótica de que o estigma linguístico, ao determinar formas “certas” e formas “erradas” têm associação com o estigma social, pois, por exemplo, “[...] sujeitos que fazem uso de *mim* em posição de sujeito em frases infinitivas antecidas por *para* acabam sofrendo preconceito linguístico, que é uma forma camuflada de preconceito social...” (NOVAIS, 2018, p. 60).

Vaz (2020), com base na Sociolinguística Laboviana, analisou a escrita de jovens brasileiros em grupos virtuais, observando o contexto do Facebook, tencionando a utilização dos pronomes *eu* e *mim* em posição de sujeito em orações infinitivas. A amostra foi constituída por 24 informantes – estratificados em uma única faixa etária (entre 18 e 28 anos); gênero (masculino e feminino) e escolaridade (universitários e não universitários). Além disso, observa o fenômeno diante de verbos simples (formar, usar, mandar, tirar, etc.) e verbos complexos (querer, fazer, ter, dizer, etc.). Para a composição do *corpus*, a referida autora utilizou-se de trechos publicados pelos informantes no ano de 2017.

Quanto aos resultados, foi constatado um baixo índice de ocorrência dos pronomes *eu* e *mim*, 23 ocorrências no total geral, sendo essas 16 ocorrências de *eu* e sete de *mim*. No tocante à escolaridade, os universitários optam pela forma considerada de prestígio, ao passo que a forma estereotipada é mais preeminente no grupo não universitário.

Em relação à variação diagenérica, notou que as mulheres tendem ao pronome *eu*, haja vista que:

[...] E, foi possível assegurar nas observações feitas que no grupo de gênero feminino não universitário a utilização do pronome *mim* como sujeito foi de 12,5% e a utilização do pronome *eu* como sujeito foi de 37,5%, um resultado significativo sobre a preferência dos falantes em relação ao uso dos pronomes *eu* e *mim*. Já no grupo de gênero feminino universitário, a utilização do pronome *eu* como sujeito foi de 43,8% e do *mim*, houve considerada redução, de 6,3% (VAZ, 2020, p. 14).

Por sua vez, para os homens universitários, os resultados foram de 12,5% do pronome *mim* e 37,5% do pronome *eu*, ao passo que, para os homens não universitários, os resultados encontrados foram 25% de utilização do *mim* e 25% de utilização do *eu*.

Ainda, quanto aos verbos, a autora reflete sobre as utilizações dos informantes, “[...] Os falantes observados, em sua maioria, utilizam verbos simples com mais frequência nas redes sociais informais...” (VAZ, 2020, p. 14).

Soares e Milani (2022) analisaram o uso do pronome *mim* em posição de sujeito nos *corpora* do NURC/RJ, nos anos 70 e 90, utilizando-se do D2 – Diálogos entre dois Informantes – e o DID – Diálogos entre Informante e Documentador, com informantes com

idades entre 25 a 56 anos. Os referidos autores fazem uma pesquisa em gramáticas e apontam que o fenômeno já ocorria no Latim.

No que concerne aos resultados, após o tratamento dos dados do Projeto NURC/RJ, foi constatado um baixo índice de ocorrências, apenas, 10 ocorrências do *mim* no total geral, sendo essas oito ocorrências no *corpus* da década de 70, e duas na década de 90. Nas palavras dos referidos autores,

[...] podemos observar que o pronome *mim* está contrariando a norma apresentada pelos gramáticos, pois a colocação pronominal de *mim* não é utilizada como objeto indireto dentro das falas dos informantes do NURC/RJ. Nesses casos, o uso do *mim* antecedidos pelo verbo no infinitivo seriam considerados um tipo de anomalia na língua formal... (SOARES; MILANI, 2022, p. 174).

Por fim, após análises de dados, os autores concluem que há um abismo entre a língua descrita pelas gramáticas normativas e as reais condições de usos da língua, uma vez que, mesmo na fala de informantes universitários, considerados representantes da fala culta, a forma estigmatizada é encontrada.

Por meio do olhar teórico sociolinguístico, o fenômeno abordado nesta tese tem um lastro, sobretudo porque diversos recortes podem ser vistos, em localidades diferentes de todo o Brasil. Isso, de certa maneira, ajuda a elucidar ainda mais não somente a realidade do quadro pronominal brasileiro, mas também permite visualizar como se encontram os pronomes pessoais de primeira pessoa, nas variedades analisadas.

Desde o primeiro estudo, nesta perspectiva teórica, Monteiro (1996) até os estudos mais recentes, percebe-se que os pronomes pessoais, no que tange à primeira pessoa do singular, apresentam modificações em relação às suas funções canônicas, quer sejam de sujeito, quer sejam de objetos e complementos, conforme apontado pelos estudiosos sociolinguistas que se debruçaram para analisar o tema em questão. Tal fator é bastante salutar, uma vez que tais estudos já apontam caminhos que servirão aqui, nesta tese, para análise e interpretação do fenômeno em todas as capitais brasileiras, com dados de sincronia mais recente, e em algumas capitais, com dados de sincronias mais distantes temporalmente.

O breve panorama descrito na perspectiva da Sociolinguística, revela que, na maioria dos estudos que retratam localidades espalhadas pelo Brasil, há uma predominância da forma *mim*, tais como evidenciam: Figueiredo (2007) – Rio de Janeiro-Rj; Maia (2016) – Ilhéus-Ba; Silva e Nogueira (2018) – Jacobina-Ba; Gomes (2019) – interior de São Paulo; Soares e Milane (2018) – Rio de Janeiro-Rj. Ao passo que, somente, em Carvalho e Almeida (2013) – Feira de Santana-Ba; em Novais, (2018) – sertão do estado de Alagoas; e, em Vaz (2020) – Brasília-DF, Facebook; o pronome *eu* foi predominante.

4.1.5 Perspectiva Dialeológica

Na dimensão dos estudos dialetais, o fenômeno ora estudado nesta tese tem as primeiras considerações, pelo que se nota, em obras de cunho monográfico, três importantes referências para o firmamento e desenvolvimento desse ramo dos estudos linguísticos, no que concerne à língua falada. A primeira consideração é observada em *O dialeto caipira*, de Amadeu Amaral, aspectos sobre a língua falada em São Paulo, a saber:

5. As cláusulas infinitivas dependentes de *para* têm por sujeito o pronome oblíquo *mim*, nos casos em que o sujeito devesse ser eu: *Êle trôxe umas fruta pra mim cumê (r). Este, como muitos outros, como quase todos os fatos da sintaxe caipira e popular de S. Paulo, repete-se nas outras regiões do país.* Um exemplo dos “Cantos populares de S. Romero [...] (AMARAL, 1920, p. 75) [grifos nossos].

Nas palavras do referido autor, tal fato pode ser notado em variedades do falar de São Paulo, mas também em outras regiões brasileiras. Bagno (2009, p. 36) reforça essa assertiva, a saber:

Apesar de tão antiga e de sua ampla difusão, essa construção ainda é alvo de rejeição por parte dos falantes letrados das outras regiões do país, de modo que não podemos tentar defendê-la, aqui, como uma regra já perfeitamente implantada em todas as variedades linguísticas brasileiras. Quem sabe alguém possa fazer essa defesa daqui a cinquenta ou cem anos, sobretudo diante de seu amplo emprego pelos professores de português que, hoje, em sua maioria, procedem de camadas sociais onde a fórmula para *mim* + infinitivo está bem implantada. Por enquanto só podemos supor que sua ascensão social prosseguirá até, talvez, ganhar espaço também nos gêneros escritos mais monitorados (BAGNO, 2009, p. 36).

A segunda consideração encontra-se em *O linguajar carioca*, de Antenor Nascentes, aspectos voltados para a língua falada no Rio de Janeiro:

As reduzidas de infinitivo com sujeito de primeira pessoa do singular e regidas pela preposição *para* trazem o pronome no caso oblíquo em vez de trazê-lo no caso reto: *Isto é para mim fazer [...]* **A impressão causada pela preposição é tão forte que o uso correto do pronome passa aos ouvidos do povo como um erro e êle o emenda do modo apontado.** Cumpre notar que o povo às vezes procura acertar; vimos casos de emendas infelizes na fonética (arfalhate, velgonha, etc.) e na morfologia (lezes); se erra, não é por falta de vontade de acertar (NASCENTES, 1953, p. 130-131) [grifos nossos].

Conforme observado, o referido autor aponta para uma hipercorreção por parte dos falantes, mediante a presença da preposição *para*. Por fim, ainda sobre essas considerações em obras monográficas pioneiras nos estudos dialetais brasileiros, a terceira obra é *A língua do Nordeste*, de Mário Marroquim, aspectos sobre a fala das Alagoas e do Maranhão. Conforme o autor,

172. A tendência para o uso do pronome nominativo faz que o matuto, da mesma forma que altera a sintaxe portuguesa com o fim de empregar o caso reto, se

conserve a ela fiel quando a língua culta autoriza o uso daquele caso [...] **Entretanto, no sul do país, segundo observação e testemunho de Antenor Nascentes e Amadeu Amaral, nessas orações, o pronome toma, no dialeto, o caso oblíquo:** “isto é para mim levar”. “Ele trôxe uma fruta pra mim comê” (MARROQUIM, 1934, p. 125) [grifos nossos].

Há indicações, segundo o referido autor, de que o fenômeno se encontra na região Sul do Brasil, ao apontar para as constatações mencionadas das duas obras precursoras.

Santos e Paim (2018), com base nos dados do Projeto ALiB, observaram as respostas fornecidas por 32 informantes de quatro capitais da região Sudeste – Vitória, Rio de Janeiro, Belo Horizonte e São Paulo. Os informantes estão estratificados conforme critérios metodológicos do projeto – sexo; faixa etária; e escolaridade – subseção 5.2. Para o estudo, as autoras utilizaram os Temas para Discurso Semidirigido (TDS), além da questão que se destina a apurar a variação entre as duas formas pronominais, *eu* e *mim*, QMS 023³².

Após o levantamento e análises dos dados, as autoras constataram 32 ocorrências, todas no QMS 023, segmentadas da seguinte maneira: 15,60% dos informantes realizaram as duas formas, *eu/mim*; 31,30% dos informantes utilizaram somente o pronome *eu*; e 53,10% dos informantes utilizaram somente o *mim*. Para os aspectos sociais analisados, na fala dos informantes homens, foram documentadas 62,5% das ocorrências de *mim*, e, na fala das informantes mulheres, 43,75%. A faixa etária I apresentou 50% de ocorrências da forma *mim*, ao passo que a faixa II 56,23%. Para a escolaridade, foram documentadas 81,25% de ocorrências da forma oblíqua na fala dos informantes com ensino Fundamental incompleto, e 25% na fala dos universitários.

Por fim, afirmam que “Os resultados encontrados revelam que o acesso à escolarização é determinante no uso das variantes *eu* e *mim*, corroborando a ideia de que a escola atua como conservadora das formas de prestígio...” (SANTOS; PAIM, 2018, p.149).

Também, ao utilizar dados do Projeto ALiB, Martins (2019) apresenta uma análise sobre os pronomes pessoais *eu* e *mim*, a partir das falas dos informantes das 25 capitais que compõem a rede de pontos do referido Projeto, observando as respostas dadas para a pergunta 023 do QMS por 200 informantes e também nos TDS.

No cômputo geral, foram catalogados 202 dados, 114 registrados com o uso *pra mim fazer*, 88 registrados com o uso do *pra eu fazer*, correspondendo a 56,5% e 43,5%, respectivamente, com poucos casos da utilização das duas formas, apontando que, em alguns casos, há sinalização de autocorreção feita por esses informantes.

³² Melhores detalhamentos sobre os questionários do Projeto ALiB serão feitos na seção 5.2.

No que tange à variação no espaço, a autora atesta que o Nordeste se destaca pelo fato de apresentar maior presença do pronome *eu* em posição de sujeito – 53.1%, com destaque para Salvador, com resposta categórica para o pronome do caso reto – enquanto nas outras regiões se encontra a maior presença do pronome oblíquo – Sul (65.4%), Sudeste (54.3%), Centro-Oeste (81%) e Norte (55.4%). Com isso, ao concluir o estudo, a referida autora apresenta uma carta experimental e indica que outras investigações sejam feitas sobre esses pronomes, principalmente nas cidades do interior.

Santos e Ribeiro (2020), em perspectiva mais ampla, também analisam os dados das capitais do Projeto ALiB, com vistas a apurar a variação entre os pronomes *eu* e *mim* em posição de sujeito em orações infinitivas iniciadas por *para/prá*, em todas as partes do Questionário do Projeto. Nessa análise, além de consultar as gramáticas tradicionais e pedagógicas, os autores traçam um breve panorama do fenômeno a partir do Latim e apresentam alguns estudos, sob orientação sociolinguística, que investigaram o *eu* e o *mim* em posição de sujeito. Diferentemente dos estudos anteriores, sob a perspectiva dialetológica, Santos e Ribeiro (2020) submetem os dados recolhidos ao pacote de programa estatístico *Goldvarb* (2001), além de fornecer cartas linguísticas com percentuais e pesos relativos.

De acordo com os referidos autores, foram catalogadas 329 ocorrências, 14% dessas foram agrupadas em Não Obtidas, ao passo que 86% foram respostas validadas entre os dois pronomes em variação. Desse total, 166 ocorrências foram obtidas no QMS 023 e 116 ocorrências foram obtidas em outras partes do inquérito. Entre as formas pronominais variantes, a distribuição foi de 27,4% da forma prescrita pela norma gramatical, *eu*, e 72,6% de uso da forma inovadora/estigmatizada, *mim*, haja vista que

Os dados referentes à estrutura sintática analisada permitem afirmar que, contrariando o imaginário popular, o qual difunde que *mim* em posição de sujeito é predominante na fala rural, indígena e/ou de baixa escolaridade, a maioria dos brasileiros, informantes das capitais, opta pela forma *mim* (SANTOS; RIBEIRO, 2020, p. 181).

Para a diatopia, os autores destacam que algumas capitais apresentam pesos relativos acima de elevados, acima de 0.50, portanto, favorecedoras à utilização do *mim* no contexto analisado, tais como: Florianópolis (0.89); Rio Branco (0.87); Campo Grande (0.84); Cuiabá (0.80); Porto Velho (0.76); São Luís (0.75); Curitiba (0.69); e Boa Vista (0.66). A capital Natal foi a que apresentou o menor peso relativo, (0.24). No que tange à escolaridade, um dos grupos selecionados pelo programa estatístico, na fala dos informantes com o nível

fundamental, quanto ao peso relativo, o pronome *mim* como regra de aplicação apresentou (0.79), ao passo que os informantes universitários (0.17).

Ao concluir o estudo, os autores apontam para a ampliação do estudo, seja através de pesquisas com dados do interior do país, seja atinente à sociohistória das localidades, bem como o controle sistemático do zero anafórico, para que, adicionados aos estudos existentes, haja um fortalecimento para o respeito às diversas formas de falar, sobretudo as que não estão previstas e não gozam de privilégio social e escolar.

Ao comparar as três obras monográficas e os três estudos, notam-se aspectos que se assemelham, uma vez que, contrariando a norma prescritiva, o pronome *mim* figura na fala dos brasileiros. Além disso, os resultados dos últimos três estudos se aproximam, pois, em todos eles, o pronome oblíquo *mim*, no contexto sintático analisado, apresenta percentuais superiores ao pronome *eu*, previsto pela norma tradicional. De acordo com Santos e Paim (2018), a escolaridade é um fator importante, quando se observa o fenômeno em questão. Martins (2019) destaca o Nordeste do Brasil como uma área favorecedora para utilização do *eu*. Por sua vez, Santos e Ribeiro (2020), conforme dados estatísticos do programa utilizado, afirmam que a escolaridade é fator de relevância, bem como a diatopia ser também um fator para identificação de áreas de presença do fenômeno estudados nesta tese com indícios de manutenção da norma de prestígio ou a veiculada na escola.

4.2 ESTUDOS SOBRE O PORTUGUÊS BRASILEIRO, PORTUGUÊS EUROPEU E O PORTUGUÊS AFRICANO

Lucchesi (2008), argumentando sobre os aspectos gramaticais brasileiros que foram afetados pelo contato entre línguas, afirma que, com base na teoria da Transmissão Linguística Irregular, conforme citado na subseção 3.1.1, os processos de variação e mudança desencadeados pelo contato entre línguas podem resultar em casos mais radicais, resultando o surgimento de línguas crioulas, e casos mais leves, resultando uma nova variedade de uma língua histórica, e exemplifica com o português popular do Brasil.

No que tange à flexão de caso dos pronomes pessoais, o referido autor compara as variedades do português culto, afro-brasileiro e cabo-verdiano, sintetizando-os em quadros:

Quadro 19 – Flexão de caso na norma culta do PB, conforme Lucchesi (2008)

Flexão de caso dos pronomes na norma culta do português brasileiro			
Pessoas	Função sintática		
	Sujeito	OD/OI	Oblíquo/adverbial
1ª pessoa do sing.	eu	me	mim/comigo
2ª pessoa do sing.	você	você ~ te ~ lhe	você, contigo
3ª pessoa do sing.	ele/ela	ele/ela ~ o/a ~ lhe	ele/ela ~ si, consigo
1ª pessoa do sing.	a gente ~ nós	a gente ~ nos	a gente ~ nós, conosco
2ª pessoa do sing.	vocês	vocês ~ lhes	vocês ~ si, consigo
3ª pessoa do sing.	eles/elas	eles/elas ~ os/as ~lhes	eles/elas ~si, consigo

Fonte: Lucchesi (2008, p. 382).

Segundo o referido autor, a configuração do quadro pronominal do português culto difere do português popular brasileiro e afro-brasileiro, uma vez que essas variedades apresentam processos de variação peculiares, tal como a variação na flexão de caso com o pronome da 1ª pessoa do singular: “Nessas variedades do PB, nem o mais pronominal dos pronomes, *eu*, escapa do processo de variação, podendo a forma do caso reto ocorrer na função de objeto direto [...], bem como na função de oblíquo ou adverbial...” (LUCCHESI, 2008, p. 383). E, ainda, posteriormente, para o português afro-brasileiro, propõe a sistematização exposta no Quadro 20.

Quadro 20 – Flexão de caso no português afro-brasileiro

Os pronomes pessoais no português afro-brasileiro				
	Função sintática			
	Sujeito	OD	OI	Obliquo/adverbial
1 ^a pessoa do sing.	eu	me ~ eu	me ~ para mim	de/ni/pra mim ~ de/ni eu ~ comigo ~ mais eu
2 ^a pessoa do sing.	você/tu	te ~ lhe ~ você ~ tu	te ~ lhe ~ a você	pra/com/ de você ~ com tu
3 ^a pessoa do sing.	ele/a	ele/a	(pra) ele/ela	de/para/ni/com ~ mais ele/a
1 ^a pessoa do sing.	nós ~ a gente	a gente ~ nos	pra /a gente ~ nós	pra/de/com ~ mais a gente ~ nós
2 ^a pessoa do sing.	vocês	–	pra vocês	pra/mais vocês
3 ^a pessoa do sing.	eles/as	elas/eles	pra (a) eles	de/ni/para/com ~ mais eles/as

Fonte: Lucchesi (2008, p. 383).

A erosão da morfologia de caso dos pronomes pessoais, segundo o autor, é resultado de um processo de uma TLI. Para os casos crioulos, não há morfologia de caso, mas existe uma variação fonológica, por meio de cliticização de formas pronominais, que podem ser sintetizadas assim:

Quadro 21 – Os pronomes pessoais do português de Cabo Verde

Os pronomes pessoais do crioulo português de Cabo Verde		
Pessoa	Sujeito	Complementos e Adjuntos Adverbiais
1 ^a pessoa do sing.	mi ~ m ~ i	me ~ mi ~ ‘m
2 ^a pessoa do sing.	bô ~ bó ~ bu	bô ~ bu ~ b’
3 ^a pessoa do sing.	êl ~ ê	l ~ êi
1 ^a pessoa do pl.	no ~ nu ~ nos	nos ~ nus
2 ^a pessoa do pl.	bosês ~ osês ~ sês	bô
3 ^a pessoa do pl.	ês	ês ~ ‘z

Fonte: Lucchesi (2008, p. 384).

Ainda, conforme postulações do referido autor, em outros crioulos de base lexical portuguesa – Guiné-Bissau, o crioulo; na ilha de São Tomé, o forro e angolar; na ilha de Príncipe, o principense –, apresenta-se, praticamente, o mesmo quadro de flexão pronominal. Desse modo, ao comparar a variedade popular do PB com as línguas crioulas, há de se levar em consideração que:

[...] Portanto, nos casos mais radicais de transmissão lingüística irregular, como os que se observam nas línguas crioulas, essas estruturas são eliminadas; nos processos menos intensos, tais estruturas são acometidas de um amplo processo de variação, sem, contudo, serem eliminadas, como ocorre nas variedades populares do português brasileiro (LUCCHESI, 2008, p. 385).

Logo, o processo de erosão de flexão casual dessas variedades assemelha-se até certo ponto, mas também se difere pela intensidade em que essa erosão se deu. No português popular brasileiro, o processo evidencia um quadro de profunda variação, ao passo que, nas variedades crioulas, o processo desencadeou flexões sem marcas de valor referencial, por exemplo, a flexão de caso dos pronomes pessoais, a morfologia verbal de pessoa e número, bem como a concordância nominal de gênero e número.

Lucchesi e Mendes (2009), ao analisarem a flexão de caso dos pronomes pessoais, comparam o PB, o PE e o PA, a fim de verificar em qual grau essas variedades de português foram afetadas, quanto à erosão da morfologia flexional. Os referidos autores apontam que o quadro pronominal brasileiro, no que tange à norma culta, encontra-se em um amplo processo de variação por duas principais razões, substituição dos pronomes canônicos por expressões nominais e fatores estruturais paradigmáticos. No que tange ao português afro-brasileiro, afirmam que

Na 1ª pessoa do singular, a forma do caso reto encontra-se em variação com as formas oblíquas nas funções de: OD (e.g., meu pai me criô / o marido num quis eu não), complemento oblíquo (e.g., todo mundo depende de mim / eles depende de eu) e adjunto adverbial (e.g., tu vem ficá aqui comigo / ela num quis ir mais eu não) [...] A forma do caso reto eu não se documenta na amostra de fala analisada na função de OI, registrando-se apenas as formas me e mim (e.g., me dá meu ‘mendum; deu pra mim). Entretanto, apesar de não ter se constatado na amostra de fala analisada, pode-se afirmar que uma frase como ela entregou pra eu é possível, tanto na fala afro-brasileira, quanto na fala popular rural. Portanto, as formas flexionadas em caso da 1ª pessoa do singular (me, mim e comigo) sofrem, na fala das comunidades rurais afro-brasileiras, a variação com a forma não marcada do caso reto, em todas as posições sintáticas, mesmo que o nível dessa variação seja baixo nos dias atuais (LUCCHESI; MENDES, 2009, p. 481).

Ao observarem o português santomense, os autores apresentam o quadro dos pronomes que demonstra a erosão da morfologia flexional, em que subsiste uma variação de cunho fonológico advinda de processos de clitização do pronome, conforme Quadro 22.

Quadro 22 – Os pronomes pessoais do português em Santo Tomé.

Os pronomes pessoais do crioulo português em Santo Tomé		
Pessoa	Sujeito	Objeto
1 ^a pessoa do sing.	n ~ i ^N	mu ⁿ ~ mu ~ m
2 ^a pessoa do sing.	bo	Bo
3 ^a pessoa do sing.	e ~ ele	e ~ ele
1 ^a pessoa do pl.	nõ	Nõ
2 ^a pessoa do pl.	nãse	Nãse
3 ^a pessoa do pl.	iné ^N ~ né ^N ~ inê	iné ⁿ ~ né ⁿ ~ inê

Fonte: Lucchesi; Mendes (2009, p. 479).

Para as variedades do português em África, crioulos de base lexical portuguesa, caboverdiano, do crioulo da Guiné-Bissau e do crioulo de São Tomé, os referidos autores apontam que “[...] o processo de transmissão linguística irregular mais radical que conduziu à formação dos crioulos portugueses da África implicou a eliminação da flexão de caso dos pronomes pessoais...” (LUCCHESI; MENDES, 2009, p. 481).

Pena-Ferreira (2015) coteja o PE com o PB, a partir de *corpora* de textos escritos e orais, analisando os diferentes usos de pronomes retos na função de complemento e usos de oblíquos na função de sujeito, que, de certa forma, são vistos como um traço que diferencia o PB. Para a análise dos textos escritos, a autora utilizou textos do: *Corpus* do português (2006); *Corpus* Mínimo de Textos Escritos em Língua Portuguesa (2006) e o Projeto *Forgotten Letters Years* (2008). Para a análise dos textos orais, foram retirados os dados do *Corpus* de Referência do Português Contemporâneo e do *Corpus* de Textos Orais do Português Santareno. Vale ressaltar que, além dos dados orais do PE e do PB, a autora analisou, também, de outras variedades do português: Angola, Cabo Verde, Guiné-Bissau, Moçambique, São Tomé e Príncipe, Goa, Macau e Timor-Leste.

A autora evidencia os resultados encontrados para as mudanças morfossintáticas que ocorreram com a inserção das formas *a gente*, para a 1^a pessoa do plural, e *você*, 2^a pessoa do singular, em todas as variedades do português que foram analisadas. No que tange ao pronome *mim*, ela destaca que:

Cruz (1991) observou que **em Odeleite o pronome pessoal complemento é usado por vezes com função de sujeito: (23) “... dá pra mim guardari”**. O caso contrário, o uso do pronome sujeito com função de complemento também se observa por vezes: (24) ‘A gente agora vai cercando elas’. (PENA-FERREIRA, 2015, p. 36) [grifos nossos].

Ademais, além de a alternância ser observada na fala popular de Odeleite, em Portugal, nota-se também que, em Angola, o pronome oblíquo tônico *mim* é utilizado como sujeito, precedido por preposição e antes de verbo no infinitivo, tal como: **“eu tive de sair mesmo da minha família e arranjar um cantinho que é para mim poder viver sozinho, trabalhar normalmente, para mim poder sobreviver”** (PENA-FERREIRA, 2015, p. 37, grifo nosso). Ao citar Medeiros (1964), a autora afirma que, nos Açores, usos semelhantes são encontrados como, por exemplo, na Ilha de São Miguel: “p’a mim comê(r); p’a mim me desembaraça(r)” (PENA-FERREIRA, 2015, p. 37). Desse modo, uma das conclusões a que se chegou o estudo foi:

Os dados apresentados revelam que o uso do pronome do caso reto em função de objeto e o uso do pronome oblíquo em função de sujeito, tidos como características do PB, ocorrem nas demais variedades da língua portuguesa, sendo, portanto, um traço peculiar a essa língua e não a uma variedade específica (PENA-FERREIRA, 2015, p. 37).

Após as análises empreendidas, para os retos na função de complemento, 136 ocorrências foram encontradas, ao passo que, para os oblíquos na função de sujeito, apenas 38 ocorrências. Dessas últimas, o uso fica restrito às formas de 1ª e 2ª pessoas do singular, *mim* e *ti*. Em relação ao pronome *mim* em posição de sujeito, essa forma só é encontrada em construções [preposição para + pronome + verbo no infinitivo], assim:

O fato de o verbo da segunda oração estar reduzido à forma nominal infinitiva contribui, acreditamos, para que o falante não considere necessária a presença de um sujeito expresso, como, talvez, acontecesse se o verbo estivesse flexionado na forma finita. Não encontramos, por exemplo, construções em que o pronome *mim* substituísse o pronome *eu* iniciando sentenças, como em * ‘*mim não quero mais nem contato com Ele*’. No discurso espontâneo, por menos escolarizado que seja o falante, esse uso não ocorre, o que nos faz pensar que a substituição do pronome oblíquo pelo pronome reto, na 1ª pessoa, esteja condicionada à presença da preposição *para* e de verbo no infinitivo. A forma [para mim] parece ser uma espécie de construção pré-fabricada³³ (PENA-FERREIRA, 2015, p. 37).

Nessa perspectiva, por analogia, o pronome oblíquo é sempre precedido por preposição e, para os falantes, mesmo em orações reduzidas de infinitivo, a forma pronominal *mim* é encaixada. Em virtude do que foi exposto, a referida autora afirma que há alternância entre os pronomes retos na função de complemento e dos pronomes oblíquos na função de sujeito. Segundo ela, são fenômenos que vêm acontecendo na língua, em todas as variedades do português, não somente na variante popular do PB.

³³ Para Bybee (2010), são qualquer expressão convencionalizada composta por duas ou mais palavras, são combinações alternativas que são utilizadas por falantes nativos.

A partir dos estudos mencionados nesse subtópico, Lucchesi (2008); Lucchesi e Mendes (2009); e Pena-Ferreira (2015), nota-se que o fenômeno ora estudado nesta tese não é exclusivo do PB, uma vez que acontece nas outras variedades do português, conforme ilustram os respectivos autores. As perspectivas teóricas são vistas como importantes, uma vez que tentam elucidar os aspectos inerentes aos pronomes, em especial, aos de primeira pessoa do singular. A TLI, de forma leve, para o PB, pode ser aventada como possível explicação para a erosão na morfologia de flexão pronominal. No entanto, também, o “para/pra + mim” faz parte de uma construção pré-fabricada. A diversidade teórica, por vezes, permite um sincretismo que possa explicar fenômenos.

4.3 SÍNTESE DA SEÇÃO

Esta seção teve a função de expor a revisão teórica específica sobre o fenômeno em estudo, a variação entre os pronomes *eu* e *mim*, em posição de sujeito, em orações infinitivas, precedidas pela preposição *para*, no que tange aos estudos que se debruçaram sobre o comportamento do fenômeno no PB.

Na subseção 4.1, nota-se que tal revisão inicia-se por meio nove gramáticas com fins teóricos/metodológicos diferentes, apresentando, também, perspectivas e ângulos plurais dos ramos da Linguística, tais como: a Gerativa, a Sociocognitiva, a Sociolinguística e a Dialetoлогия. Na subseção 4.2, são apresentados estudos que cotejaram o fenômeno em variedades do PB contrastando com outras variedades, tais como o PE e o PA.

Tais subseções revelam aspectos importantes, uma vez que atende a um dos objetivos traçados, apresentar uma revisão de literatura sobre o fenômeno. No que tange às gramáticas, há perspectivas, ainda, que preconizam a um quadro pronominal canônico e estanque, tais como em Bechara (2010) e Lima (2018), ao passo que as gramáticas históricas, descritivas e pedagógicas/escolares convergem para quadros pronominais flexíveis, em que o pronome *mim* pode exercer outras funções para além das estabelecidas prototipicamente.

Na seção seguinte, 5, são apresentados dois grandes projetos sobre a língua falada no Brasil, Projeto NURC e Projeto ALiB.

5 TRAJETÓRIAS DO PROJETO NURC E DO PROJETO ALiB

Não é portanto, desanimadora, de modo algum, a avaliação que se pode fazer sobre o conhecimento atual do português do Brasil embora, realisticamente, ainda se tem que afirmar não que está tudo por fazer, mas que há muito a ser feito e as dificuldades são muitas sobretudo porque pesquisas da natureza das anteriormente referidas exigem muitos pesquisadores bem preparados e financiamentos de peso e todos nós sabemos que não são de extração fácil, mas até mesmo muito difícil, os financiamentos para pesquisas que não apresentem rentabilidade imediata .

(SILVA, 1994, p. 217).

Nesta seção, serão expostos, de forma breve, alguns aspectos dos dois *corpora* – Projeto NURC e do Projeto ALiB – a fim de demonstrar a relevância desses projetos para o estudo da língua falada no Brasil, além de poder evidenciar as distintas metodologias deles e como se chegou a metodologia adotada na tese. Destacam-se as inéditas propostas de periodização para os dois projetos.

5.1 O PROJETO DE ESTUDO DA NORMA LINGUÍSTICA URBANA CULTA NO BRASIL – PROJETO NURC

Quando se pensa em estudos sobre a língua oral no Brasil, não se nega a importância do projeto pioneiro, surgido nos fins da década de 60, com objetivos voltados à investigação empírica e à descrição da variedade culta da língua falada, mas também que visavam ao ensino. Por esse motivo, “Considero o Projeto NURC como um divisor de águas na questão da(s) norma (s) linguística(s) culta(s) no Brasil, porque através dele, nas camadas hegemônicas universitárias, introduziu-se uma nova visão, uma nova ideologia enfim, quanto ao respeito à diversidade de normas no processo de ensino...” (MATTOS E SILVA, 1995, p. 44).

Sendo assim, alvorece o Projeto de Estudo da Norma Linguística Urbana Culta no Brasil – Projeto NURC – inspirado em um projeto sobre o estudo do espanhol falado na América de língua espanhola³⁴. O Projeto NURC ganha corpo em 1969, com grupos de trabalho espalhados pelas cinco capitais alvos da pesquisa, a saber: Porto Alegre (POA), Rio de Janeiro (RJ), São Paulo (SP), Salvador (SSA) e Recife (REC). Na época, as gravações

³⁴ Um pouco mais da história e do surgimento do Projeto NURC podem ser encontrados no livro NURC 50 ANOS - 1969 – 2019, organizado por Miguel Oliveira Jr.

foram feitas em fitas magnéticas profissionais de rolo, pois era um material que garantia uma qualidade na gravação do som, embora seus reprodutores fossem equipamentos caros.

Os informantes participavam da pesquisa por meio de um dos três tipos de gravação previstos pela metodologia adotada pelo Projeto, quer sejam: Elocução em situação Formal (EF); Diálogo entre informante e documentador (DID); e Diálogo entre dois informantes (D2). Interessava ao NURC a fala de informantes oriundos dessas cinco capitais, que foram estratificados em sexo – homem e mulher –, faixa etária – I - de 25 a 35 anos; II - de 36 a 55 anos; III - de mais de 56 anos – e todos deveriam possuir formação universitária completa, esses, conforme pensamento dos pesquisadores da época, representavam a “fala culta”. Assim,

[...] a norma culta era objeto do projeto Norma Urbana Culta – NURC, que, embora não tenha seguido a orientação metodológica da sociolinguística variacionista, possibilita o controle de aspectos estilísticos (diferentes tipos de coleta em função da atenção à fala, como as elocuições formais, o diálogo entre entrevistador e informante e o diálogo entre informantes) e sociais/dialetais... (FREITAG, 2016, p. 453).

Nesse sentido, vale ressaltar que

A complexidade social brasileira se reflete na complexidade sociolinguística do Brasil, e os limites empíricos entre o que é “fala culta” e “fala popular” se tornam cada vez mais imprecisos e fluidos, quando não inexistentes, embora ainda se conservem como elemento ideológico estruturante de uma sociedade marcada pelas formas mais cruéis de injustiça. Por tudo isso é que o corpus do NURC, material precioso, retrata uma “norma urbana culta” já em boa medida ultrapassada, ainda que revele uma distância larga e funda com relação ao que era previsto na tradição gramatical normativa. Seu valor como documentação de uma importante fase histórica da constituição do português brasileiro contemporâneo permanece, apesar de tudo, intocável... (BAGNO, 2019, p. 17).

Como é consabido, o expressivo acervo do Projeto NURC registra falas de informantes nas cinco capitais pesquisadas, vem servindo de material de pesquisa para diversos pesquisadores, a fim de proporcionar retratos do português falado no Brasil, embora seja de um grupo específico da população, das décadas de 70 e 90. Nessa direção,

[...] é potencialmente a amostra linguística oral diacrônica mais antiga do português brasileiro; e, novamente, se seguido o protocolo de documentação, a comparação entre a amostra original e a nova amostra permite a realização de estudos de mudança linguística em tempo real. Além disso, o NURC permite a observação dos efeitos de monitoramento, contribuindo para a descrição estilística (FREITAG, 2019, p. 133).

Além das inúmeras teses, dissertações, monografias de especialização e trabalhos de conclusão de curso³⁵ elaborados a partir do banco de dados do Projeto NURC no Brasil,

³⁵ Um levantamento detalhado com alguns trabalhos pode ser visto em Mota (2019, p. 76-79).

destacam-se as publicações das transcrições grafemáticas, que compõem o *corpus* compartilhado³⁶, conforme Quadro 23:

Quadro 23 – Publicações da coleção do Projeto NURC

CIDADE	ANO	AUTOR	OBRA	MATERIAL
São Paulo SP	1986	Ataliba de Castilho; Dino Preti.	A linguagem falada culta na cidade de São Paulo. Materiais para seu estudo.	EF
	1987	Ataliba de Castilho; Dino Preti.	A linguagem falada culta na cidade de São Paulo. Materiais para seu estudo.	D2
	1990	Dino Preti; H, Urbano.	A linguagem falada culta na cidade de São Paulo. Materiais para seu estudo.	DID
Rio de Janeiro RJ	1992	Dinah Callou.	A linguagem falada culta na cidade do Rio de Janeiro. Materiais para seu estudo.	EF
	1993	Dinah Callou; Célia Lopes.	A linguagem falada culta na cidade do Rio de Janeiro. Materiais para seu estudo.	DID
	1994	Dinah Callou; Célia Lopes.	A linguagem falada culta na cidade do Rio de Janeiro. Materiais para seu estudo.	D2
Salvador SSA	1994	Jacyra Motta; Vera Rollemberg.	A linguagem falada culta na cidade de Salvador. Materiais para seu estudo.	DID
	2006	Jacyra Motta; Vera Rollemberg.	A linguagem falada culta na cidade de Salvador. Materiais para seu estudo.	EF
	2011	Jacyra Motta; Marcela Paim; Vera Rollemberg.	A linguagem falada culta na cidade de Salvador. Materiais para seu estudo.	D2
Recife REC	1996	M ^a da Piedade Sá entre outros.	A linguagem falada culta na cidade de Recife. Materiais para seu estudo.	DID
	2005	M ^a da Piedade Sá entre outros.	A linguagem falada culta na cidade de Recife. Materiais para seu estudo.	EF
	2017	M ^a da Piedade Sá entre outros.	A linguagem falada culta na cidade de Recife. Materiais para seu estudo.	D2

³⁶ Para que em todas as cinco capitais brasileiras que estavam envolvidas na pesquisa tivessem cópias dos materiais coletados, foi desenvolvido o que se denominou “*corpus* compartilhado”. Um conjunto composto de 90 entrevistas: “18 inquéritos de cada capital, sendo três elocuições formais, dez diálogos entre informante e documentador e cinco diálogos entre dois informantes” (MOTA, 2019, p.60).

Porto Alegre POA	1997	José Hilgert	A linguagem falada culta na cidade de Porto Alegre. Materiais para seu estudo.	DID
	2007	José Hilgert	A linguagem falada culta na cidade de Porto Alegre. Materiais para seu estudo.	EF
	2009	José Hilgert	A linguagem falada culta na cidade de Porto Alegre. Materiais para seu estudo.	D2

Fonte: Elaborado pelo autor (2023).

Com base nessas produções, disponíveis e de fácil acesso, e tendo em vista as contribuições dos *corpora* do Projeto NURC para o entendimento da língua falada em solo brasileiro, ressalta-se que

[...] no Brasil, identificamos a implementação dos estudos sociolinguísticos, a partir dos desdobramentos dos estudos dialetais. O Projeto NURC, a esse propósito, é considerado um expoente não somente por ter sido a partir dele que se fundaram preocupações variacionistas no país, mas principalmente por ter fomentado a configuração de outros projetos [...]. Atualmente, muitos desses projetos continuam operantes e são responsáveis pelo entendimento da situação sociolinguística do português do Brasil (MOTA; SILVA, 2019, p. 30).

Após meio século de seu surgimento, um banco de dados com bastante vigor e que fomenta o interesse de diversos pesquisadores, sobretudo, dialetólogos e sociolinguistas, o NURC “[...] conhece agora uma nova etapa de sua história cinquentenária: sua transformação em dados digitalizados segundo as mais avançadas técnicas de tratamento informatizado de corpus...” (OLIVEIRA JR., 2019, p.17). O NURC Digital já é uma realidade, a partir da iniciativa do grupo de pesquisadores do Projeto NURC/Recife³⁷.

5.1.1 Esboço de uma periodização do Projeto de Estudo da Norma Linguística Urbana Culta – Projeto NURC

Ao considerar a história e a representatividade do Projeto de Estudo da Norma Linguística Urbana Culta – Projeto NURC – para os estudos dialetológicos e sociolinguísticos brasileiros, quando se vislumbram esses dois ramos dos estudos linguísticos, nota-se que o referido Projeto assume o pioneirismo e a vanguarda, assim, se constituindo o primeiro banco de dados com amostras de fala do país. Cabe ressaltar, mais uma vez, as incontáveis pesquisas acadêmicas que tomaram por base o acervo desse grande empreendimento, quer sejam monografias, dissertações ou teses.

³⁷ O Portal NURC digital encontra-se disponível em <https://fale.ufal.br/projeto/nurcdigital/>.

Nesse sentido, hoje, sugere-se pensar o Projeto NURC em fases. Lembrando que, muito se distancia da pretensão de uma periodização fixa e categórica, mas sim aberta, flexível e passível de modificações. Com isso, propõe-se, então, uma segmentação estabelecida em cinco fases, a saber: i. 1ª fase; ii. 2ª fase; iii. 3ª fase, iv. 4ª fase e v. 5ª fase, conforme ilustra o Quadro 24, por meio de marcos temporais e, respectivos, marcos alusivos. Um aspecto que merece ser ressaltado é que os limites entre as fases não são estanques, pois há momentos em que os fatos e marcos entre o final de uma fase e início de outra se fundem, há também as fases que são contínuas, por exemplo, as 3ª e 4ª fases, pois as obras podem ser publicadas a qualquer tempo. Nesse sentido, portanto, busca-se, com essa proposta, destacar os objetivos nucleares e prototípicos de cada fase, pois eles não se anulam, se complementam.

Quadro 24 – Proposta de periodização do Projeto da Norma Urbana Linguística Culta – Projeto NURC³⁸

FASES	MARCOS	
	Início	Término
1ª	1968, jan. Relatório apresentado ao IV Simpósio do Programa Interamericano de Linguística e Ensino de Idiomas (PILEI).	1969 Implantação do Projeto NURC no Brasil.
2ª	1970 Início da fase de constituição do <i>corpus</i> .	1978 Final da constituição do <i>corpus</i> .
3ª	1986 Publicação da primeira obra da coleção NURC - São Paulo, Elocuções Formais.	2003 Publicação do vol. 3 ³⁹ da coleção NURC – Salvador Diálogos entre dois informantes
4ª	1990/1991 Início da coleção	2022 Gramática do Português

³⁸ Essa proposta foi reformulada, após a defesa da tese, a partir das sugestões feitas pela Profa. Dra. Jacyra Andrade Mota, integrante do Projeto NURC-Ssa e membro da banca.

³⁹ O Volume IV da coleção NURC - Salvador encontra-se em vias de publicação.

	Gramática do Português falado, vol. I.	falado, vol. VIII.
5 ^a	2016 Publicação do site do NURC Digital.	

Elaborado pelo autor (2023).

A 1ª fase compreende o início do Projeto NURC, em 1968 – a partir do relatório apresentado pelo professor Nelson Rossi ao IV Simpósio do Programa Interamericano de Linguística e Ensino de Idiomas (PILEI). A pretensão inicial brasileira era ser uma extensão do *Proyecto de Estudio Coordinado de la Norma Lingüística Culta de las Principales Ciudades de Iberoamérica y de la Península Ibérica*, composto por alguns países de língua espanhola na América Latina, representados por algumas cidades – Bogotá, Buenos Aires, Caracas, Havana, Lima, Madrid, México, Puerto Rico e Santiago do Chile – e se estende até os anos 1969, por meio da implementação do Projeto NURC no Brasil.

Destacam-se, também, nessa primeira fase, inicialmente, escolha dos coordenadores dos trabalhos nas cinco capitais, e, posteriormente, formação das equipes; reuniões com os coordenadores, com produção de atas dos acordos firmados para a solificação da empreitada. Foi um momento de fundamentar as bases, a fim de construir o primeiro banco de dados sobre a língua falada no Brasil, o que hoje é considerado o testemunho mais antigo de língua, no que tange à base empírica, em amostras orais a nível nacional.

A 2ª fase compreende de 1970 – com o início da fase de coletar dados, por meio das entrevistas, momento de composição do *corpus*, por meio das entrevistas feitas nas cinco capitais – e se estende até 1978 – momento em que, para as amostras dos anos 70, se finda a fase de constituição do *corpus*, em Salvador.

Como toda pesquisa *in loco*, cabe trazer à tona aspectos importantes para a composição de um banco de dados linguístico. Nessa direção, no que tange ao cenário político, o Brasil vivia um fase de ditadura cívil-militar. Logo, as censuras ditavam os comportamentos e, por vezes, os discursos e narrativas.

A 3ª fase compreende o período de 1986 – momento em que a equipe de São Paulo publica a primeira obra da série *A linguagem culta falada em São Paulo: Materiais para o seu estudo*. Essas obras possuem datas de publicação diversas, com a primeira delas feita pela equipe do NURC/São Paulo, em 1986, Elocuções Formais. Ressalta-se a importância da

referida série e não se descartam as contribuições advindas dela para o entendimento da língua falada no Brasil, embora o recorte dos dados se restringisse a algumas capitais e a uma parcela de brasileiros, conforme critérios metodológicos já mencionados, informantes com ensino universitário, oriundos de cinco capitais brasileiras – e se estende até 2003 – momento em que a equipe do NURC/Salvador publica o volume 3 da referida série.

Incluem-se, na terceira fase, o período de recontato de informantes, pois ele visa uma ampliação do *corpus*, objetivando, a partir dos resultados, construir uma análise em tempo aparente - a partir do comportamento dos locutores por faixas etárias – com o resultado das análises em tempo real, aos moldes sociolinguísticos. Nesse sentido, as novas entrevistas também são inclusas, haja vista que retratam a língua falada no Brasil, por informantes universitários, na década de 90.

Ademais, há diversas obras publicadas pela equipe NURC/SP, que dizem respeito a *Projetos Paralelos*, com base no material do NURC, a título de exemplo, cita-se a obra *Análise de Textos Orais* (PRETI, 2001).

A 4ª fase inicia-se, no início dos anos 90, em 1990/1991 - com a publicação da coleção Gramática do Português falado (coleção iniciada e coordenada por Ataliba T. de Castilho), que, tomando por base o *corpus* do Projeto NURC, que visa a uma gramática de referência que espelha a variante culta do português falado no Brasil – e se estende até 2022, com a publicação do volume 8 dessa coleção (organizado por Maria Bernadete M. Abaurre e Angela C. S. Rodrigues).

As publicações dessa coleção, com base em dados do NURC, revelam aspectos e nuances da língua falada em vários níveis de análise da língua. Do primeiro volume ao último, até então publicado, o volume 8, há um vasto material em que se pode notar estudos e análises de um material representativo para as pesquisas sobre a língua falada no Brasil.

A 5ª fase inicia-se em 2016, a partir da publicação do site do NURC Digital, com dados de Recife e se estende até os dias atuais. Vale mencionar que, anteriormente, os dados do Projeto NURC/Rj foram digitalizados e disponibilizados no site⁴⁰, numa tentativa de informatização e acessibilidade ao acervo, mesmo sem levar em consideração aspectos importantes para elaboração de um banco de dados digital. Nesse sentido, embora o grupo do Rio de Janeiro tenha dado o primeiro passo, a era digital do Projeto ganha fôlego e se concretiza com a equipe da Universidade Federal de Alagoas, por meio dos dados dos recifenses. Com isso,

⁴⁰ Informações e dados disponíveis em <https://nurcrj.lettras.ufjf.br/>.

Um dos objetivos centrais do Projeto NURC Digital foi disponibilizar os dados do Projeto NURC/Recife ao público em geral, através de um site dedicado. Essa ação beneficia diretamente a comunidade científica, que passa a ter disponíveis, para consulta otimizada, dados – anteriormente de difícil acesso – em formato digital de alta qualidade, devidamente catalogados, etiquetados e transcritos. Nesse sentido, os dados foram tratados de maneira a adequá-los para diversos usos online, com busca avançada, opções de download, acesso a informações não-sensíveis etc... (OLIVEIRA JR., 2019, p. 188).

A iniciativa, audaciosa e louvável, da equipe de Recife, em tornar digitais os dados do NURC, instaurou mais uma meta para esse Projeto que possui, nesses 50 anos de existência, inegável importância para o entendimento da língua falada no Brasil. Almeja-se que tal conquista se estenda às demais capitais, a fim de preservação do extenso acervo linguístico nacional. O desejo é que recursos sejam investidos nessa empreitada, não somente recursos humanos, mas, sobretudo, financeiros, pois urge que essa tarefa seja concretizada, para salvaguardar as primeiras amostras de falas do português brasileiro em *corpus* oral diacrônico. Registra-se, com isso, a obra *NURC 50 anos – 1969 - 2019* – (OLIVEIRA JR., 2019) como importante para o entendimento e para a compreensão dessa segunda fase, mas também alusiva quanto à proposição de início de uma nova fase.

A partir do que foi exposto nessa subseção, sugere-se pensar em um alvorecer amadurecido da 5ª fase, no que tange à periodização, para o Projeto NURC. Tal fase tem a incumbência de tornar os materiais do NURC públicos e mais acessíveis, por meio das ferramentas das Tecnologias Digitais da Informatização e Comunicação – TDIC, a fim de continuar a empreitada e seguir os caminhos da equipe de Recife. Conforme Oliveira Jr., (2016) pontua:

O processo de digitalização dos dados do NURC/Recife evidenciou a fragilidade com que se encontra esse material. Dada a riqueza dos dados do Projeto NURC e a incontestável importância que esse material teve e tem para a pesquisa em linguística desenvolvida no Brasil, torna-se imperativo a adoção de práticas urgentes para garantir a sua preservação nas demais capitais em que o projeto foi sediado: Salvador, Rio de Janeiro, São Paulo e Porto Alegre. Esperamos que este artigo sirva como estímulo para isso (OLIVEIRA JR., 2016, p. 170).

Assim, as trilhas estão postas, as ferramentas da era digital também estão, porém, faltam recursos, quer sejam humanos e, sobretudo, financeiros, a fim de dar continuidade à 5ª fase do Projeto NURC, fase digital, salvaguardando em ambientes mais modernos as amostras linguísticas do PB das cidades de Salvador, Rio de Janeiro, São Paulo e Porto Alegre.

5.2 O PROJETO ATLAS LINGUÍSTICO DO BRASIL – PROJETO ALiB

No alvorecer dos anos 90, mais especificamente em 1996, o Projeto Atlas Linguístico do Brasil – Projeto ALiB – surge a partir do *Seminário Nacional Caminhos e Perspectivas para a Geolinguística no Brasil*⁴¹, que teve o objetivo principal debater o futuro para uma política geolinguística brasileira, utilizando-se dos princípios da Geolinguística Pluridimensional, isto é, que abarca o registro diatópico, mas também onde se incluem os registros diastráticos, como apresentado na secção 2.2.1.

A retomada do antigo desejo – Decreto presidencial nº 30.643/1952 e Portaria nº 536/1952 – de ter um atlas nacional foi estabelecida por convênios que foram firmados por universidades federais e estadual, configurando-se o Projeto ALiB, dessa maneira, pois, um Projeto de dimensão interinstitucional. Além disso, foi estabelecido um Comitê Nacional, composto por autores de atlas estaduais publicados e atlas em andamento; também foram estabelecidas equipes regionais de trabalho; comissões de informatização e cartografia, além de equipes de inquiridores, inquiridores auxiliares, transcritores e bolsistas de apoio técnico, bem como os bolsistas de iniciação científica, um verdadeiro exército de dialetólogos em busca do atlas.

Ao ter por meta o mapeamento nacional, abarcam-se, na rede de pontos, 250 localidades brasileiras, considerando-se a densidade demográfica de cada estado/região; a sugestão de localidades para a composição de um atlas linguístico para o Brasil, prevista em Nascentes (1958); os limites interestaduais e internacionais; bem como as características históricas e culturais. A rede de pontos, conforme Figura 22, é composta por 25 capitais⁴² e 225 cidades interioranas.

⁴¹ Realizado em Salvador, na UFBA, entre os dias 03 e 05 de novembro de 1996, com a presença de dialetólogos de todo o Brasil e do professor Michel Contini.

⁴² Na composição da rede de pontos, segundo critérios pré-estabelecidos, ao considerar o período de fundação de cada capital, excluem-se do *corpus* as capitais do Estado do Tocantins, Palmas, e do Distrito Federal, Brasília, contemplando-se todas as demais.

Figura 22 – Rede de Pontos do Projeto ALiB



Fonte: Cardoso et al. (2014b).

De maneira sistemática, os informantes estão distribuídos em duas faixas etárias – faixa I, composta por informantes de 18 a 30 anos; e a faixa II que possui informantes de 50 a 65 anos – em dois sexos (homem e mulher) e com nível de escolaridade – fundamental incompleto (localidades do interior e capital) e universitário completo (capitais), perfazendo um total de 1100 brasileiros nativos que forneceram suas falas para obtenção dos dados, são oito pessoas por capital e quatro por localidade do interior.

Geralmente, as entrevistas duram, em média, 3h. Os avanços tecnológicos foram importantes auxílios para as pesquisas, uma vez que possibilitaram a gravação em meio magnético e/ou digital.

No que tange à versão atual do questionário, o Questionário ALiB (2001), utilizado nos inquéritos, é oriundo de duas versões anteriores, feitas em 1998 e 2000, e se subdivide em partes distintas: são sete partes, que buscam retratar várias perspectivas de estudos, tais como o fonético, o semântico-lexical, o morfossintático, o pragmático e o discursivo, como pode ser visto no Quadro 25.

Quadro 25 – Subdivisão do Questionário ALiB (2001)

Questionário	Nº de perguntas
QFF – Questionário Fonético-Fonológico	159 com mais 11 de prosódia
QSL – Questionário Semântico-Lexical	202
QMS – Questionário Morfossintático	49
QP – Questão de Pragmática	04
TDS – Temas para Discurso Semidirigido	04
PM – Perguntas Metalinguísticas	06
LE – Texto para Leitura	Parábola dos Sete Vimes / Texto adaptado.

Fonte: Elaborado pelo autor (2023).

O Questionário Fonético-Fonológico (QFF) contém questões que se voltam à apuração de fatos fônicos segmentais, tais como: QFF 04 – “Como se chama aquele aparelho onde se pode ver novela, jogo, programas...?” e QFF 46 – “Como se chama um bichinho que voa e tem as asas bonitas e coloridas?” (COMITÊ NACIONAL..., 2001). Além das questões de prosódia, a fim de verificar aspectos entonacionais, compostas por situações que levam o informante a fazer declarações de frases interrogativas, frases afirmativas e frases imperativas, respectivamente, por exemplo: “Se você/ o (a) senhor (a) quer saber se alguém **vai sair hoje**, como é que você / o (a) senhor (a) pergunta?”; “Se você / o (a) senhor (a) quer dizer a algumas pessoas que estão presentes que você / o (a) senhor (a) **está muito feliz com o resultado do trabalho**, como é que você o (a) senhor (a) diz?” e, “Se você/ o (a) senhor (a) quer chamar muitos **meninos** que estão reunidos para que **venham almoçar**, como é que você/ o (a) senhor (a) diz?” (COMITÊ NACIONAL..., 2001).

O Questionário Semântico-Lexical (QSL) busca apurar a variação lexical e semântica dos itens no nível da palavra e do significado, está subdividido em 14 campos semânticos, a saber: acidentes geográficos, fenômenos atmosféricos, astro e tempo, atividades agropastoris, fauna, corpo humano, ciclos da vida, convívio e comportamento social, religiões e crenças, jogos e diversões infantis, habitação, alimentação e cozinha, vestuário e acessórios e vida urbana. Por exemplo, o QSL 08 – “Como se chama um clarão que surge no céu em dias de chuva?”; QSL 049 – “Como se chama o grão coberto por uma casquinha dura, que se come assado, cozido, torrado ou moído?” e QSL 136 – “Como se chama a pessoa que fala demais?” (COMITÊ NACIONAL..., 2001).

O Questionário Morfossintático (QMS) é composto por questões relativas à apuração de elementos gramaticais no nível da palavra e da sentença. Em sua subdivisão interna,

encontram-se áreas voltadas para a apuração da variação ou mudança em: artigo, substantivo, adjetivo, pronome, inclusive a QMS 23, alvo da pesquisa aqui empreendida, verbo e advérbio. Algumas questões são: QSM 01 – “Tem filhos / irmãos? Como se chamam? O que eles fazem?”; QMS 09 – “Se, na Presidência da República, estivesse uma mulher, ela seria o quê?”; QMS 26 – “O que vocês fazem no fim-de-semana?”; QMS 37 – “O que você / o (a) senhor (a) fez ontem (de diferente)?”; e QMS 048 – “Você / o (a) senhor (a) já viu disco voador, não é?” (COMITÊ NACIONAL..., 2001).

As Questões de Pragmática (QP) são referentes à observação e análise do comportamento linguístico dos informantes em algumas situações específicas, mais precisamente em como se dá a relação ou o modo de tratamento interpessoal, por exemplo: “Um objeto (carteira, lenço, chave) caiu do bolso de um rapaz jovem e ele não viu. Como um outro rapaz jovem chama a atenção desse rapaz?”; e “Um objeto (carteira, lenço, chave) caiu do bolso de uma mulher idosa e ela não viu. Como um outro rapaz jovem chama a atenção dessa mulher?” (COMITÊ NACIONAL..., 2001).

Nos Temas para Discurso Semidirigido (TDS), os inquiridores direcionam os informantes para elocuições mais livres, menos monitoradas, por meio de temas dirigidos como: Relato pessoal: “Relate um acontecimento marcante em sua vida (casamento, namoro...); Comentário: “De que programas de televisão você / o (a) senhor (a) gosta mais? Por quê?”; Descrição: “Você / o (a) senhor (a) trabalha em quê? Fale um pouco sobre o seu trabalho?”; e Relato não pessoal: “Conte um caso / um fato de seu conhecimento (de que tenha ouvido falar, que tenha acontecido com um amigo, etc)” (COMITÊ NACIONAL..., 2001).

As Perguntas Metalinguísticas (PM) são perguntas que se destinam à apuração das percepções dos informantes sobre a língua que falam, bem como as suas avaliações sobre a fala de outras pessoas: “Como se chama a língua que você / o (a) senhor (a) fala?”; “Poderia dar um exemplo do modo como falam essas pessoas que “falam diferente”?”; e “No passado, falavam diferente aqui?” dentre outras (COMITÊ NACIONAL..., 2001). Além disso, por fim, há um Texto para a Leitura, “A Parábola dos Sete Vimes”, quando se buscam algumas variáveis que estão distribuídas por todo o texto (COMITÊ NACIONAL..., 2001).

Hoje, tendo se passado mais de duas décadas da elaboração desse instrumento de pesquisa, entende-se que há necessidade de aprimoração em alguns aspectos, sejam eles de ordem da formulação e adequação ao contexto, haja vista que a sociedade mudou bastante durante o transcorrer desses anos e, por vezes, algumas formulações podem ferir crenças e

religiões, por exemplo, a questão QSL-149 “O que certas pessoas fazem para prejudicar alguém e botam, por exemplo, nas encruzilhadas?” (COMITÊ NACIONAL..., 2001, p. 33).

Outras questões, tais como QSL 156: “Como se chama as coisinhas redondas de vidro com que os meninos gostam de brincar?” (COMITÊ NACIONAL..., 2001, p. 34), que pode sugerir um direcionamento para o sexismo, diferindo brincadeiras exclusivas para meninos e meninas. Nesses casos, modalizadores linguísticos podem atenuar algumas construções.

Ainda, nessa direção, cita-se a QSL-164 “Como se chama uma brincadeira em que as crianças ficam em círculo, enquanto uma outra vai passando com uma pedrinha, uma varinha, um lenço que deixa cair atrás de uma delas e esta pega a pedrinha, a varinha, o lenço e sai correndo para alcançar aquela que deixou cair?” (COMITÊ NACIONAL..., 2001, p. 34), são inclusas do rol daquelas que apresentam índices baixos de obtenção de respostas válidas – suscitando outro tipo de investigação também importante para a pesquisa dialetal, os aspectos que envolvem o não dado, conforme se discutiu na seção 2.4 – uma vez que, somente após a aplicação do questionário, pode-se repensá-lo e tentar chegar mais próximo de um melhor instrumento para as futuras coletas e pesquisas *in loco*.

No que tange ao questionário utilizado para obtenção dos fenômenos linguísticos da língua falada, às técnicas e aos métodos da exegese destes, os ideais são sempre àqueles que ainda serão elaborados, aplicados e, assim, testados, ou seja, os atuais e os vigentes são passíveis de aperfeiçoamento e lapidações, em virtude do fazer dialetológico ser contextual e dinâmico.

Para além desses dois aspectos mencionados anteriormente, é de fundamental importância, na contemporaneidade, atentar-se para o tempo de duração das entrevistas. Não se estima, de antemão, um tempo de duração médio ideal, haja vista que a entrevista nos moldes da pesquisa dialetal/sociolinguística envolve nuances que, por vezes, fogem do querer controlador do inquiridor. No entanto, quanto mais extenso e numeroso for o instrumento de coleta, a probabilidade de que a duração do inquérito seja maior, isso implica na obtenção dos fenômenos, por exemplo, com os informantes jovens, sobretudo, os homens jovens, que se cansam mais facilmente da dinâmica no ato da entrevista / do inquérito.

Após a apresentação do que se denominou o tripé básico da pesquisa geolinguística – rede de pontos, informantes e questionário – conforme postulações de Cardoso (2010). Ressalta-se, também, a relevância do (s) inquiridor (es) – principal e auxiliar – para o bom andamento da pesquisa dialetal. Afirma-se que, portanto, a atividade desenvolvida pelos inquiridores é *conditio sine qua non* para que todo o fazer dialetal *in loco*. Tal afirmação leva em conta as reflexões de Santos (2016), quando o referido autor sugere a inclusão de

mais um pé de apoio para o sustento das pesquisas em Dialectologia, para além do tripé elementar. Assim, há algumas recomendações para exercer o papel de inquiridor, tais como:

a) ter bom ouvido e gosto pela linguagem popular; b) conhecer bem a região e a língua, sobretudo se se tratar de um estrangeiro; c) ter tato e simpatia pessoal, para merecer a confiança e a boa vontade dos informantes que, em caso contrário, podem não responder ou, o que é bem possível, e pior, responder propositalmente errado; d) conhecer bem o questionário que se vai aplicar (SILVA NETO, 1957, p. 31).

Ademais, conforme Ferreira et al., (1996):

De qualquer modo, para a realização de um bom inquérito, os conhecimentos linguísticos e etnográficos do inquiridor sobre a localidade são sempre importantes, pois podem auxiliar na escolha do informante, despertar a memória de quem responde, suscitar informações suplementares ou mesmo aferir a qualidade das respostas. O questionário, indispensável, orienta o inquérito, mas o inquiridor pode transformar o interrogatório numa conversa agradável [...] Hoje em dia, parece preferível manter em campo vários inquiridores, dado que o trabalho de equipa, pela visão plural que suscita, contribui também para a qualidade dos dados finais... (FERREIRA et al., 1996, p. 486).

Por fim, ainda, no que tange às funções do inquiridor:

De um lado, a ele cabe formular e eventualmente reformular as perguntas de modo claro e objetivo para ter sucesso na obtenção de respostas válidas, insistir para obter outras variantes ou, se não tiver sucesso na obtenção de resposta(s) espontânea(s) e por meio da insistência, perguntar sobre determinadas variantes que supostamente podem ser conhecidas/usadas pelo inquerido (FIGUEIREDO JUNIOR et al., 2021, p. 3).

Nesse sentido, quando se observa a coleta para o Projeto ALiB, destaca-se a diversidade de inquiridores que, de forma incansável, pode proporcionar a construção de um banco de dados com amostras orais do português falado no Brasil em tempos hodiernos, constituindo um acervo linguístico que servirá para pesquisas de diversas naturezas e propósitos, além da elaboração dos volumes do atlas nacional.

O transcritor, também, tem um papel elementar nas pesquisas sobre a língua falada, ele “[...] é quem converte os dados brutos gravados recebidos em notação técnica, a qual pode ser grafemática, esquemática ou fonético-fonológica, com uso de sinais e símbolos convencionados, ou ainda via representação imagética de aspectos físicos ou acústicos da fala humana...” (FIGUEIREDO JUNIOR et al., 2021, p. 4).

Transcrever dados linguísticos é uma atividade que tenta ser fidedigna ao que foi coletado *in loco*, uma vez que transpõe do oral para o escrito, ou seja, é uma atividade que deve ser desenvolvida com seriedade e precisão. Caso contrário, uma transcrição incorreta e não fidedigna pode levar o pesquisador a afirmações vacilantes e não condizentes com a fala da localidade pesquisada. Nesse sentido, há algumas escolhas que devem ser feitas para que

a transcrição seja feita de forma que represente os intentos do pesquisador e se adeque ao nível de linguístico enfocado, a saber:

[...] Os níveis morfológico, sintático, textual e pragmático são transcritos grafemática ou esquematicamente. O nível dos sons individuais, quando não abordado física ou acusticamente, é objeto de transcrição fonético-fonológica. Já o nível físico ou acústico dos sons e dos padrões entoacionais é representado imagetivamente, segundo os diferentes modelos representacionais existentes... (FIGUEIREDO JUNIOR et al., 2021, p. 4).

A transcrição pode ser feita manualmente e/ou utilizando software específico, seja qual for a maneira eleita, a atividade requer pesquisadores treinados e conhecedores dos fenômenos pesquisados, além dos conhecimentos sobre os símbolos fonéticos, nos casos das pesquisas nesse nível de análise da língua. No caso dos dados do Projeto ALiB, geralmente, tal tarefa é atribuída aos bolsistas de iniciação científica, a partir de execuções manuais, com revisões atentas feitas pelos professores pesquisadores.

De certo modo, o Projeto ALiB possui diversos produtos, desde o seu surgimento até os dias atuais, tendo se passado 27 anos da fundação. Esses produtos espelham a grandiosidade de uma empreitada frutífera. Então, vale mencionar, dentre esses, a série *Documentos*, que pode retratar os caminhos trilhados pelos pesquisadores do Projeto, a fim da consecução do objetivo principal e antigo desejo dialetal brasileiro.

Além das obras da referida série, outra que, também, conta um pouco sobre os frutos do trabalho desenvolvido pela equipe do Projeto ALiB, diz respeito à coleção *Atlas Linguístico do Brasil: descrevendo a língua e formando jovens pesquisadores*, volumes I, II, III, IV, V e VI. Nesse sentido, destaca-se mais uma das potencialidades de um projeto de amplitude nacional, o interesse na formação de novos pesquisadores. O Projeto ALiB tem, nesse aspecto, se tornado uma escola. Após 27 anos de história, notam-se dezenas de pesquisadores que foram formados por uma equipe comprometida com a ciência, com o fazer dialetal e empenhada no compromisso em multiplicar os saberes, ao compartilhá-los com os jovens pesquisadores em formação.

Em 2014, por ocasião do *III Congresso de Dialectologia e Sociolinguística – III CIDS* – em homenagem às professoras Suzana Alice Marcelino Cardoso e Jacyra Andrade Mota, na cidade de Londrina-PR, foram publicados os dois primeiros volumes do *Atlas Linguístico do Brasil*. O volume 1 contém a Introdução e o volume 2 contém as Cartas Linguísticas com alguns dados das capitais.

A publicação dos primeiros volumes do *ALiB* vem fomentar ainda mais a Geolinguística no país, para além disso, concretiza um desejo antigo de um atlas nacional e

preenche uma lacuna, no que diz respeito a ter dados uniformes que possibilitem a comparação da fala de brasileiros, do Oiapoque ao Chuí, ou seja, permitindo a apresentação de dados sobre o português brasileiro pelo prisma diatópico, mas também considerando o prisma social, ao observar a variação nos planos diageracional, diastrático, diassexual e/ou diafásico.

Publicado recentemente, o volume 3⁴³ contém análises sobre as cartas do volume 2. Além disso, estão sendo projetados os volumes subsequentes, 4 e 5, que conterão estudos sobre as capitais, e os volumes 6 e 7, os pioneiros a trazerem dados linguísticos coletados nas cidades interioranas (OLIVEIRA et al., 2018). Ademais, ainda conforme as referidas autoras,

Assim, encontra-se na agenda de trabalho do Projeto ALiB a realização do Alibinho, atlas linguístico do Brasil para crianças, idealizado por Suzana Cardoso, no intuito de contribuir para: o ensino por meio de dados que apresentam a realidade espacial do domínio do português, evidenciando as distinções e semelhanças que se registram no território brasileiro; a discussão sobre a diversidade linguística, como fenômeno de todas as línguas; a diminuição de preconceitos relativos aos juízos de valores que denunciam noções equivocadas de “correto”, “incorreto”, “horroroso” e “lindo” e discriminações sociais com base na realidade da língua; a visualização de uma amostragem sobre como convivem diferenças e convergências, reconhecendo, porém, a validade da existência de um padrão culto necessário à comunicação oficial, à ministração do ensino, à efetivação do discurso formal e às opções de grupos de falantes; o fornecimento de elementos para a construção da história da língua portuguesa no Brasil, quer pela indicação dos caminhos seguidos, quer pela natureza das opções de mudança feitas, quer pelo estabelecimento de camadas caracterizadas linguística, social e geograficamente, quer, ainda, pela referência a resultados de contato com outras línguas ou à adoção de empréstimos linguísticos (OLIVEIRA et al., 2018, p. 420).

Nesse sentido, é inegável a contribuição do Projeto ALiB para o ensino de língua portuguesa, atingindo, desse modo, dois dos seus seis objetivos, a saber:

3. Examinar os dados coletados na perspectiva de sua interface com outros ramos do conhecimento — história, sociologia, antropologia, etc. — de modo a poder contribuir para fundamentar e definir posições teóricas sobre a natureza da implantação e desenvolvimento da língua portuguesa no Brasil. [...] 5. Contribuir para o entendimento da língua portuguesa no Brasil como instrumento social de comunicação diversificado, possuidor de várias normas de uso, mas dotado de uma unidade sistêmica (COMITÊ NACIONAL..., 2001, p. 16).

Assim, a riqueza do material linguístico, no tocante a língua falada, permite a elaboração dos materiais didáticos, bem como a formação de professores, haja vista que, com os dados coletados e publicados, têm-se, dessa maneira, um acervo em potencial que pode ser explorado, a fim de que o ensino contemple a variação e a diversidade linguística

⁴³ O volume 3 do atlas nacional foi publicado no XV WorkALiB, realizado em novembro de 2023, em Campo Grande, na Universidade Federal do Mato Grosso do Sul.

existentes no país, sobretudo, por meio das cartas linguísticas, utilizada enquanto gênero textual.

Além disso, por possuir amostras de falas de informantes das capitais e interiores, com diferentes escolaridades, os dicionários dialetais produzidos, através do Projeto Dicionário Dialetal Brasileiro – DDB⁴⁴ podem, nesse sentido, também, ser utilizados em sala de aula e outros espaços, como consultórios médicos, por exemplo. Assim sendo, outro objetivo do Projeto ALiB se cumpre:

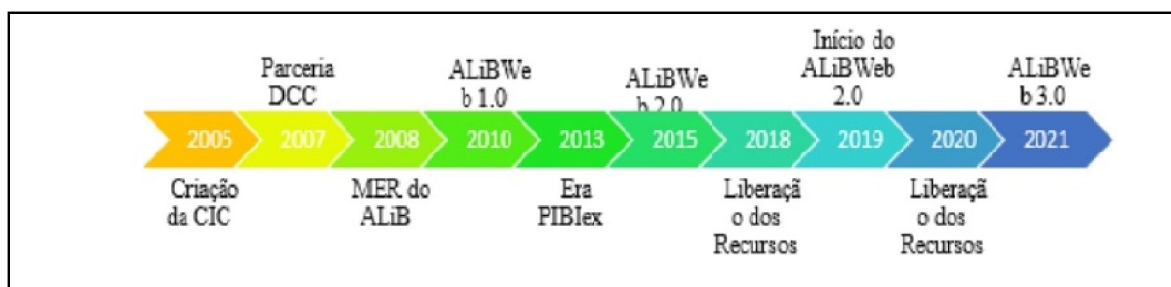
4. Oferecer, aos interessados nos estudos linguísticos, um considerável volume de dados que permita aos lexicógrafos aprimorarem os dicionários, ampliando o campo de informações; aos gramáticos atualizarem as informações com base na realidade documentada pela pesquisa empírica; aos autores de livros didáticos adequarem a sua produção à realidade cultural de cada região; aos professores aprofundar o conhecimento da realidade linguística, refletindo sobre as variantes de que se reveste a língua portuguesa no Brasil e, conseqüentemente, encontrando meios de, sem desprestigiar os seus dialetos de origem, levar os estudantes ao domínio de uma variante tida como culta (COMITÊ NACIONAL..., 2001, p. 16).

Em pleno desenvolvimento, ao comemorar 27 anos de sua existência, o Projeto ALiB segue com metas a serem alcançadas, para além das publicações dos volumes posteriores, há a intenção do lançamento, em breve, do ALiBWeb, empreendimento que vem sendo perseguido pela equipe de Cartografia e Informatização do Projeto, a fim de disponibilizar aos consulentes cartas linguísticas, com os respectivos sons e formas pronunciadas pelos informantes, haja vista que

Cientificamente, o sistema ALiBWeb trará contribuições tanto para a Dialetologia quanto para a área da Ciência da Computação, visto que já há dissertações de mestrado e teses de doutorado sendo desenvolvidas em ambas as áreas. Espera-se que, com a informatização deste atlas, novas pesquisas em ambas as áreas possam ser desenvolvidas, contribuindo para o crescimento da pesquisa na área de Dialetologia e, também, na área de Processamento de Linguagem Natural, especificamente na Extração da Informação (CLARO; PAIM; JESUS, 2021, p. 286).

Ainda sobre o empreendimento, transposição dos dados do atlas nacional para um atlas digital, no que tange à língua portuguesa, notam-se os avanços pelos quais a informatização do Projeto ALiB vêm alcançando, conforme Figura 23:

⁴⁴ Mais informações podem ser obtidas em: www.gruponemesis.ufba.br/projetos/dicionario-dialetal-brasileiro.

Figura 23 – Etapas do processo de Informatização do Projeto ALiB

Fonte: Claro; Paim; Jesus (2021, p. 286).

Com isso, espera-se que “O resultado imediatamente esperado do ALiBWeb é, evidentemente, o Atlas Falante, no qual será possível a audição da voz do informante no momento da entrevista, além das cartas linguísticas dinâmicas, geradas de acordo com os requisitos do usuário do sistema...” (CLARO; PAIM; JESUS, 2021, p. 286). Dessa forma, facilitando o acesso de pesquisadores e consulentes. Assim,

O módulo de geração automática das cartas linguísticas e dos mapas falantes visa oferecer aos usuários do sistema ALiBWeb a representação das informações através de mapas, permitindo a identificação geográfica dos aspectos linguísticos pesquisados. As cartas linguísticas podem ser obtidas através das consultas aos dados fonéticos, semânticos e sintáticos, fornecendo assim várias possibilidades de produzi-las. Além destes aspectos, os áudios referentes às palavras cortadas dentro dos questionários poderão ser acessados dentro destas cartas linguísticas, sendo assim denominado de Mapa Falante. A busca fonética também poderá ampliar a utilização dos fones dentro das cartas linguísticas (CLARO; TELES; RIBEIRO, 2017, p. 305).

Por fim, embora se possa expor aqui as inúmeras contribuições que o banco de dados do Projeto ALiB oferece, ressalta-se a possibilidade de um novo traçado dialetal para o Brasil, com isoglossas fonéticas, lexicais e/ou morfossintáticas, que retratem, de fato, os atuais limites dialetais do país, desta feita, com base em dados empíricos coletados de forma sistemática, semelhantes aos do Projeto, conforme sinalizado pelo próprio Nascentes “[...] Aguardemos o Atlas Lingüístico do Brasil (até quando?), para um trabalho definitivo” (NASCENTES, 1955, p.99). E ressaltado por Cardoso (2016, p. 47), “O caminho está aí. Necessário se faz continuar a percorrê-lo para atingir a meta almejada no que diz respeito à divisão dialetal do Brasil.”

5.2.1 Esboço de uma periodização do Atlas Linguístico do Brasil – ALiB

A história dos estudos dialetais no Brasil é consagrada por um avanço progressivo, datado e periodizado, de forma precisa e fundamentada, por dialetólogos e dialetólogas. Considerando as cinco divisões propostas, conforme subseção 2.1.3, ressalta-se a validade delas, devido à consistência teórica e à pertinência dos marcos alusivos.

Em acréscimo, no que tange à língua portuguesa falada em solo brasileiro, pensa-se que a pesquisa dialetal no Brasil trilhou caminhos singulares, ganhando forma e face abrasileiradas, espelhados por todo o fazer dialetológico – obras monográficas; produtos geolinguísticos; técnicas e métodos empregados na pesquisa *in loco*, na elaboração do instrumento de coleta e na exegese dos dados. Esses demarcam temporalidades, isto é, circunscrevem e delimitam um antes, um durante e um depois do atlas nacional, haja vista que

[...] a Geolinguística no Brasil segue o seu próprio caminho desde o início da segunda década do século XXI já com uma metodologia sedimentada e com base teórica consistente [...] Esse “desmembramento” da Geolinguística brasileira do modelo europeu originou-se da percepção dos geolinguístas de que este país possui especificidades geográficas, históricas e sociais (SILVA, ROMANO, 2022, p. 1718).

No Quadro 26, intenta-se traçar uma cronologia de fatos relevantes para o surgimento, à construção e à solidificação do Projeto ALiB, alguns desses são também representativos para a história da Dialetologia brasileira, a saber:

Quadro 26 – Síntese da história do Projeto Atlas Linguístico do Brasil – Projeto ALiB

Ano	Marco
1952	Decreto presidencial nº 30.643, que institui o Centro de Pesquisas da Casa de Rui Barbosa e dispõe sobre seu funcionamento e as funções da Comissão de Filologia.
1952	Portaria nº 536, que detalha o Decreto e assenta as funções da Comissão de Filologia e explicita a necessidade de elaboração do atlas nacional.
1957	<i>Guia para os estudos dialetológicos</i> , de Serafim da Silva Neto.
1958/1961	<i>Bases para a elaboração do atlas linguístico do Brasil</i> , de Antenor Nascentes.
1963	<i>Atlas Prévio dos Falares Baianos</i> , de Nelson Rossi ⁴⁵ .
1996	Instituição do Projeto Atlas Linguístico do Brasil (Projeto ALiB).
2001	Primeiro inquérito realizado pela equipe do Projeto ALiB, em Quirinópolis-GO.
2013	Último inquérito feito pela equipe do Projeto ALiB, em Limoeiro - PE.

⁴⁵ O APFB representa aqui todos os atlas estaduais/regional e de pequeno domínio que solidificaram o caminho para o ALiB.

2014	Publicação dos dois primeiros volumes do <i>Atlas Linguístico do Brasil</i> – ALiB.
2021	Seminário 25 anos do Projeto ALiB.
2023	Publicação do terceiro volume do <i>Atlas Linguístico do Brasil</i> – ALiB.

Fonte: Elaborado pelo autor (2023).

Vale ressaltar, de antemão, que a história do Projeto ALiB é atravessada por histórias diversas, de seus membros, dos informantes, e, também, da história do país, uma vez que os caminhos trilhados por um grande empreendimento, tal qual o Projeto ALiB, são construídos com base nessas histórias e disponibilidades dos recursos humanos e financeiros, sem os quais a pesquisa não tem êxito. No quadro 26, há uma tentativa de sumarizar os diversos fatos importantes que materializam esses anos. No entanto, essa seleção é passível de ajustes – inclusões e/ou exclusões – a fim de um delineamento mais profícuo.

A elaboração do atlas linguístico geral do Brasil guiou as aspirações dialetológicas por todo o país, pode-se estabelecer que, desde o Decreto de 1952 aos dias atuais, quando são projetados volumes novos do Atlas Linguístico do Brasil – ALiB –, há períodos que evidenciam peculiaridades e características, decorrentes de aspectos elementares para o alcance desse antigo desiderato, e que permitem uma proposta de segmentação. Postas essas considerações, consoante o Quadro 27, sugere-se uma divisão quanto aos períodos.

Quadro 27 – Proposta de periodização do Atlas Linguístico do Brasil – ALiB

FASES	MARCOS	
	Início	Término
1 ^a (<i>mens dialectus</i> ⁴⁶)	1952 Decreto nº 30.643/1953 e a Portaria nº 536/1952.	1996 Carta de Salvador Implementação do Projeto Atlas Linguístico do Brasil.
2 ^a (<i>corpus dialectus</i>)	1996 Carta de Salvador Implementação do Projeto Atlas Linguístico do Brasil.	2014 Publicação dos volumes 1 e 2 do Atlas Linguístico do Brasil.
3 ^a (<i>anima dialectus</i>)	2014 Publicação dos volumes 1 e 2 do Atlas Linguístico do Brasil.	2023 Publicação do volume 3 do Atlas Linguístico do Brasil.
4 ^a (<i>aequilibrium dialectus</i>)	2023 até os dias atuais Publicação do volume 3 do Atlas Linguístico do Brasil.	

Fonte: Elaborado pelo autor (2023).

⁴⁶ Termo alusivo ao preconizado por Serafim da Silva Neto, *mentalidade dialetológica*, em 1957.

A partir do que demonstra a historiografia da literatura da Dialectologia brasileira, esses períodos podem ser segmentados internamente, quando possível, de modo a evidenciar, em perspectiva cronológica, o fazer dialetal em cada fase, a saber:

a) 1ª fase – momento em que se cria a *mens dialectus* (mentalidade dialetológica) – inicia-se em 1952 – a partir do decreto nº 30.643/1952 que determina a elaboração do atlas nacional, reforçado pela Portaria nº 536/1952 – até 1996 – por ocasião do *Seminários Caminhos e Perspectivas para a geolingüística no Brasil*, com a implementação do Projeto Atlas Linguístico do Brasil.

Quanto às obras, o primeiro período está segmentado em i. *O linguajar carioca* (NASCENTES, 1953); *Guia para os estudos dialetológicos* (SILVA NETO, 1957); e *Bases para elaboração do atlas linguístico do Brasil* (NASCENTES 1958/1961); ii. *O Atlas Prévio dos Falares Baianos* (ROSSI, 1963); o *Esboço de um Atlas Linguístico de Minas Gerais* (RIBEIRO et al., 1977); o *Atlas Linguístico da Paraíba* (ARAGÃO; BEZERRA DE MENEZES, 1984); o *Atlas Linguístico de Sergipe* (FERREIRA et al., 1987); e o *Atlas Linguístico do Paraná* (AGUILERA, 1994); e, por fim, iii. *A Geografia linguística no Brasil* (BRANDÃO, 1991); *A Dialectologia no Brasil* (FERREIRA; CARDOSO, 1994) e *Diversidade do português do Brasil: estudos de dialectologia rural e outros* (FERREIRA et al., 1994).

Além disso, a esse primeiro momento pertence às 3ª e 4ª fases, no tocante à periodização dos estudos dialetais brasileiros. Por meio do exposto, “A tão almejada mentalidade dialetológica pôde ser criada e está sendo alimentada nas diversas instituições de ensino superior, e tem rendido safras de bons trabalhos científicos na área...” (SILVA; AGUILERA, 2009, p. 337).

b) 2ª fase – o período que se cria o *corpus dialectus* (corpo dialetológico) – inicia-se em 1996 – por ocasião do *Seminários Caminhos e Perspectivas para a geolingüística no Brasil*, com a implementação do Projeto Atlas Linguístico do Brasil – até 2014 – com a publicação dos volumes 1 - Introdução e 2 - Cartas Linguísticas do Atlas Linguístico do Brasil (CARDOSO et al., 2014ab).

A segunda fase está segmentada em i. reuniões do Comitê nacional iniciadas em março de 1997⁴⁷; *A geolingüística no Brasil: caminhos e perspectivas* (AGUILERA, 1998); Inquéritos experimentais realizados com o fito de testar o instrumento de coleta, utilizando

duas versões – 1998 e 2000 – que sucederam ao questionário definitivo; *Questionários ALiB* (COMITÊ NACIONAL...2001); *Documentos 1* (AGUILERA et al., 2004); *Documentos 2* (MOTA; CARDOSO, 2006); ii. *Geolinguística: tradição e modernidade* (CARDOSO, 2010); *A geolinguística no Brasil: trilhas seguidas, caminhos a percorrer* (AGUILERA, 2013). Esse período pertence às fases 4ª e 5ª.

c) 3ª fase – período em que se cria a *anima dialectus* (alma dialetológica) – inicia-se em 2014 – a partir publicação dos dois primeiros volumes do Atlas Linguístico do Brasil (CARDOSO et al., 2014ab) – até a publicação do terceiro volume do atlas (MOTA et al., 2023), em 2023 –, classificado como o período de ressignificação, quando já se pensa, com o fazer atual, o momento porvir.

Para a 3ª fase, destacam-se alguns pontos levantados que, de certa maneira, carecem de mais discussões e refinamentos, a saber: i. a dissertação de Santos (2016) com a sugestão de reformulação do tripé básico da pesquisa dialetal; a proposição da cartografia do não dado por Santos (2016/2020), conforme seção 2.4; a obra *A geolinguística no Brasil: caminhos percorridos, resultados alcançados* (AGUILERA; ROMANO, 2017); a tese de doutorado de Teles (2018), com o mapa dialetal proposto por Nascentes em 1953, a partir do georreferenciamento da Cartografia automatizada, e a proposição da 5ª fase para a Dialetologia brasileira; e, mais recentemente, a obra *Tendências da Geolinguística brasileira e a nova geração de Atlas Linguístico* (SILVA, ROMANO, 2022). Ademais, a terceira fase pertence às 4ª e 5ª fases, no que tange à periodização dos estudos dialetais do Brasil.

As novas proposições, tais como, a partir dos incontáveis debates contemporâneos, suscitar aproximações entre o letramento racial e os bancos de dados dialetais e sociolinguísticos, em virtude da história social e da história linguística brasileira, haja vista que “As categorias sociodemográficas estratificadas nos bancos de dados sociolinguísticos brasileiros não contemplam raça/etnia, nem classe social...” (FREITAG, 2016, p. 455). Tal aspecto, que pode ser considerado como lacunar, estende-se também aos bancos de dados dialetológicos brasileiros.

Nessa direção, encontram-se parâmetros nos trabalhos de Dias (2017); Cabral (2017); Sá (2018); e Costa (2019), que tencionaram aspectos sobre a fala quilombola e indígena, isto é, estudos de dimensão microdialetais dão início a essas perspectivas elementares. Espera-se que, com a solidificação e a expansão, estudos com dimensões macrodialetais sejam empreendidos. Assim, dando base às ações voltadas para o cumprimento do decreto 7.387/2010, que institui a criação do Inventário Nacional da Diversidade Linguística –

INDL⁴⁸ (BRASIL, 2010). Além disso, a incorporação dos aspectos atinentes ao gênero e à sexualidade nos *corpora* dialetais, bem como a reformulação dos questionários da pesquisa *in loco*, conforme menção na seção 5.2, são tarefas necessárias para o avanço e a atualização/modernização desse ramo dos estudos linguísticos.

d) 4ª fase, – período em que se cria o *aequilibrium dialectu* (equilíbrio dialetológico) – inicia-se em 2023 – a partir publicação do terceiro volume do Atlas Linguístico do Brasil (MOTA et al., 2023) – até os dias atuais. Nessa fase, desenvolve-se a Dialetologia brasileira do futuro, aludida ao entre os períodos *mens – corpus – anima*. Geralmente, em relação aos aspectos futurísticos, quando se almejam o crescimento e a prosperidade, projetam-se passos no momento atual, a fim de que o empreendimento seja bem-sucedido.

A atividade dialetal brasileira, no tempo presente, de forma a continuar sendo um ramo de estudos linguísticos contextual, pode arquitetar um futuro promissor, ao buscar um equilíbrio, com base no pretérito fundamental e no presente frutífero, haja vista que os avanços científicos são de extrema importância. No entanto, para isso, faz-se necessário tomar como base os antecessores, de modo a contemplar os feitos, quando possível, ressignificá-los, e, na linha tênue, entre o ideal e o possível, vislumbrar os passos subsequentes. Por isso, para os bancos de dados das pesquisas sobre a língua falada,

[...] **no Brasil, precisamos resgatar este lado mais social [...], contemplando não só coletas que não sigam estratificações fechadas e homogeneizantes, mas que percebam a comunidade como um todo.** Isso requer um engajamento maior com a comunidade e com o informante. Meyerhoff (2016) destaca que o próprio Labov, ao revisitar a coleta inicial do “The Social Stratification of English in New York City”, recomenda mais do que um único encontro entre o entrevistador e o entrevistado, relevando a importância do trabalho articulado entre a pesquisa de ampla amostragem e a observação participante. Os encontros mais frequentes fazem com que o vernáculo emergja... (FREITAG, 2016, p.458) [grifos nossos].

Nesse entendimento, acredita-se que presente, passado e futuro estão estritamente interligados, conforme ilustra a Figura 24.

⁴⁸ Dentre os objetivos centrais do INDL, destaca-se o que promover, além de valorizar a diversidade linguística brasileira, com ênfase em línguas brasileiras, consideradas como patrimônio cultural nacional, a saber: a) indígenas; b) imigração; c) comunidades afro-brasileiras; d) sinais; e) crioulas e; f) língua portuguesa e suas variações dialetais.

Figura 24 – Proposta de periodização do Atlas Linguístico do Brasil – ALiB

1ª fase
(1952-1996)
2ª fase
(1996-2014)
3ª fase
(2014-2023)
4ª fase
(2023...)

Fonte: Elaborado pelo autor (2023).

Vale destacar que essa proposta não pode ser considerada definitiva, embora lance, ainda mais, luz aos esforços empreendidos por gerações de dialetólogos. Entende-se que tal proposição está sujeita à apreciação da comunidade dialetológica, ou seja, carece de contribuições e, conseqüentemente, caso seja necessário, de refacção. Ressalta-se que, também, os limites fixados entre os períodos são fluídos, haja vista que eles fazem parte de um *continuum* dialetal perene.

É consabido que a Dialetologia à brasileira se desenvolveu em torno de um grande objetivo, o Atlas Linguístico do Brasil, no tocante à língua portuguesa falada, dessa maneira, nota-se que existem aspectos que mudaram, por interferência desse propósito que inspirou gerações de pesquisadores nessa área dos estudos linguísticos. Por tais razões, sugere-se um alvorecer de um momento de autoavaliação, que considera e valoriza os aspectos dos momentos pretéritos, mas pensa-os criticamente, além de projetar, por meio dos feitos atuais, o período vindouro. Logo, fica a questão que não será respondida agora, para além do ALiB, é coerente pensar no surgimento de uma Dialetologia brasileira crítica?

Nessa direção, há caminhos que apontam para ações programáticas com vistas ao fortalecimento e à expansão dos projetos que se debruçam sobre o português falado no Brasil, sobretudo àqueles que possuem banco de dados, quer sejam fundamentados pela Dialetologia, quer sejam pela Sociolinguística. Dessa maneira,

[...] muito já se sabe sobre a realidade dialetal e sociolinguística do país. Agora é preciso a conjugação de teorias e métodos diversos bem como a exploração dos dados já disponíveis e a coleta de outros novos para uma agenda de pesquisa deste século XXI. Deve-se agregar aos trabalhos a inter- e a transdisciplinaridade e pensar em formas de contribuição ao ensino de língua portuguesa. É urgente, portanto, rever os procedimentos metodológicos (pensando, inclusive, em coletas alternativas), incorporar as novas tecnologias, realizar mais estudos comparativos e ampliar as formas de divulgação dos resultados para que ultrapassem os círculos acadêmicos e cheguem à sociedade como um todo e, sobretudo, aos envolvidos com o ensino-aprendizagem da língua portuguesa (OLIVEIRA; GONÇALVES; MOTA, 2021, p. 299).

Assim, essas ações podem ser vistas como frentes de pesquisas, ao mesmo tempo, independentes e interligadas, a saber: i. autoavaliação - preservação - adequação; ii. informatização e digitalização do *corpus*; iii. aproximação do ensino de língua portuguesa; iv. divulgação das pesquisas; e, por fim, v. formação de novos pesquisadores.

As ações descritas em i. autoavaliação - preservação - adequação têm como objetivo fazer um balanço de modo a revisar aspectos que necessitam de mudança, tais como métodos e instrumento de coleta, por exemplo, visando à adequação ao contemporâneo. Além disso, revisão e revisitação dos arquivos – sonoros e fotográficos –, com o objetivo de construir de um memorial, a fim de que as histórias sejam preservadas. Nesse quesito, ações de entrevistas com percussores da área também são fundamentais.

No item ii. informatização e digitalização do *corpus*, como descrito nas subseções 5.1 e 5.2, os Projetos NURC e ALiB já são exemplos para que os acervos das amostras de fala sejam melhor condicionados. O objetivo principal é salvaguardar, em meios mais modernos, os registros/testemunhos orais da fala brasileira. As interfaces com outras áreas do saber são bem-vindas, tais como com a História, a Ciência da Informação e a Geografia. As cartas com QR-code, por exemplo, ilustram possibilidades frutíferas para essas interfaces. Sugere-se, então, que tais ações sejam fomentadas pelo financiamento público, haja vista que as pesquisas sobre a língua também são importantes para memória e história de um povo.

Para a terceira ação, iii. aproximação do ensino de língua portuguesa, há muito o que ser feito, haja vista que, ainda, tentativas tímidas são feitas, embora já logrem bons frutos. As políticas sobre o desenvolvimento do respeito linguístico (SCHERRE, 2019) podem ser fomentadas na escola, a partir dos resultados de pesquisas empíricas, como exemplo, citam-se as cartas linguísticas, enquanto instrumentos eficazes no que tange à demonstração da configuração de um determinado fenômeno, extinguindo pensamentos enraizados no preconceito e sem fundamento. Quiça, materiais didáticos que contemplem a diversidade linguística, desta feita, sem estereótipos, a exemplo do trabalho de Silva e Santos (2022), ou seja, fomentando uma vertente frutífera, a Dialetoologia educacional.

O fazer ciência é para servir a uma coletividade, assim, o item iv. divulgação das pesquisas necessita de ações que se voltem para a promoção e valorização de toda produção das pesquisas sobre a língua falada em termos de conhecimento, resgate e preservação. Nesse sentido, parcerias com as mídias, quer sejam televisivas, radiofônicas e/ou digitais são de grande importância, com vistas a, por meio da visibilidade, atrair recursos diversos, sobretudo financeiros. Também, com maior alcance das pesquisas sobre a língua falada, a

intenção é propagar que “[...] Unidade e diversidade não se defende, constata-se” (FERREIRA; CARDOSO, 1994, p. 21).

Por seu turno, as ações voltadas para o item v. formação de novos pesquisadores objetivam preparar, da graduação à pós-graduação, novos pesquisadores, a fim de que todo trabalho, fomentado e sustentado por gerações de pesquisadores sobre a língua falada no Brasil, seja mantido, alimentado e desenvolvido/aperfeiçoado. Tais ações necessitam estar conjugadas com as ações anteriormente descritas, haja vista que cursos de extensão, os subprojetos de pesquisa de iniciação científica, o trabalho em sala de aula, bem como seminários de pesquisa voltados aos convidados externos, visando o plantar novas sementes e novos fôlegos para o fortalecimento da pesquisa e das agendas sobre o português falado no Brasil, principalmente coleta de dados, elaboração de atlas – de macro extensão e/ou micro extensão –, dentre outros feitos.

Dessa maneira, ao congrega essas ações, pensa-se em um letramento sociodialetoológico cuja definição aponta para um processo amplo e contínuo que possibilita o desenvolvimento de habilidades, comportamentos e práticas fundamentais para o labor da pesquisa sobre a língua falada – sociolinguística e/ou dialetológica – e permite aos sujeitos a instrumentalização para compreender a esfera discursiva das pesquisas nessas duas áreas.

5.3 SÍNTESE DA SEÇÃO

Esta seção, 5, tem por finalidade trazer à tona aspectos sobre dois grandes marcos para os estudos linguísticos brasileiros, haja vista que, por meio do banco de dados desses projetos, é possível expor, de forma panorâmica, fotografias da língua no Brasil em diferentes épocas, sobretudo, em diferentes graus de formalidade em relação à utilização da língua falada.

Nas subseções 5.1 e 5.2, busca-se apresentar um breve histórico dos Projetos NURC e ALiB, respectivamente. Além disso, de forma inédita, embora consideradas não definitivas, em virtude de ser um esboço, são apresentadas periodizações para esses projetos – NURC com três fases e ALiB com quatro fases –, considerando fatos e marcos alusivos a grande relevância para os estudos sociolinguísticos e dialetológicos.

É consabido que, na história dos estudos sobre a língua falada no Brasil, tanto na Dialetologia quanto na Sociolinguística Variacionista, os projetos de pesquisa, quer sejam de grande dimensão, quer sejam de dimensão menor, possuem um local de destaque, frente ao entendimento que se tem hoje, pois espelham e refletem peculiaridades sobre os aspectos espaciais, sociais e históricos, desde os primórdios, que moldam a forma pela qual

determinadas comunidades utilizam a língua. Portanto, para que os Projetos NURC e ALiB continuem vigorosos, as ações programáticas descritas na subseção 5.2.1 são de extrema relevância, uma vez que outros projetos, certamente, se espelham nas trajetórias trilhadas por eles.

6 ROTAS METODOLÓGICAS

[...] a Dialectologia, mais do que qualquer outra especificidade linguística, tende a fixar-se num reduzido elenco de princípios gerais, teóricos, abstratos, em benefício de uma extrema flexibilidade de métodos, de uma enorme variedade de processos para solução dos problemas concretos que se lhe ofereçam.

(ROSSI, 1967, p. 90).

Nesta seção, os materiais e os métodos utilizados serão explicitados, a fim de chegar aos procedimentos necessários para a posterior análise dos dados. Após a exposição dos fundamentos históricos e metodológicos dos dois grandes projetos, NURC e ALiB, faz-se necessário, agora, explicitar quais foram os critérios metodológicos adotados nessa tese, perpassando pelos critérios peculiares aos *corpora*, os parâmetros estabelecidos para o tratamento estatístico, bem como para a cartografia e a exploração sobre a metalinguagem dos dados.

Os dados do Projeto NURC foram consultados a partir das transcrições grafemáticas publicadas⁴⁹. Nesse sentido, foram 44 pertencentes às Elocuções Formais (EF); 228 pertencentes ao Diálogo entre informante e documentador (DID); e 85 pertencentes ao Diálogo entre dois informantes (D2). Salienta-se que, mais uma vez, os informantes se enquadram em perfis previamente delimitados com os seguintes critérios: todos possuem formação universitária, são homens e mulheres, e estão subdivididos em três faixas etárias: faixa I - de 25 a 35 anos; faixa II - de 36 a 55 anos; e faixa III - de mais de 56 anos, oriundos das cinco capitais eleitas, POA, SP, RJ, SSA e REC. Esses inquéritos foram estratificados e a divisão da amostra pode ser visualizada no Quadro 28.

⁴⁹ Os inquéritos do NURC/RJ também foram consultados pelo site <https://nurcrj.letas.ufrj.br/> onde há disponível um acervo maior do que o que consta nas publicações, sobretudo de inquéritos DIDs.

Quadro 28 – Inquéritos consultados do Projeto NURC

CIDADE	QUANTIDADE	MATERIAL
São Paulo SP	06	EF
	06	D2
	09	DID
Rio de Janeiro RJ	06	EF
	08	D2
	160 ⁵⁰	DID/70
	15	DID/90 ⁵¹
Salvador SSA ⁵²	06	EF/70
	06	D2
	12	DID
	06	EF/90
	12	DID/90
Recife REC	12	EF
	08	D2
	12	DID
Porto Alegre POA	08	EF
	07	D2
	08	DID
Total: 307 inquéritos		

Fonte: Elaborado pelo autor (2023).

Como se nota no Quadro 28, são 21 inquéritos de São Paulo – 27 informantes; 189 do Rio de Janeiro – 197 informantes; 42 de Salvador – 48 informantes; 32 de Recife – 40 informantes; e, por fim, 23 de Porto Alegre – 30 informantes, desse modo, totalizando 307 inquéritos consultados e 342 informantes.

Vale ressaltar que não existia, no Projeto NURC, nenhum tipo de questionário, muito menos uma pergunta específica direcionada à apuração do(s) fenômeno(s), embora, no DID e no D2, as falas fossem em torno de um tema fornecido pelo pesquisador. Logo, todos os inquéritos foram analisados, pois as ocorrências poderiam estar presentes em qualquer parte dos diálogos e/ou elocuições, em qualquer uma das falas, em qualquer tipo de situação gravada.

Os dados do Projeto ALiB são inéditos, foram coletados e estão dispostos no banco de dados do Projeto. Objetivando a análise e o tratamento destes dados, foi solicitada a autorização, junto ao Comitê Nacional do Projeto ALiB, para compor o *corpus*. Após a

⁵⁰ Os inquéritos do NURC/RJ estão disponibilizados no site www.nurcj.letas.ufrj.br, por isso eles foram consultados em maior número.

⁵¹ Os inquéritos consultados foram da amostra complementar da década de 90.

⁵² Os inquéritos da década de 90, EF e DID, estão no prelo, organizados no volume 4, pela equipe NURC/SSA.

liberação, passou-se a buscar os áudios e ouvi-los atentamente, a fim de coletar os materiais que serviriam de análises e tratamentos estatísticos, além de consultar, quando necessário, as transcrições dos inquiridos, bem como a ficha dos informantes.

Como já foi destacado, os dois *corpora* aqui analisados possuem natureza e propósitos distintos. Nesse sentido, vale frisar que foram consultados 200 inquiridos, oriundos das 25 localidades, e a estratificação dos informantes segue a mesma utilizada pelo Projeto ALiB, conforme Quadro 29, no que tange às capitais.

Quadro 29 – Estratificação dos informantes do Projeto ALiB – capitais

Faixa etária	Sexo/ Escolaridade				subtotal faixa etária	
	Fundamental		Universitário			
	Homem	Mulher	Homem	Mulher		
Faixa I	25	25	25	25	100	
Faixa II	25	25	25	25	100	
subtotal	Sexo/	50	50	50	50	Total geral: 200 informantes.
Escolaridade						

Fonte: Elaborado pelo autor (2023).

Ainda, no que concerne aos aspectos que diferem os *corpora*, nota-se que há, no QMS do Projeto ALiB, uma questão que se destina a apurar o fenômeno em descrição e análise nesta tese, QMS 023, por meio de duas formulações, a saber:

1. Alguém pede para você / o (a) senhor (a) fazer uma tarefa. Mas outra pessoa acha que a tarefa era para ela. Então, você / o (a) senhor (a) diz: Essa tarefa, na verdade, é para ____ fazer.
2. [Vamos supor que você / o (a) senhor (a) precisa distribuir tarefas por três pessoas. Por exemplo: a limpeza é para alguém fazer, a roupa é para outra pessoa lavar e a terceira tarefa (...) é para você / o (a) senhor (a). Como você / o (a) senhor (a) diria? (COMITÊ NACIONAL... 2001, p. 40) [grifos nossos].

Sabe-se que a catalogação de fatos morfossintáticos é passível de ser realizada em todas as partes do inquirido linguístico. Por isso, optou-se por considerar a análise empreendida sobre as estruturas *para/prá + eu/mim + infinitivo* em todo o Questionário ALiB (2001), não somente catalogar no QMS 023.

No que tange à questão destinada a coletar as formas pronominais, *eu* e *mim*, para a posição de sujeito, precedidos pela preposição *para*, observa-se que, na primeira formulação, o inquiridor elabora uma situação hipotética na qual permite que o informante apenas preencha a lacuna deixada pela formulação. Tal indução também impede que o informante

pronuncie a estrutura de forma mais espontânea, coletando, muitas vezes, na fala do informante, apenas um dos pronomes. Assim,

Nesta oportunidade, convém lembrar uma das principais dificuldades na formulação das perguntas de um questionário morfossintático: ele se distingue dos outros dois questionários por exigirem, como resposta, mais que uma única palavra. Espera-se do informante uma frase completa que ele nem sempre pronunciará na primeira tentativa (KOCH, 2004, p.72).

Caso a primeira formulação não tenha sido suficiente para a obtenção do dado, o inquiridor opta pela segunda, que, pelo que pode ser visto, já concede maior liberdade para que o informante realize a estrutura completa, não somente preencha a lacuna deixada pelo inquiridor. Portanto, a apuração de dados morfossintáticos a partir da aplicação de questionário do Projeto ALiB, muitas vezes, torna-se uma tarefa complexa, pois

Parte-se, para esta reflexão, da afirmação preliminar de que toda aplicação de questionário — de qualquer sorte — apresenta-se como uma técnica monitorada, e busca-se analisar, a partir, de exemplos do Projeto ALiB, como se comporta ou deve se comportar o pesquisador, como deve lidar com esse ou aquele tipo de questionário e como interpretar os resultados, pondo-se em destaque o “efeito gatilho” que se pode observar na formulação de determinadas questões [...] o “efeito gatilho” é um fato real, mas a sua atuação pode ser acionada ou não como se pôde ver nos casos trazidos à consideração em que ora a formulação do inquiridor não interfere na resposta do informante, ora o discurso do inquiridor condiciona a própria formulação do informante (CARDOSO, 2009, p. 4507-4510).

No que tange à rede de pontos, as atenções voltam-se para os dados das 25 capitais, conforme Figura 25. Também, elas estão sistematizadas no Quadro 30, com os respectivos números dos pontos.

Figura 25 – Capitais pertencentes à Rede de Pontos do Projeto ALiB.

Fonte: Teles (2018) – Projeto ALiB – Comissão de Informatização e Cartografia.

Quadro 30 – Capitais pertencentes à Rede de Pontos do Projeto ALiB

Região Geográfica	Estado UF	Rede de Pontos ALiB N.º Ponto	Localidade Nome da Capital
Norte	Amapá	02	Macapá
	Roraima	03	Boa Vista
	Amazonas	06	Manaus
	Pará	12	Belém
	Rio Branco	20	Acre
	Rondônia	21	Porto Velho
Nordeste	Maranhão	26	São Luís
	Piauí	35	Teresina
	Ceará	41	Fortaleza
	Rio Grande do Norte	53	Natal
	Paraíba	61	João Pessoa
	Pernambuco	70	Recife
	Alagoas	77	Maceió
	Sergipe	79	Aracaju
Bahia	93	Salvador	

Centro-Oeste	Mato Grosso	108	Cuiabá
	Mato Grosso do Sul	115	Campo Grande
	Goiás	123	Goiânia
Sudeste	Minas Gerais	138	Belo Horizonte
	São Paulo	179	São Paulo
	Espírito Santo	190	Vitória
	Rio de Janeiro	202	Rio de Janeiro
Sul	Paraná	220	Curitiba
	Santa Catarina	230	Florianópolis
	Rio Grande do Sul	243	Porto Alegre

Fonte: Elaborado pelo autor (2023).

Notam-se, também, aspectos que levam ao cotejo, uma vez que são convergentes, a saber: cinco capitais fazem parte dos dois *corpora*, Porto Alegre (POA), São Paulo (SP), Rio de Janeiro (RJ), Salvador (SSA) e Recife (REC). Além disso, há informantes que possuem nível universitário. Nesse sentido, salvaguardando as diferenças metodológicas dos Projetos, os dados poderão ser vistos a partir de um panorama que demonstre semelhanças e/ou diferenças, caso existam, dos informantes com o nível universitário dessas cinco capitais nos anos 70/90/2000.

6.1 O LEVANTAMENTO DOS DADOS E OS CRITÉRIOS PARA O TRATAMENTO ESTATÍSTICO E INTERPRETATIVO

Nesta subseção, encontram-se os critérios procedimentais utilizados para o tratamento dos dados do Projeto NURC e do Projeto ALiB, a partir dos 507 inquéritos analisados.

Para manipulação dos dois *corpora*, então, algumas etapas foram feitas, a saber:

- (1) Inicialmente, uma consulta aos dados – acesso às transcrições de ambos e audição dos áudios do Projeto ALiB;
- (2) Levantamento criterioso de todas as formas coletadas, *eu*, *mim* e *zero anafórico*;
- (3) Identificação dessas formas, a partir das elocuições, atentando-se para o contexto em que elas apareceram;
- (4) Organização dos dados em planilhas do Excel, com dados contextuais, a fim de se perceber aspectos linguísticos que pudessem interferir na escolha pronominal, para além da estrutura *para/pra + eu/mim + infinitivo*;

- (5) Análise dos contextos nos quais os informantes forneceram pistas para análises posteriores, destacando que, às vezes, esses contextos evidenciam um pensamento e/ou um juízo de valor metalinguístico e social sobre o uso pronominal;
- (6) Registro de comentários ao longo da audição e das análises, objetivando, assim, possíveis notas explicativas de casos singulares;
- (7) Codificação dos dados;
- (8) Rodadas no programa *Goldvarb X* (SANKOFF; TAGLIAMONTE; SMITH, 2005);
- (9) Análises e interpretações dos percentuais e pesos relativos fornecidos pelo programa estatístico;
- (10) Elaboração das cartas linguísticas, a fim de favorecer a visualização espacial de como está configurado o fenômeno nas capitais brasileiras, nas décadas de 70, 90 e nos anos 2000.

As não respostas (ausência de dados), no caso dos *corpora* do Projeto NURC, foram encontradas nas Elocuções Formais (EF), como se detalhará a diante. No caso do *corpus* do Projeto ALiB, sejam elas por motivos diversos, foram reunidas em um grupo: NO – Não Obtida: são todas as ocorrências em que a resposta não foi dada para o QMS 023, embora o inquiridor tenha tentado, por meio de várias reformulações, e não tenha obtido êxito e/ou em todo o inquérito. Há casos, também, em que o inquiridor, ao formular a pergunta, pronunciou a resposta. Além dessas, são assinalados os casos de problemas técnicos nos áudios. De imediato, considera-se como resposta não obtida.

Somam-se a esses os casos de NS – Não sabe e NL – Não lembra, pois, na maioria dos casos, aventa-se ser a interferência do estereótipo que cerca a utilização da forma estigmatizada, *mim*. Tais dados receberão um tratamento qualitativo/interpretativo, assim como a metalinguagem dos/nos dados.

A planilha foi organizada de modo a facilitar, *a posteriori*, o tratamento dos dados referentes ao fenômeno em estudo. Sendo assim, as respostas foram elencadas por ordem de ocorrência.

Após verificações preliminares, a partir das análises dos dados, optou-se por considerar a variável dependente binária, logo, os dados com o *zero anafórico* foram descartados. Tal deliberação foi motivada por: i) não haver casos de variação estrita com

zero com sujeito em matriz de primeira pessoa; e ii) não haver casos de variação estrita com zero com sujeito em matriz de pessoa diferente.

Além disso, as construções com *pr'eu* ou construções com elementos gramaticais entre a forma pronominal e o verbo serão alvos de descrição qualitativa e foram separadas do cômputo geral, pois foram poucas ocorrências. Desse modo, somente as ocorrências dos pronomes *eu* e o *mim* foram consideradas para as análises estatísticas, obedecendo à estrutura *para/prá + eu/mim + infinitivo*.

Salienta-se que, inicialmente, para o tratamento estatístico, os dados coletados foram codificados e submetidos ao *Goldvarb X* de forma separada, uma vez que os projetos possuem diferentes naturezas metodológicas e peculiaridades, conforme já mencionadas nesta seção. A variável dependente e as variáveis independentes, no que tange à base linguística, foram semelhantes. Por seu turno, as variáveis extralinguísticas, algumas foram semelhantes e outras diferiram. No entanto, houve um cotejo com as variáveis comuns aos dois *corpora*, conforme se explicitará adiante. E, para o *corpus* do Projeto ALiB, houve uma rodada de cruzamento entre as variáveis escolaridade e localidade.

6.1.1 A variável dependente

Foi estabelecida como variável dependente, para esse estudo, a alternância binária entre os pronomes do caso reto, *eu*, e o pronome do caso oblíquo tônico, *mim*. No PB, a variante do caso reto é a forma de prestígio. Ao passo que a forma do caso oblíquo tônico é a forma estigmatizada, que, muitas vezes, de forma pejorativa, é atribuída à língua indígena falada e/ou associada ao “erro linguístico”.

6.1.2 As variáveis independentes

Nesta subseção, apresentam-se as variáveis independentes, a fim de verificar o encaixamento do fenômeno na estrutura linguística e na estrutura social.

6.1.2.1 Variáveis linguísticas

Para escolha e definição das variáveis linguísticas, optou-se por aquelas que podem ter influência na escolha pronominal dos informantes. Nesse sentido, foram eleitas as seguintes variáveis: estilo de registro; natureza semântica do verbo da oração principal; papel

semântico do pronome sujeito; função sintática da oração infinitiva; conjugação verbal; e a forma da preposição. Essas variáveis foram utilizadas na análise dos dois *corpora*, vale ressaltar.

6.1.2.1.1 Estilo de registro

O controle dessa variável decorreu da necessidade de saber se os informantes se comportam linguisticamente de forma diferente, mediante o tipo de questionário durante a entrevista e/ou situação comunicativa, quer sejam situações mais tensas, monitoradas, ou menos tensas, livres.

Os fatores escolhidos para a observação dos dados do Projeto NURC foram considerados a partir do tipo de inquérito ao qual o informante foi submetido, uma vez que eles possuem naturezas distintas, por isso entende-se que o discurso irá diferir também. Logo, tencionou-se, por meio do controle desse grupo, verificar de que modo os estilos contextuais interferem nas escolhas pronominais. A hipótese é de que quanto mais próximo de estilos monitorados, os informantes utilizam a forma canônica, *eu*. Nesse sentido, os fatores são:

a. ++ monitorado –EF – Elocução Formal;

(05) INF: Às vezes, fica difícil *pra eu ver* o que ela está escrevendo...

(Inq. 334. Recife, feminino, faixa I, universitário – EF).

b. + monitorado – DID – Diálogo entre informante e documentador;

(06a) INF: Eu posso comer em casa...Os outros dias todos dá *pra mim almoçar* em casa.

(Inq. 104. Salvador, masculino, faixa I, universitário – DID).

(06b) INF: ...eu penso que...não...não sei bem como eu digo precisaria definir bem religião *pra eu poder dizer se... É ou não é não é? o que é...uma demonstração de fé indiscutível...*

(Inq. 242. São Paulo, masculino, faixa III, universitário – DID).

c. + monitorado – D2 – Diálogo entre dois informantes.

(07a) INF: Você vai comprar uma casa que tenha um jardim na frente *pra eu brincar* com terra é água?

(Inq. 269. Rio de Janeiro, feminino, faixa II, universitário – D2).

(07b) INF: ... eu sempre vejo o aspecto teórico, vejo o outro lado da coisa...então... fica muito mais difícil *pra mim ver*, mas eu estou procurando agora dialogar contigo....

(Inq. 283. Porto Alegre, masculino, faixa I, universitário – D2).

Para os dados do Projeto ALiB, quanto ao estilo de registro, a hipótese é semelhante ao do Projeto NURC. Os fatores considerados foram:

a. + monitorado – QFF; PR, QSL e QMS;

(08a) INQ: O que você / o (a) senhor (a) faria se ganhasse na loteria?

INF: Se eu ganhasse na loteria, eu ia comprar um carro *pra mim ajudar* minha vó, né? Dar uma casa melhor pra ela também e curtir um pouco da vida.

(Inf. 01 – 093: Salvador, homem, faixa I, fundamental – QMS 044).

(08b) INQ: Como se chama a parte do dia quando começa a clarear?

INF: ... já era de manhã porque a minha menina tinha escola lá pertinho lá de casa e eu ia todo dia cedinho *pra mim apagar a luz* pra num... por causa num num num ir muito alto a conta de luz né?

(Inf. 03 – 108: Cuiabá, homem, faixa II, fundamental – QSL 022).

b. ++ monitorado – QMS 023 – foi considerada como um fator monitorado, de forma isolada das demais questões do QMS;

(09a) INQ: Alguém pede para você / o (a) senhor (a) fazer uma tarefa. Mas outra pessoa acha que a tarefa era para ela. Então, você / o (a) senhor (a) diz: Essa tarefa, na verdade, é para _____ fazer.

INF: É *pra eu fazer*.

(Inf. 01 – 093: Salvador, homem, faixa I, fundamental – QMS 023)

(09b) INQ: Alguém pede para você / o (a) senhor (a) fazer uma tarefa. Mas outra pessoa acha que a tarefa era para ela. Então, você / o (a) senhor (a) diz: Essa tarefa, na verdade, é para _____ fazer.

INF: Isto, essa tarefa é *pra mim fazer*.

(Inf. 07 – 243: Porto Alegre, homem, faixa II, universitário – QMS 023).

c. + livre – PM, TDS e LE.

(10a) INQ: Isso. Tá certo. Aí essa leitura.

INF.- *Pra mim ler* isso daqui?

(Inf. 06 – 03: Boa Vista, mulher, faixa I, universitário – LE).

(10b) INQ: – Tá bom. E antigamente, será que o pessoal falava diferente?

INF: – *Pra eu falar* assim, num sei não.

(Inf. 02 – 079: Aracaju, mulher, faixa I, fundamental – PM 06).

6.1.2.1.2 Natureza semântica do verbo da oração principal

Buscou-se, por meio do controle desse grupo, observar a relevância da natureza semântica exercida pelo verbo que se encontra na oração principal da sentença. A hipótese é de que os verbos elocutivos e os verbos existenciais são favorecedores aos contextos mais formais, desse modo, favorecendo o pronome *eu*. Assim, esses verbos formam seis fatores:

a. Ação ou mudança de estado;

(11a) INF: Eu geralmente boto só *pra mim escutar*.

(Inf. 03 – 026: São Luís, homem, faixa II, fundamental – QMS 032).

(11b) INF: Por exemplo, eu, eh, eh, roupa de trabalho eu compro roupa *pra eu trabalhar*.

(Inq. 233. Rio de Janeiro, masculino, faixa II, universitário – DID).

b. Avaliativos e modais;

(12a) INF: É muito fácil *pra mim estabelecer* amizades, embora eu seja, de um certo modo, fechada.

(Inq. 56. Rio de Janeiro, feminino, faixa II, universitário – DID).

(12b) INF: mas não é comum *pra mim falá* alvorecendo...

(Inf. 7 – 093: Salvador, homem, faixa II, universitário – QSL 022).

c. Elocutivos;

(13a) INF: Eu vou perguntando logo que é *pra eu saber* qual é a posição que eu vou tomar.

(Inq. 237. Rio de Janeiro, masculino, faixa III, universitário – DID).

(13b) INF: ... venha conversar *pra eu dar* uns bons conselhos...

(Inq. 279. Recife, feminino, faixa III, universitário – D2).

d. Estativos;

(14a) INF: Minha filha, eu num tenho possibilidade, nem ninguém da minha família tem fazenda *pra mim ir*.

(Inf. 4 – 079: Aracaju, mulher, faixa II, fundamental – QSL 022).

(14b) INF: ... faltava acho que uns três meses *pra mim completar* seis anos...

(Inf. 2 – 20: Rio Branco, mulher, faixa I, fundamental – TDS 01).

e. Existenciais;

(15a) INF: Geralmente não tem pessoas assim conhecidas e parentes longe *pra eu poder* ter oportunidade de escrever cartas. Eu uso telegrama.

(Inq. 15. Rio de Janeiro, feminino, faixa I, universitário – DID).

(15b) INF: Havia os bondes...*pra eu ter* aula.

(Inq. 210. Rio de Janeiro, masculino, faixa III, universitário – DID).

f. Transferenciais;

(16a) INF: Ele me cedeu a casa *pra mim morar*...

(Inf. 3 – 108: Cuiabá, homem, faixa II, fundamental – QFF 011).

(16b) INF: Meu pai pedia *pra eu deixá* meu marido

(Inf. 4 – 20: Rio Branco, mulher, faixa II, fundamental – PM 06).

6.1.2.1.3 Papel semântico do pronome sujeito

Examinou-se, com a observação desse grupo de fatores, ainda sob os aspectos semânticos da construção, o papel desempenhado semanticamente pelas formas pronominais ora analisadas. Foram encontrados nos *corpora* três papéis semânticos, a saber: i. agente – aquele capaz de desencadear uma ação, responsável por agir com controle –; ii. beneficiário – aquele ente que recebe o benefício da ação descrita –; iii. experienciador – aquele ser animado que sofreu mudança ou está em certo estado mental, perceptual ou psicológico (CANÇADO, 2005, p.111-112). A hipótese testada é a de que os papéis de agente e beneficiário tendem ao pronome *eu*.

a. Agente;

(17a) INF: Eu faço tudo *pra mim andá* certo com as minha coisa.

(Inf. 04 – 243: Porto Alegre, feminino, faixa II, fundamental – TDS 01).

(17b) INQ: E que você faz durante o dia? [...]

INF: Então, então vai lá pra sala reparti vai, vai, vai lá, vai lá reparti e... né, tem dia que a gente tira o dia pra lavá, tem um dia que tira pra passá né, e eu tamém, assim, eu quando sobra um tempinho, que às vezes num sobra, mas eu tenho que fazê esse tempo sobrá *pra mim estuda* a música (risos).

(Inf. 02 –123: Goiânia, mulher, faixa I, fundamental – QMS 033).

b. Beneficiário;

(18a) INF: Assim que eu gosto de matar...*pra eu descansar* um pouco a cabeça, sabe?

(Inq. 293. Rio de Janeiro, feminino, faixa II, universitário – DID).

(18b) INF: Ter isso implica em, em realmente ser alguma coisa a mais *pra mim, pra eu, pra eu estar* me sentindo bem ou não, entende?

(Inq. 34: Rio de Janeiro, feminino, faixa I, universitário – DID).

c. Experienciador;

(19a) INF: Tô gripada né... num sei, *pra mim sintí* chero é difícil.

(Inf. 04 – 06: Manaus, homem, faixa II, fundamental – QSL 041).

(19b) INQ: - Bom, você falou que tem irmãos, né? Agora me fala os nomes deles. [...]

INQ: - Tão namorando?! (risos) Aí, que bom, né? Vocês estão contentes?

INF: - É, que nem eu canso de falá, né? É um sonho *pra mim ver* meu pai e minha mãe junto de novo.

(Inf. 01 – 220: Curitiba, homem, faixa I, fundamental – QMS 01).

6.1.2.1.4 Função sintática da oração infinitiva

Mediante o controle desse grupo, buscou-se observar a relevância da função sintática exercida pela oração infinitiva em correlação com a oração principal. A hipótese é de que as orações adverbiais finais, em função do caráter finalístico da oração, sejam favorecedoras ao pronome *eu*. Para isso, estabeleceram-se cinco fatores:

a. orações substantivas subjetivas;

(20a) INF: ai...agora tem tantas (que) *pra mim lembrar* é difícil...

(Inq. 121. Porto Alegre, feminino, faixa I, universitário – DID).

(20b) INF: É muito fácil *pra mim estabelecer* amizades, embora eu seja, de um certo modo, fechada.

(Inq. 56. Rio de Janeiro, feminino, faixa III, universitário – DID).

b. orações substantivas objetivas;

(21a) INF: A pessoa pediu *pra mim guardar* ou eu guardei?

(Inf. 04 – 243: Porto Alegre, feminino, faixa II, fundamental – QMS 042).

(21b) INQ: Você / o a senhor a tinha que trazer uma encomenda para alguém. Mas você o a senhor a não fez isso. Se a pessoa perguntar pela encomenda, o que é que você / o a senhor a diz?

“Infelizmente, eu não a encomenda”.

INF: Eu não pude trazer a encomenda. Infelizmente não deu *pra, pra eu trazê*.

INQ: Aí ela vai perguntar assim: “você... você”.

INF: Você não pôde trazer?

INQ: “Cadê minha encomenda?”

INF: “Não deu *pra, pra eu trazer*”. É isso?

(Inf. 05 – 093: Salvador, homem, faixa I, universitário – QMS 041).

c. orações adverbiais finais;

(22a) INF: Então ele me deu um rádio pequenininho *pra eu poder transportar*.

(Inq 162. Rio de Janeiro, feminino, faixa II, universitário – DID).

(22b) INF: Eu vou perguntando logo que é *pra eu saber* qual é a posição que eu vou tomar.

(Inq 237. Rio de Janeiro, masculino, faixa III, universitário – DID).

d. orações completivas nominais;

(23a) INF: É um sonho *pra mim ver* meu pai e minha mãe junto de novo.

(Inf. 01 – 220: Curitiba, homem, faixa II, fundamental – QMS 01).

(23b) INF: tô doido *pra mim arrumar* um dinherin’ a mais...

(Inf. 03 – 123: Goiânia, homem, faixa II, fundamental – QMS 044).

6.1.2.1.5 Conjugação verbal

Examinou-se, com o controle dessa variável, a influência da natureza do verbo no infinitivo sobre o fenômeno em observação. Hipotetiza-se que os verbos terminados em -ar e -er, por serem em maior número no PB, são mais proeminentes, quando da utilização do pronome *eu*. Com isso, delimitaram-se três possibilidades:

a. verbos da 1ª conjugação, terminados em (-ar);

(24a) INQ: O que você / o (a) senhor (a) faria se ganhasse na loteria?

INF: Primeiramente, iria resolvê os problemas econômico-financêros da família como um todo. Faria doações. A posteriori, iria conversar com o professor Alcir que já conhece o mundo todo para me dizer os melhores locais *para eu viajar*.

(Inf. 06 – 034: Teresina, mulher, faixa I, universitário – QMS 044).

(24b) INQ: – Alguém pede para os enhor catalogar uma pedra. E aí o senhor deixou passar um tempinho tal, e vem um outro geólogo pra catalogar, o senhor diz: “Não, não, esse aí é pra...” quem catalogar?

INF: – *Pra mim catalogar, né?*

(Inf. 07 – 03: Boa Vista, homem, faixa II, universitário – QMS 023).

b. verbos da 2ª conjugação, terminados em (-er);

(25a) INQ: Alguém pede para você / o (a) senhor (a) fazer uma tarefa. Mas outra pessoa acha que a tarefa era para ela. Então, você / o (a) senhor (a) diz: Essa tarefa, na verdade, é para _____ fazer.

INF: *Para mim fazê.*

(Inf. 02 – 061: João Pessoa, mulher, faixa I, fundamental – QMS 023).

(25b) INF: eu tenho realmente muito cuidado...*e pra eu manter* isso na hora do almoço eu não posso carregar demais na hora da, do café.

(Inq. 328. Rio de Janeiro, mulher, faixa III, universitário – DID).

c. verbos da 3ª conjugação, terminados em (-ir);

(26a) INF: Se bem que eu uso a feira...*pra eu sair*, me deslocar pra zona sul pra comprar um legume, uma fruta, eu acho que não é econômico, entendeu?

(Inq. 273. Rio de Janeiro, feminino, faixa II, universitário – DID).

(26b) INQ: – O senhor ouve rádio bem alto ou baixinho?

INF: – Hum... médio... de..., de... depende, depende do barulho que tá em volta. Quando eu tô no carro, eu boto um pouquinho mais alto, *pra eu ouvir* a música.

(Inf. 05 – 243: Porto Alegre, homem, faixa I, universitário – QMS 035).

6.1.2.1.6 Forma da preposição

Por meio do controle desse grupo, buscou-se observar como a forma da preposição incide na utilização dos pronomes. Infere-se que, ao observar as escolhas preposicionadas e pronominais, há correlação, de forma proeminente, entre a preposição plena *para* e o pronome canônico *eu*. Nesse sentido, estabeleceram-se dois fatores:

a. Forma plena (*para*);

(27a) INQ: Alguém pede para você / o (a) senhor (a) fazer uma tarefa. Mas outra pessoa acha que a tarefa era para ela. Então, você / o (a) senhor (a) diz: Essa tarefa, na verdade, é para ____ fazer.

INF: *Para eu fazer.*

(Inf. 05 – 61: João Pessoa, homem, faixa I, universitário – QMS 023).

(27b) INQ: - Eu peço pra você, por exemplo levar uma carta no correio, né? Aí você foi até lá fora, chegou a Rosana e pegou a carta. Aí você fala: “Não, Rosana! Essa carta é para...” INF.- Para mim.

INQ: - Essa carta é pra... levar ao correio. É para...

INF: - É *para mim levar* ao correio.

(Inf. 01 – 220: Curitiba, homem, faixa I, fundamental – QMS 023).

b. Forma contracta (*pra*);

(28a) INQ: Fala um pouquinho do seu trabalho, quando você trabalhava. Esse último (...)

INF: Eu não sou trabalhadora, num gosto de trabalhar. Mais eu cheguei, digo: “Vou fazer, aquilo é *pra mim fazer*.” Pra mim, eu faço... faço bem assim.

(Inf. 04 – 03: Boa Vista, mulher, faixa I, fundamental – TDS 03).

(28b) INQ: Alguém pede pra você é... fechar uma porta, aí você fecha não fecha a porta, aí a pessoa vai lá pra fechar você fala não deixa essa porta é pra... quem fechar? INF.- *Pra eu fechar*.

(Inf. 06 – 06: Manaus, mulher, faixa I, universitário – QMS 023).

6.1.3 Variáveis extralinguísticas

Para escolha e definição das variáveis extralinguísticas, optou-se por aquelas que podem ter influência na fala dos informantes. Salienta-se que, como os *corpora* possuem naturezas diferentes, foram eleitas as seguintes variáveis para o Projeto NURC: década; faixa etária; localidade; e região geográfica. Ao passo que, para o Projeto ALiB, as variáveis foram: sexo; faixa etária; escolaridade; localidade; e região geográfica.

6.1.3.1 Variáveis extralinguísticas – Projeto NURC

Conforme já se comentou anteriormente e segundo as peculiaridades dos informantes que compõem as amostras, para os dados do Projeto NURC, no que tange ao extralinguístico, controlam-se quatro variáveis, a saber:

- a. Década: (70 e 90);
- b. Faixa etária: (faixa 1: 25 a 35 anos; faixa 2: 36 a 55 anos; faixa 3: mais de 56 anos);
- c. Localidade: (Porto Alegre; Rio de Janeiro; São Paulo; Salvador e Recife);
- d. Região Geográfica (Sul; Sudeste e Nordeste).

6.1.3.2 Variáveis extralinguísticas – Projeto ALiB

Consoante as particularidades dos informantes que compõem os dois *corpora*, para os dados do Projeto ALiB, no que tange ao extralinguístico, controlam-se cinco variáveis:

- a. Sexo: (Homem e Mulher);
- b. Faixa etária: (faixa etária I: 18 a 30 anos; faixa etária II: 50 a 65 anos);
- c. Escolaridade: (fundamental e universitário);
- d. Localidade: (25 capitais brasileiras);
- e. Região Geográfica: (Sul, Sudeste, Centro-Oeste, Nordeste e Norte).

6.1.4 Variáveis linguísticas e extralinguísticas – agrupamento dos *corpora*

Na definição das variáveis, por ocasião do agrupamento das amostras dos dois Projetos, priorizaram-se os possíveis aspectos intercomparáveis entre eles, a saber:

a. linguísticos: (natureza semântica do verbo da oração principal; papel semântico do pronome sujeito; função sintática da oração infinitiva; conjugação verbal; e forma da preposição);

b. extralinguísticos: (década: 70/90 – Projeto NURC; 2000 – Projeto ALiB; localidade: cinco capitais coincidentes – Porto Alegre; Rio de Janeiro; São Paulo; Salvador e Recife; região geográfica – Sul, Sudeste e Nordeste).

Ressalta-se que todos os informantes, nesse referido agrupamento, possuem o nível universitário, tanto os do Projeto NURC quanto os do Projeto ALiB, ou seja, os dados orais dos informantes pertencentes ao nível fundamental, *corpus* do Projeto ALiB, foram retirados para essa etapa.

6.1.5 A codificação dos dados

Codificar dados linguísticos é uma arte. Esse ato, por muitas vezes, árduo e minucioso, visa a transformar os dados linguísticos coletados em uma sequência de caracteres legíveis para determinado programa estatístico. Desse modo, códigos são utilizados para definição das variáveis em estudo, bem como para a criação dos grupos de fatores. Para a codificação dos dados, foi utilizado o programa editor do Bloco de Notas do Windows.

Essa etapa é primordial e antecede a etapa da submissão e do processamento dos dados, devidamente, preparados para que o programa estatístico revele, em números, as relevâncias e as predominâncias de determinados grupos de fatores. Assim, dando indícios sobre os fatores que favorecem e/ou desfavorecem a realização de um dado fenômeno linguístico.

6.1.6 O *Goldvarb X*

A utilização do programa de regra variável para o tratamento dos dados permite ao pesquisador obter mais credibilidade em suas afirmações, uma vez que a ferramenta estatística e matemática fornece aparatos para eliminar desconfiças sobre o fenômeno, trazendo maior credibilidade para o estudo, desse modo, afastando algumas interferências

subjetivas. A incorporação de programas computacionais na análise de dados linguísticos é oriunda das postulações sociolinguísticas, haja vista que a variação linguística é passível de ser mensurada, sob diversos efeitos dos variados fatores envolvidos nos processos de variação e mudança linguísticas.

O pacote de programas estatísticos *Goldvarb X* (SANKOFF; TAGLIAMONTE; SMITH, 2005) é bastante utilizado em estudos variacionistas, haja vista que ele fornece aos pesquisadores os resultados sobre a atuação de fatores individuais de cada variável explanatória sobre a variável dependente. Com a utilização dessa ferramenta, podem ser obtidos: a frequência de realização das variantes; o peso relativo de cada um dos fatores que compõe as variáveis; a seleção das variáveis relevantes para o favorecimento do fenômeno; além do nível de significância dos resultados quantificados pelo *software* específico para este fim, *Goldvarb X*, que consiste na versão para *Windows* do pacote de programas *Varbrul* (LAWRENCE; TAGLIAMONTE, 2001).

Com isso, vale dizer que, a partir desse pacote de programas, é possível “[...] investigar situações em que a variável linguística em estudo é investigada por vários elementos do contexto, ou seja, múltiplas variáveis independentes sobre a ocorrência das realizações da variável que está sendo tratada como dependente” (GAY; ZILES, 2007, p. 105).

Inicialmente, a partir das variáveis independentes e os fatores estabelecidos para elas, o programa apresenta para o pesquisador valores que são globais e percentuais de ocorrências das variantes analisadas pelo estudo. Destaca-se que, nesse momento, por vezes, há casos chamados de *nocautes*, ou seja, há ausência de variação, 0% ou 100% de ocorrência de uma das variantes em análise. Quando isso acontece, precisa-se necessariamente adotar outros procedimentos, por exemplo, agrupar ou excluir os fatores em que apresentaram *nocautes*.

Posteriormente, outras rodadas são feitas, as chamadas rodadas multivariadas por duas maneiras: a) as variáveis são adicionadas e contrastadas entre si, *step-up.*; b) alguns grupos vão sendo retirados da análise e só permanecem aqueles os quais o programa considerou como relevantes, *step-down*.

Vale dizer que o programa utiliza um molde matemático, logo, os valores em pesos sempre estarão entre 0 e 1,0. Assim, os valores que se aproximarem de zero indicam que a probabilidade de se encontrar a variante em análise, quando o fator está presente, é nula. Ao passo que os valores mais próximos a um indicam alto grau de favorecimento do fato observado. Já os que apresentam valores iguais ou mais próximos de 0,50 são neutros para a

análise da variável. No entanto, o pesquisador, de posse dos dados fornecidos pelo programa, precisa interpretá-los e cotejar com as hipóteses delineadas para a análise de cada grupo estabelecido.

6.1.7 A cartografia⁵³

Nesta subseção, *a priori*, pretende-se demonstrar alguns elementos importantes para que se entenda a representação linguística por meio da cartografia. *A posteriori*, serão descritos os parâmetros pelos quais as cartas da tese foram produzidas.

6.1.7.1 Afinal, o que é e para que serve uma carta linguística?

Como mencionado na subseção 2.1.2, a Geolinguística pluridimensional, hoje, no Brasil, pode ser representada por obras que vão desde os primórdios aos dias atuais, revelando facetas diversas desse método do trabalho dialetal. Assim, algumas explicações são pertinentes, com vistas ao entendimento sobre o que se convencionou nomear de carta linguística/ mapa linguístico e qual a finalidade destes.

Antes mesmo de apontar esses esclarecimentos, de forma breve, define-se um atlas linguístico como um produto geolinguístico cujo objetivo é reunir “[...] um conjunto de mapas de um território, mais ou menos vasto, que representam e localizam as realizações dos paradigmas linguísticos em estudo (de natureza fonética, lexical, morfológica ou sintáctica), registrando as respectivas variações geográficas...” (FERREIRA et al., 1996, p. 484). Como um empreendimento precioso, um atlas linguístico possui um valor de um instrumento capaz de documentar os registros de uma língua ainda viva, fornecendo possibilidades de interpretação e reflexão sobre ela, além de caminhos de novas investigações (FERREIRA et al., 1996).

Após o preâmbulo, busca-se, a partir das linhas seguintes, responder aos questionamentos próprios dessa subseção. Desse modo, a fim de que se considerem uma carta ou um mapa, o produto cartográfico “[...] deve apresentar informações essenciais como: orientação, sistema de projeção, sistema de referência para as coordenadas e escala...” (TELES; RIBEIRO, 2014, p. 115). E sua finalidade, além do conjunto delas compor um atlas, no que concerne a uma carta temática

⁵³ As primeiras versões das cartas com QR-Code foram criadas na plataforma digital QR-Code Generator, e as cartas ficaram disponíveis na plataforma Scribd.

[...] é a representação dos fenômenos localizáveis de qualquer natureza e de suas correlações, não se limitando apenas à representação do terreno. Permite, dessa forma, que sejam feitas análises sobre o tema (ou o assunto) cartografado e que se possam ser relacionadas as informações obtidas a partir de outros instrumentos (por exemplo, dados de censos ou de outros tipos de cadastros, simplesmente) (TELES; RIBEIRO, 2014, p. 115).

As cartas linguísticas, em relação à tipologia, podem ser expositivas – são aquelas que demonstram as respostas de uma unidade em um mapa, em que a tabulação dos dados é colocada em dimensão geográfica – interpretativas – são aquelas que fazem uma exposição de forma mais geral do fenômeno, ao recorrer a distribuição das formas encontradas com predominância entre as localidades estudadas (CHAMBERS; TRUDGILL, 1994). Ademais, nesse sentido, conforme intentos do pesquisador, as cartas linguísticas podem representar as variantes fonéticas de um vocábulo em estudo; variantes que designam um mesmo conceito, mesmo com variação fônica; variação de certos aspectos morfológicos; e, também, variações de estruturas sintáticas (BRANDÃO, 1991).

Acredita-se que as cartas de um atlas linguístico podem ser de grande serventia para o ensino, uma vez que aspectos da língua são, por meio delas, registrados, permitindo a contemplação, de forma científica, de como a língua está configurada em determinado espaço geográfico. Assim,

Não basta pensar em modas, como declínio da fonética, mas demandas linguísticas atuais métodos de trabalho muito mais refinados do que Gilliéron inventou. Cada um poderá decidir pelo procedimento que ele queira, ou que considere melhor; afinal, os dogmas eles estão fora da ciência. Mas devemos lembrar que depois da ALF já não era possível fazer cartografia linguística como como Weigand o projetou, tomo por exemplo; depois do AIS seria difícil tentar fazer uma cartografia linguística como na época de Gilliéron; após as experiências do NALF e as interrupções ou atrasos do ALR, do ALI e do ALPI, será necessário pensar em uma terceira era da cartografia linguística (ALVAR, 1969, p. 199)⁵⁴ [tradução nossa].

Nessa direção, ao pensar em uma Dialectologia feita para todos, linguistas e não linguistas, as cartas temáticas, de orientação geolinguística, necessitam de representações e/ou meios que facilitem a leitura, independentemente do tipo de instrução do leitor e/ou consulente. Dessa maneira, permitirão um maior acesso a elas e aos resultados do laborioso trabalho dialetal, sobretudo, com a união desse ramo dos estudos linguísticos que é, por

⁵⁴ No basta pensar en modas pasajeras, como el declive de la fonética, sino que la lingüística actual exige métodos de trabajo mucho más refinados que los que inventó Gilliéron. Cada uno podrá decidir por el procedimiento que quiera, o que considere mejor; después de todo, los dogmas están fuera de la ciencia. Pero debemos recordar que después del ALF ya no era posible hacer cartografía lingüística como la diseñó Weigand, por ejemplo; después del AIS sería difícil intentar hacer una cartografía lingüística como en la época de Gilliéron; Después de las experiencias del NALF y las interrupciones o retrasos del ALR, el ALI y el ALPI, será necesario pensar en una tercera era de la cartografía lingüística.

natureza, contextual e interdisciplinar, com outras áreas, usufruindo, cada vez mais, das TDICs.

6.1.7.2 As cartas linguísticas

A cartografia foi elaborada atentando-se para as semelhanças e diferenças linguísticas entre as formas em análise, a fim de fornecer um panorama de como se configuram os pronomes *eu* e *mim* em estrutura do tipo *para/prá + eu/mim + infinitivo* nas capitais brasileiras nos *corpora* do Projeto NURC e no *corpus* do Projeto ALiB.

Cartografar dados linguísticos é, sempre, a parte da pesquisa que requer da Dialetoлогия uma interface com outras áreas. No caso do atlas nacional, “[...] Os rigores estabelecidos pela Linguística e pela Cartografia têm de ser perseguidos e concretizados. Por essa razão, à equipe do ALiB foram incorporados profissionais das ciências cartográficas e da computação” (TELES; RIBEIRO, 2014, p. 113).

Desse modo, diante do exposto, registra-se que: (i) a análise linguística foi realizada pelo autor desta tese, ex-bolsista IC do Projeto ALiB, doutorando em Língua e Cultura pela UFBA; (ii) o planejamento cartográfico, a edição da base e a composição da cartografia temática foram elaboradas por Weldon Ribeiro Santos, ex-bolsista IC do Projeto ALiB, doutorando em Geografia pela UFBA, tendo como fonte a Base Cartográfica do IBGE (bCIMd), versão 2.0 e a utilização do *software* ARcGis (versão 1.0); e, por fim, (iii) os ajustes quanto aos aspectos da Ciência da Computação foram feitos por Mariana e Nicolas, funcionários de uma empresa de serviços gráficos.

As cartas elaboradas seguem alguns critérios para a cartografia temática, são eles: a) representação das formas documentadas nas amostras analisadas; b) consideração da frequência em localidades com três modos (formato em pizza, formato em barra e presença *versus* ausência); c) foram elaboradas cinco cartas linguísticas experimentais, duas cartas com dados do Projeto NURC, duas cartas com dados do Projeto ALiB e uma carta que agrega dados dos dois Projetos; d) essas recebem numeração de 01 a 05, o M corresponde a ser uma carta com dados morfossintáticos; e) indicação dos resultados em cada localidade pelo formato pizza, nas cartas experimentais M01, M02 e M04, com valores em percentuais; formato pizza, na Carta experimental M05, com valores presença *versus* ausência; e, por fim, formato em barras, na Carta experimental M03, com valores em peso relativo.

Ainda sobre a cartografia, as cartas estão subdivididas de modo a espelhar alguns aspectos, a saber: a Carta experimental M01: representa todas as ocorrências válidas dos

pronomes em estudo nos *corpora* do NURC, décadas de 70 e 90; a Carta experimental M02: representa todas as ocorrências válidas das formas *eu* e *mim* no *corpus* do Projeto ALiB; a Carta experimental M03: ainda com os dados do Projeto ALiB, mas espelha os resultados quanto à distribuição por nível de escolaridade, fundamental e universitário; a Carta experimental M04: representa o agrupamento dos *corpora* dos dois Projetos, com aspectos intercomparáveis, tais como as cinco localidades e os informantes universitários; por fim, a Carta experimental M05: traz a cartografia do “não dado”, a partir dos dados do Projeto NURC, Elocuções formais, evidenciando ser possível fornecer ao leitor tanto a informação do dado linguístico apurado quanto do dado não obtido.

Contemporaneamente, a sociedade vem utilizando invenções tecnológicas para facilitar e modernizar as relações, sejam elas em qualquer âmbito. Nesse sentido, o *Quick Response Code* - QR Code - se configura como um código de barras bidimensional, que possui uma extensa capacidade de armazenamento de dados, cuja função, quando escaneado, é direcionar para *links* de páginas da *web*. Geralmente, vem sendo utilizado para dinamizar as relações comerciais entre outras.

Assim sendo, ao utilizar o QR-code a favor da Linguística, em especial, da Dialetoлогия e da Sociolinguística, visa-se que os produtos desses ramos de estudos sobre a língua estejam mais facilmente ao alcance da sociedade, fornecendo, com base em pesquisas, informações confiáveis sobre a língua falada, auxiliando, de modo geral, o conhecimento sobre fenômenos linguísticos e, conseqüentemente, o ensino de língua portuguesa no Brasil. Os QR-codes foram do tipo estáticos e não expiram, também não sendo possível editá-los e nem rastrear e controlar os escaneres. Para tal fim, foi utilizada a plataforma digital ME-QR.com⁵⁵, local em que é possível a criação gratuitamente.

Quanto ao *locus* para hospedar a leitura/análise da carta feita pelo pesquisador, autor desta tese, optou-se por um ambiente virtual, que fosse público e de fácil acesso, e que não se solicita permissão, uma vez que o objetivo é popularizar o fazer dialetológico. Nessa direção, utilizou-se o mesmo site em que foi criado o QR-code.

A partir do movimento em congregar o método geolinguístico com outros, foi possível concretizar o desejo do linguista/dialetólogo em disponibilizar, de forma mais democrática, dados linguísticos, em especial, as cartas linguísticas. Acredita-se que as cartas linguísticas podem funcionar como instrumentos didático-pedagógicos eficazes para o ensino de línguas, inclusive, podendo ser utilizadas em salas de aula, quer sejam do Ensino Básico, quer sejam

⁵⁵ Mais informações estão disponíveis em <https://me-qr.com/>.

do Ensino Universitário, disponibilizadas em materiais de cunho pedagógico, como atividades e/ou livro didático.

Ademais, para além dos espaços de Educação formal, as cartas com QR-code podem possibilitar o maior alcance, quando utilizadas, por exemplo, nas redes sociais, fomentando o combate ao preconceito linguístico, pois os interessados terão acesso à leitura, interpretação e aos dados da pesquisa que geraram a carta. Para a Dialetologia e a Geolinguística Pluridimensional, o ganho é imensurável, haja vista que se pretende que os resultados sirvam a uma coletividade e, de preferência, que não se restrinja à comunidade acadêmica, mas que, de algum modo, possam contribuir para o progresso do fazer científico.

6.1.8 Metalinguística nos/dos dados

Ao coletar os dados, observou-se se os informantes, quer sejam do Projeto NURC, quer sejam do Projeto ALiB, emitiam alguma percepção durante a fala na realização de uma construção linguística tão cercada pelos estigmas sociolinguísticos, como a ora analisada. Nesse sentido, foram extraídas dos *corpora* algumas falas que revelam que tal estrutura da língua, ainda, é alvo de um monitoramento, quando requerida em certos contextos comunicativos.

Atentar-se para a variação diarreferencial (THUN, 2000, p. 191), explorando a percepção metalinguística dos informantes é de suma importância, pois pode fornecer caminhos para que, cada vez mais, os estudos sobre a língua falada ofereçam materiais empíricos que sirvam de combate ao preconceito e aos estigmas criados e fomentados no seio social, haja vista que

[...] os juízos de certo e de errado podem ser objeto de formação de um *corpus* não linguístico a ser analisado, juntamente com os juízos metalinguísticos/epilinguísticos e etnográficos em geral — que vem a constituir a dimensão diarreferencial —, com potencial de auxiliar o entendimento de condicionamentos externos à realidade de fala do agrupamento humano — situado num determinado tempo e num determinado espaço — focalizado pela pesquisa geolinguística (FIGUEIREDO JÚNIOR et al., 2021, p. 16).

A dimensão diarreferencial, as atitudes e as crenças dos informantes podem apontar para raízes dos estigmas e preconceito linguísticos, tais como os relacionados aos fenômenos fonéticos, lexicais, morfossintáticos e entonacionais. Nota-se, por vezes, que determinados fenômenos sofrem maiores vigilância e repressão. Nesse caso, nesta tese, a forma oblíqua *mim*, quando utilizada na posição de sujeito precedida da preposição *para* e em contexto específico, pode ser enquadrada no grupo de fenômenos que recebem monitoramentos

severos, isto é, o preconceito linguístico está relacionado com valoração pejorativa das formas diferentes da norma linguística do conhecimento individual, coletivo e/ou regional do universo dos falantes, uma vez que frente a consciência sociolinguística

[...] a atitude que é tomada por estes frente àquilo que ele conhece/acredita sobre sua variedade e a do outro. Tal consciência é composta por uma rede ampla e complexa de elementos sócio-históricos que podem ser acentuados, a depender, como verificamos, do grau de escolaridade do falante. (AGUILERA; SILVA; MORATTO, 2023, p. 216).

Conforme Figueiredo Junior (2019, p. 16), os juízos de valor encontrados nas elocuições dos informantes podem ser classificados em dois grandes polos, a saber: i. metalinguístico/epilinguístico – a) sinônimos; b) definições semânticas; c) jocosidade associada; d) informação diageracional; e) *status* temporal de formas: antigas vs. recentes; f) *status* popular vs. técnico de formas g) percepção sobre frequência; h) percepção de caída em desuso; i) juízos de valor negativo; j) juízos de “certo” vs. “errado –, ii. etnográfico – a) tabus; b) lendas; c) crenças religiosas; d) superstições laicas.

Na mesma direção da proposição de Figueiredo Jr., (2019), sugere-se pensar que os juízos de valor, emitidos pelas atitudes e pelas crenças, expressas pelos comentários metalinguísticos, podem ser classificados em: i. preconceito linguístico de xenofobia: a) interna: (regional/local); b) externa: (outro país); ii. preconceito linguístico sexista: sexo⁵⁶; iii. preconceito linguístico étnico-racial: negro (pretos e pardos), indígena, quilombola, cigano etc.; iv. preconceito linguístico educacionista: escolaridade (geralmente, as pessoas com pouca ou sem escolaridade); e, por fim, v. preconceito linguístico auto-metalinguístico (autoavaliação com a própria maneira de uso da língua).

Segundo as proposições de Thun (2000) e Aguilera (2008ab), sabe-se que as percepções dos falantes, no aspecto sobre a metalinguagem, são ferramentas que denunciam o modo subjetivo que eles enxergam a si, o outro, a língua, e, por consequência, os fenômenos dela decorrentes. Desse modo, por vezes, obtém-se, por meio das elocuições dos informantes, quais são as raízes que alimentam o estigma contra certas formas da língua, e quais os motivos de, no imaginário coletivo de uma comunidade linguística, algumas formas gozarem de privilégios em detrimento de outras.

⁵⁶ Ressalta-se que gênero e sexualidade carecem de definições próprias e mais assertivas. Portanto, não estão incluídas nessa proposição de preconceito linguístico de cunho sexista.

6.2 SÍNTESE DA SEÇÃO

Esta seção, 6, teve por meta descrever o passo a passo da pesquisa, com fins de facilitar o entendimento das escolhas teóricas e metodológicas utilizadas para tratar o fenômeno estudado.

Na subseção 6.1, explicitam-se os procedimentos adotados na tese, a partir do cotejo estabelecido em juntar dois *corpora* distintos em suas respectivas naturezas e propósitos, para além dos tempos cronológicos, 70/90/2000. Por fim, no 6.2, são expostos os procedimentos quanto ao levantamento e tratamento dos dados, bem como os critérios utilizados para o tratamento estatístico, por meio do *Golvarb* (2001). Além dos utilizados para o tratamento interpretativo dos dados, a partir da análise da metalinguagem e do não dado, as interfaces entre Geolinguística, Cartografia e TDICs, a fim da elaboração das cartas linguísticas com QR-code, tornando a Geolinguística brasileira a pioneira nesse quesito.

Vale mencionar a importância desta seção, haja vista que, por meio dela, obtém-se o designer metodológico desta tese. Logo, frisa-se que os critérios estabelecidos, tanto para o tratamento estatístico, quanto para o interpretativo, seguem aspectos que visam a espelhar os resultados obtidos, no que concerne ao fenômeno ora analisado.

Após essas exposições, na seção 6, demonstram-se a análise e os resultados dos dados catalogados, depois das investigações e correlações feitas.

7 RUMOS DA ANÁLISE DE DADOS

O que se deu pode perfeitamente ser comparado com o processo de multiplicação vegetal conhecido pelo nome de mergulhia: até se criarem raízes, o ramo vergado e mergulhado no solo é sustentado pela seiva do vegetal que sofre a operação; uma vez brotadas raízes que possam haurir seiva para a nova planta, um corte no ramo a separa e ela vai separadamente prosseguir sua evolução.

(NASCENTES, 1953, p. 11).

Nesta seção, serão descritos e discutidos os dados referentes às escolhas pronominais entre o pronome pessoal reto, *eu*, e o pronome oblíquo tônico, *mim*, em posição de sujeito, em orações infinitivas, precedidos pela preposição *para*, em estrutura tal como *para/prá + eu/mim + infinitivo*, obtidos por meio dos dados do Projeto NURC e Projeto ALiB, nas capitais brasileiras contempladas nesse estudo.

Inicialmente, a descrição e a discussão serão realizadas a partir das rodadas dos dados feitas pelo programa estatístico *Goldvarb X*, utilizando o *eu* como regra de aplicação, seguidas de leituras interpretativas com o arcabouço teórico dialetológico e sociolinguístico sobre os dados do Projeto NURC e, posteriormente, sobre os dados do Projeto ALiB. Em seguida, a leitura e a interpretação serão sobre o agrupamento feito com os dados das cinco capitais coincidentes nos dois *corpora*, a saber: Porto Alegre, São Paulo, Rio de Janeiro, Salvador e Recife, somente com dados de fala de informantes universitários. Por fim, no viés interpretativo/qualitativo, analisam-se os aspectos metalinguísticos encontrados e os aspectos relacionados às respostas não obtidas.

7.1 OBSERVAÇÕES INICIAIS – PROJETO NURC

A partir dos *corpora* das cinco capitais analisadas pelo Projeto NURC, Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre, nas décadas de 70 e 90, considerando as elocuições observadas em 307 inquéritos, foi obtida uma amostra de 144 ocorrências no cômputo total, das quais 50 ocorrências foram de dados com o zero anafórico, ou seja, 34,7%. Ao passo que foram contabilizadas 94 ocorrências válidas com a utilização dos pronomes *eu* e *mim*, 65,3%, conforme Tabela 1.

Tabela 1 – Uso pronominal na estrutura *para/prá + eu/mim + infinitivo* – dados das capitais do Projeto NURC: todas as respostas

Dados coletados	Número de ocorrências	%
Eu/Mim	94	65,3%
0	50	34,7%
Total:	144	100%

Fonte: Elaborado pelo autor (2023).

Dessas 94 ocorrências válidas⁵⁷, 31 ocorrências, ou seja, 33% foram do pronome oblíquo tônico *mim*, em estruturas tais como as analisadas pelo estudo, ao passo que 63 ocorrências, isto é, 67% foram da forma do caso reto, *eu*, como se ilustra no Gráfico 1.

Gráfico 1 – Uso pronominal na estrutura *para/prá + eu/mim + infinitivo* – dados das capitais do Projeto NURC: frequência geral das respostas válida



Fonte: Elaborado pelo autor (2023).

Desse modo, conforme o Gráfico 1, é notório que a forma *mim*, no cômputo geral dos dados, obtém um menor percentual de respostas válidas. Os dados referentes à estrutura sintática analisada permitem afirmar que, nos dados coletados pelas equipes do Projeto NURC, nas décadas de 70 e de 90, o pronome *eu*, em posição de sujeito (precedido da preposição *para* e seguido de verbo no infinitivo), é predominante na fala dos informantes escolarizados, com nível universitário, oriundos das cinco capitais elencadas para a composição dos *corpora* do Projeto NURC.

⁵⁷ Desse total, pouco mais de 10% são pertencentes aos anos 90.

7.2 OS CONDICIONADORES PARA A REALIZAÇÃO DO FENÔMENO – PROJETO NURC

Nessa subseção, serão apresentados alguns aspectos, os mais relevantes, para a obtenção dos resultados, ao utilizar o programa estatístico, conforme os dados do Projeto NURC, com vistas a delinear a configuração do fenômeno nesses *corpora*.

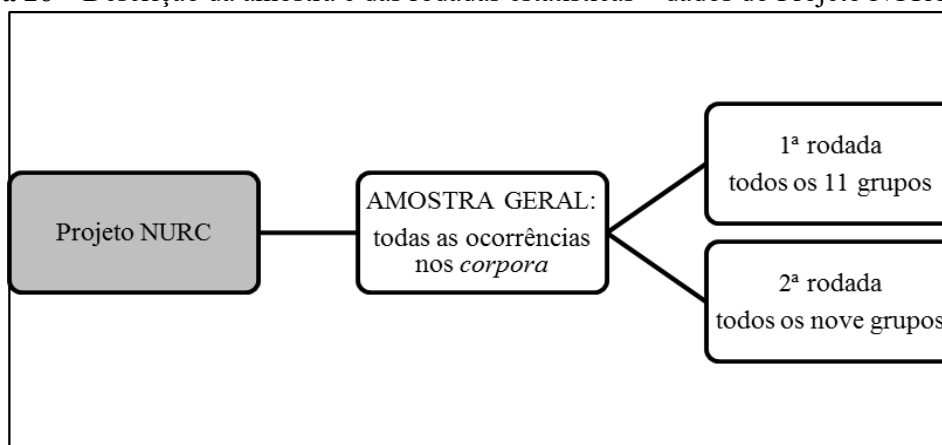
7.2.1 Alguns comentários sobre as rodadas no programa estatístico

Inicialmente, foi realizada uma rodada com todos os dados válidos relativos às elocuições dos 307 inquéritos das cinco capitais, das décadas de 70 e de 90. As variáveis controladas foram consideradas isoladamente, desse modo, cabe lembrar os grupos, conforme seção 6.4:

- (1) Variantes;
- (2) Estilo de registro;
- (3) Natureza semântica do verbo da oração principal;
- (4) Papel semântico do pronome sujeito;
- (5) Função sintática da oração infinitiva;
- (6) Conjugação verbal;
- (7) Forma da preposição;
- (8) Década: (70 e 90);
- (9) Faixa etária: (faixa 1: 25 a 35 anos; faixa 2: 36 a 55 anos; faixa 3: mais de 56 anos);
- (10) Localidade: (Porto Alegre; Rio de Janeiro; São Paulo; Salvador e Recife);
- (11) Região Geográfica (Sul; Sudeste e Nordeste).

Na primeira rodada, dos 11 grupos de fatores analisados, os grupos de fatores que apresentaram *nocautes* foram: estilo de registro, no que tange às Elocuições Formais; a forma da preposição, no caso da forma plena *para*; a década, no que concerne aos anos 90. Todos esses foram descartados pelo programa estatístico na observação da variação em estudo, pois apresentaram unanimidade para a forma *eu*, na estrutura *para/prá + eu + infinitivo*.

Para a segunda rodada, dois grupos foram descartados, forma da preposição e década, pois eram grupos binários, assim, ficando, apenas, nove grupos. Estes foram submetidos ao programa novamente. Diante dos resultados das rodadas iniciais, a fim de que se entendam os caminhos feitos para a obtenção dos valores percentuais e dos pesos relativos, as rodadas foram esquematizadas na Figura 26.

Figura 26 – Descrição da amostra e das rodadas estatísticas – dados do Projeto NURC

Fonte: Elaborado pelo autor (2023).

Salienta-se que, nas duas rodadas, 1ª e 2ª, os grupos foram considerados isoladamente.

7.2.2 As variáveis independentes – intralinguísticas

Neste subtópico, serão abordados os grupos de fatores intralinguísticos. A ordem da exposição se deu da seguinte maneira: inicialmente, o único grupo, desse conjunto, considerado relevante para o programa estatístico, foi a função sintática da oração infinitiva. Os outros grupos, a saber: o papel semântico do pronome sujeito, a natureza semântica do verbo da oração principal, a conjugação verbal, e, por fim, a forma da preposição não foram selecionados como estatisticamente relevantes pelo *Goldvarb X*.

7.2.2.1 Função sintática da oração infinitiva

A inclusão dessa variável na pesquisa decorreu da necessidade de controlar a relevância da função sintática exercida pela oração infinitiva. Para isso, observando-a em quatro aspectos, conforme contextos dos usos catalogados, a saber: as orações substantivas subjetivas; as orações substantivas objetivas; as orações adverbiais finais; e as orações completivas nominais. Dispõem-se, na Tabela 2, em ordem decrescente, os valores obtidos, ao considerar os pesos relativos.

Tabela 2 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme função sintática da oração infinitiva – dados do Projeto NURC

SUBSTANTIVAS OBJETIVAS		
APLIC./TOTAL	%	P.R.
9/10	90%	0.69
ADVERBIAIS FINAIS		
APLIC./TOTAL	%	P.R.
46/54	85,2%	0.59
SUBSTANTIVAS SUBJETIVAS		
APLIC./TOTAL	%	P.R.
2/5	40%	0.14
COMPLETIVAS NOMINAIS		
APLIC./TOTAL	%	P.R.
2/8	25%	0.07

Input: 0.798 Significância: 0,001

Fonte: Elaborado pelo autor (2023).

Diante desses valores, vê-se que as orações substantivas objetivas e as orações adverbiais finais são as favorecedoras ao pronome *eu*, com pesos relativos, respectivos, 0.69 e 0.59. Ao passo que os dois outros tipos de orações, substantivas subjetivas e completivas nominais, com pesos relativos de 0.14 e 0.07, são desfavorecedoras. Essa variável – função sintática da oração infinitiva – na pesquisa de Figueiredo (2007), foi considerada como a mais relevante.

Como foi possível observar, então, corroborando os resultados da pesquisa do referido autor, a função sintática exercida pela oração infinitiva é um fator que deve ser levado em consideração, quando observados os pronomes ora estudados.

7.2.2.2 Papel semântico do pronome sujeito

Procurava-se, por meio da observação dessa variável, averiguar se o papel desempenhado pelo pronome sujeito exercia influência na escolha pronominal da oração subordinada. Na Tabela 3, encontram-se dispostos os valores obtidos.

Tabela 3 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme papel semântico do pronome sujeito – dados do Projeto NURC

EXPERIENCIADOR		
APLIC./TOTAL	%	P.R.
6/8	75%	0.71
AGENTE		
APLIC./TOTAL	%	P.R.
46/56	82,1%	0.53
BENEFICIÁRIO		
APLIC./TOTAL	%	P.R.
7/13	53.8%	0.24

Input: 0.804 Significância: 0.002

Fonte: Elaborado pelo autor (2023).

A partir desses valores, nota-se que o papel semântico de experienciador é favorecedor quanto à utilização do pronome *eu*, em posição de sujeito, com peso relativo de 0.71. O papel semântico de agente, com peso 0.53, aparece próximo ao ponto de neutralidade. Ao passo que o beneficiário apresenta desfavorecimento quanto à utilização do *eu*, com peso relativo de 0.24.

Em contraponto, quando observada essa variável nos dados do estudo de Gomes (2019), os papéis semânticos do pronome sujeito agente, ambíguo e beneficiário e/ou experienciador são mais proeminentes para as variantes *mim* e para o *zero anafórico*. Ao passo que, para a forma *eu*, os valores obtidos foram bem menores.

7.2.2.3 Natureza semântica do verbo da oração principal

Buscou-se controlar esse grupo de fatores a fim de observar a relevância do papel semântico exercido pelo verbo que se encontra na oração principal da sentença. Estabeleceram-se como fatores os verbos transferenciais; existenciais; estativos; ação ou mudança de estado; elocutivos; e avaliativos e modais. Os resultados obtidos para esse grupo constam na Tabela 4.

Tabela 4 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme natureza semântica do verbo da oração principal – dados do Projeto NURC

EXISTENCIAIS		
APLIC./TOTAL	%	P.R.
7/9	77.8%	0.93
TRANSFERENCIAIS		
APLIC./TOTAL	%	P.R.
9/10	90%	0.72
ELOCUTIVAS		
APLIC./TOTAL	%	P.R.
5/6	83.3%	0.61
AÇÃO/MUDANÇA DE ESTADO		
APLIC./TOTAL	%	P.R.
22/29	75.9%	0.40
ESTATIVOS		
APLIC./TOTAL	%	P.R.
10/13	76.9%	0.30
AVALIATIVOS E MODAIS		
APLIC./TOTAL	%	P.R.
6/10	60.0%	0.21

Input:0.850 Significância: 0.009

Fonte: Elaborado pelo autor (2023).

Com base nos valores apresentados na Tabela 4, a situação dos pronomes em estudo demonstrou valores em pesos relativos que permitem subdividi-los em dois grupos: i. os que apresentam um comportamento favorável à utilização do pronome do caso reto, os existenciais, com peso relativo 0.93; os transferenciais, com peso relativo de 0.72; e os elocutivos, com peso relativo de 0.61; ii. Em contraste, no outro grupo, estão ação e mudança de estado, com peso relativo 0.40; os estativos, com peso relativo 0,30; e, por fim, os avaliativos/modais, com peso relativo 0.21.

Em comparação, no estudo de Gomes (2019), os verbos elocutivos e os transferenciais também são os mais proeminentes para os contextos de usos dos pronomes de primeira pessoa. Nesse sentido, é possível hipotetizar que os verbos bitransitivos são favorecedores quanto ao pronome *eu*.

7.2.2.4 Conjugação Verbal

Tencionou-se, com o controle desse grupo de fatores, avaliar as possíveis interferências das três conjugações verbais, quanto ao verbo da oração infinitiva, tendo em vista que, na língua portuguesa falada em solo brasileiro, a mais produtiva é a terminada em -ar. Os resultados obtidos se encontram na Tabela 5.

Tabela 5 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme conjugação verbal – dados do Projeto NURC

2 ^a (-ER)		
APLIC./TOTAL	%	P.R.
21/26	80.8%	0.56
1 ^a (-AR)		
APLIC./TOTAL	%	P.R.
27/37	73%	0.48
3 ^a (-IR)		
APLIC./TOTAL	%	P.R.
11/14	78.6%	0.41

Input: 0.850

Significância: 0.009

Fonte: Elaborado pelo autor (2023).

Constata-se, por meio dos resultados expostos na Tabela 5, que os verbos terminados na segunda conjugação, com peso relativo de 0.56, possuem o maior peso relativo, tendendo ao favorecimento da forma *eu*. Os verbos terminados na terceira e primeira conjugações, com pesos relativos 0.48 e 0.41 demonstram um aparente desfavorecimento, embora a terceira conjugações esteja próxima do ponto neutro.

Inferências de várias ordens são passíveis de serem feitas, pois se há uma produtividade maior de verbos terminados em -ar no PB, eles estariam sendo utilizados com maior frequência com a forma inovadora, *mim*, na estrutura analisada? Dados mais recentes do PB falado nas capitais podem responder ao referido questionamento, podendo ofertar caminhos de explicação sobre a conjugação verbal e as possíveis correlações com os pronomes *eu* e *mim* em posição de sujeito.

7.2.2.5 Estilo de registro

A partir do controle desse grupo de fatores, busca-se compreender a influência sobre os usos dos pronomes *eu* e *mim*, conforme os tipos de inquéritos utilizados pelas equipes do

Projeto NURC, quer sejam: EF - Elocução Formal; DID - Diálogo entre informante e documentador; e D2 - Diálogo entre dois informantes.

Da observação geral, como já notificado no início desta seção, esse grupo de fator não foi considerado relevante pelo programa estatístico. Nas Elocuções formais, como já referido, a ocorrência foi categórica quanto ao pronome do caso reto. Vale destacar que, ao considerar os diferentes graus de monitoramento, nesse tipo de gravação, os informantes estavam submetidos a situações de uso formal da língua, ou seja, aulas, conferências, palestras e etc. Logo, os dados catalogados refletem usos linguísticos formais, quando considerados os contextos em que foram gravados. Os resultados que estão expostos na Tabela 6 referem-se aos outros tipos de registro.

Tabela 6 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme estilo de registro – dados do Projeto NURC

D2		
APLIC./TOTAL	%	P.R.
9/12	75%	0.77
DID		
APLIC./TOTAL	%	P.R.
49/63	77.8%	0.45

Input: 0.850 Significância: 0.009

Fonte: Elaborado pelo autor (2023).

Como é possível observar, a maior probabilidade de ocorrência do *eu*, na estrutura em análise, pode ser percebida no D2 - Diálogo entre dois informantes, com peso relativo de 0.77. Com isso, o DID - Diálogo entre informante e documentador é, conforme o peso relativo de 0.45, um contexto desfavorecedor quanto à utilização do *eu*. Comprovou-se, assim, que, nos contextos em diálogo com outros informantes, o pronome *eu* é favorecido, embora, nos contextos em que há diálogos com o documentador, em que se espera um maior monitoramento linguístico, há um número maior de ocorrências catalogadas.

Desse modo, algumas ilações são possíveis, tais como: no D2, os monitoramentos são mais frequentes, pois, geralmente, o diálogo é entre pessoas sem relação de proximidade? No DID, o documentador agiu de modo a atenuar o paradoxo do observador? Tais respostas podem ser, quiçá, fomentadoras de estudos futuros sobre as amostras do referido Projeto e sobre as possíveis interferências na interlocução entre o documentador e o entrevistado.

7.2.2.6 Forma da preposição

Pretendeu-se, a partir do estabelecimento dessa variável, avaliar influência sobre os usos dos pronomes *eu* e *mim*, conforme as formas da preposição, quer seja na forma plena, *para*, ou na forma contracta, *pra*. Como já notificado no início desta seção, esse grupo de fator não foi considerado relevante pelo programa estatístico. A forma contracta, com peso relativo de 0.53, demonstrou ter uma leve tendência ao favorecimento à utilização da forma pronominal *eu*, embora esteja também próxima do ponto neutro. Ao passo que, a forma preposicional *para* demonstrou-se categórica ao pronome do caso reto.

A leitura sobre esses dados alude a uma associação entre a forma plena, *para*, e o pronome *eu*, canônico para a função de sujeito na estrutura em análise. Tal contexto é, inclusive, esperado, quanto se observa o grau de monitoramento da fala dos informantes, embora a forma contracta já figure, em muitos contextos, tanto com a forma pronominal canônica quanto com a forma inovadora.

7.2.3 As variáveis independentes – extralinguísticas

Nesta subseção, serão abordados os grupos de fatores de ordem extralinguísticos. A ordem da exposição se deu da seguinte maneira: a região geográfica, em seguida, a localidade, a faixa etária e, por fim, a década. Ressalta-se que nenhum grupo extralinguístico foi considerado como relevante pelo *Goldvarb X*.

7.2.3.1 A distribuição diatópica – região geográfica

Com o controle da variável região geográfica, buscou-se observar se o fenômeno se encontraria estratificado nas diferentes regiões geográficas brasileiras, desse modo, sendo possível apontar indícios referentes às implicações sociais do fenômeno no espaço regional.

Na Tabela 7, a partir dos valores globais, percentuais e pesos relativos obtidos, observa-se que, no panorama das regiões geográficas brasileiras, no que tange à escolha pronominal em contexto de *para/pra + eu/mim + infinitivo*, a região Sudeste, com peso relativo de 0.52, quanto à utilização do *eu*, apresenta maior peso relativo.

Tabela 7 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme a região geográfica – dados do Projeto NURC

SUDESTE		
APLIC./TOTAL	%	P.R.
46/58	79.3%	0.52
NORDESTE		
APLIC./TOTAL	%	P.R.
11/14	78.6%	0.45
SUL		
APLIC./TOTAL	%	P.R.
2/5	40%	0.39

Input: 0.850 *Significância:* 0.009

Fonte: Elaborado pelo autor (2023).

A partir dos pesos relativos demonstrados na Tabela 7, percebe-se que, nas três regiões geográficas, Sudeste, com peso de 0.52; Nordeste, com peso de 0.45; e, por fim, Sul, com peso de 0.39, não se pode afirmar favorecimentos quanto à utilização da forma canônica, *eu*. Esses dados, portanto, embora demonstrem que a região Sudeste, nos *corpora* analisados, do Projeto NURC, apresenta um peso relativo superior as demais regiões, nota-se uma proximidade com o ponto neutro. Ao passo que as regiões Nordeste e Sul apresentam valores com pesos relativos que estão desfavorecendo a forma pronominal canônica.

7.2.3.2 A distribuição diatópica – localidade

Buscou-se, por meio do controle dessa variável, a confirmação de uma das hipóteses principais do estudo: o fenômeno tem um comportamento diferente em determinadas localidades brasileiras. Os resultados estão apresentados na Tabela 8.

Tabela 8 – O pronome *eu* em posição de sujeito, nas capitais do Brasil – distribuição conforme a localidade – dados do Projeto NURC

CAPITAL	APLIC./TOTAL	%	PESO RELATIVO
São Paulo	4/5	80%	0.61
Rio de Janeiro	42/53	79.2%	0.60
Salvador	8/9	88.9%	0.54
Recife	3/5	60%	0.08
Porto Alegre	2/5	40%	0.06

Input: 0.850 *Significância:* 0.009

Fonte: Elaborado pelo autor (2023).

Embora os dados numéricos sejam quantitativamente poucos, verifica-se que, por meio deles, alguns indícios, quanto à escolha pronominal da forma *eu*, São Paulo e Rio de Janeiro, com pesos relativos, 0.61 e 0.60, respectivamente, são capitais favorecedoras. Salvador, com 0.54, apresenta-se próxima ao ponto de neutralidade.

Por fim, as capitais Recife e Porto Alegre são desfavorecedoras, com pesos relativos, respectivos, 0.08 e 0.06, quando observados os dados do Projeto NURC dos anos 70, uma vez que, nos anos 90, foram encontrados dados categóricos quanto ao pronome *eu*. A Figura 27 demonstra a configuração espacial dos resultados, considerando os valores percentuais.

Figura 27 – Carta experimental M01 – distribuição dos pronomes *eu* e *mim* quanto à localidade – dados do Projeto NURC.



Fonte: Elaborado pelo autor (2023).

Diatopicamente, no que tange à localidade, observa-se, por meio da Carta experimental M01, um aspecto que leva a ilações no sentido do século e da data de fundação dessas cinco capitais, conforme a Tabela 8 e a Figura 27. Considerando os valores percentuais, haja vista que a capital mais nova, criada no século XVIII, Porto Alegre, apresenta um comportamento

que tendência para utilização da forma inovadora, *mim*, enquanto as outras capitais, criadas no século XVI, Recife, Salvador, Rio de Janeiro e São Paulo, portanto, mais antigas, tendem a preferir a forma *eu*, a forma pronominal que goza de prestígio na norma linguística brasileira.

Desse modo, um questionamento é pertinente, a saber: as capitais mais antigas apresentam um comportamento linguístico mais conservador, por terem sido fundadas na primeira fase do Brasil colônia, se aproximando, dessa maneira, do padrão lusitano?

Diante disso, ainda conforme os fatos históricos e sociais, o comportamento diferenciado da capital sulista, tais como a localização geográfica, ao Sul do Brasil, em fronteira com variedades da língua espanhola – Argentina e Uruguai –, além da presença e contato com as línguas de imigração podem ser aspectos que moldem essas escolhas pronominais. Sugerem-se, então, pesquisas que tenham a dimensão diatópico-cinética (THUN, 1998), a fim de apurar com mais detalhes as suposições aqui levantadas.

7.2.3.3 A distribuição por faixa etária

Buscou-se, por meio do controle dessa variável, verificar se a faixa etária exercia influência sobre os usos das duas formas pronominais em estudo. Nesse sentido, podendo demonstrar aspectos sobre a ação do tempo sobre a língua.

Tabela 9 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme faixa etária – dados do Projeto NURC

FAIXA II (36 – 55 anos)		
APLIC./TOTAL	%	P.R.
25/31	80.6%	0.56
FAIXA III (+ de 56 anos)		
APLIC./TOTAL	%	P.R.
12/15	80%	0.53
FAIXA I (25 – 35 anos)		
APLIC./TOTAL	%	P.R.
22/31	71%	0.42

Input: 0.850 *Significância:*0.009

Fonte: Elaborado pelo autor (2023).

Observa-se que, conforme os dados dispostos na Tabela 9, os valores percentuais e relativos entre os informantes da faixa II e da faixa III são bem próximos, indicando que os informantes da faixa intermediária, faixa II, são usuários da forma de prestígio, com pesos

relativos de 0.56, favorecedores ao pronome *eu*. Os informantes mais velhos, da faixa III, apresentam um comportamento linguístico, conforme os dados, que não podem ser determinantes quanto ao favorecimento, com 0.53, pela proximidade com o ponto neutro. Logo, é possível, de imediato, a partir dos valores apresentados, notar que os informantes que possuem a faixa etária menor, os mais jovens, tendem à forma inovadora, com peso relativo de 0.41, desfavorecedor quanto ao pronome *eu*.

Esses dados revelam uma direção oposta aos que foram encontrados por Carvalho e Almeida (2013), na variedade do português falado em Feira de Santana, na Bahia, quando as autoras demonstram que, sobre o recorte diageracional, a fala dos informantes mais velhos apontam para a preferência da forma oblíqua *mim*, ou seja, os informantes com faixas etárias mais avançadas tendem ao uso da forma inovadora, ao passo que os mais jovens tendem ao conservador, conforme estudo das referidas autoras.

7.2.3.4 A distribuição por década

A partir do controle dessa variável, buscou-se demonstrar se os usos dos pronomes *eu* e *mim*, catalogados em duas décadas, 70 e 90, foram mais proeminentes, podendo, quiçá, evidenciar aspectos sobre a variação dessas formas pronominais e fornecer caminhos de investigação quanto à concorrência delas para o preenchimento da posição de sujeito.

Como já notificado no início desta seção, esse grupo de fatores não foi considerado relevante pelo programa estatístico. Nos anos 70, foram contabilizadas 47 ocorrências do *eu*, com valores percentuais de 74.6%. Como todos os dados catalogados nos anos 90 foram categóricos quanto ao pronome do caso reto, *eu*, esse grupo não foi posto na segunda rodada.

Nesse sentido, fatores diversos podem ser aludidos para a interpretação e para a leitura desses resultados, tais como o grau de monitoramento quanto ao estilo da documentação linguística, uma vez que, os dados dos anos 90, são oriundos do DID/RJ e do DID/SSA e das EF/SSA, ou seja, são situações de usos da língua que requerem um maior monitoramento por parte dos informantes, sobretudo, no que tange às EFs. Salienta-se que somente duas capitais, Rio de Janeiro e Salvador, possuem dados pertencentes à década de 90, ou seja, as duas primeiras capitais do Brasil nos dados recoletados e nas novas entrevistas, são categóricos quanto à forma de prestígio, *eu*.

7.3 OBSERVAÇÕES INICIAIS – PROJETO ALiB

A partir do *corpus* das 25 capitais analisadas, considerando as elocuições observadas dos 200 informantes do Projeto ALiB, foi obtida uma amostra de 639 ocorrências no cômputo total, das quais 245 ocorrências foram de dados com o zero anafórico, ou seja, 38,4%. Ao passo que foram contabilizadas 312 ocorrências válidas com a utilização dos pronomes *eu* e *mim*, 48,8%, além das 82 ocorrências agrupadas em NO, 12,8%, conforme Tabela 10.

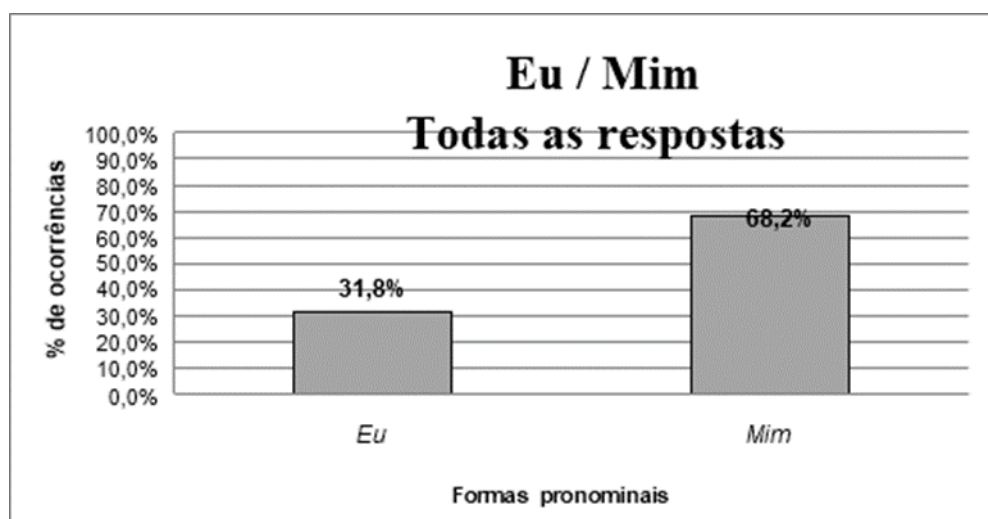
Tabela 10 – Uso pronominal na estrutura *para/prá + eu/mim + infinitivo* – dados das capitais do Projeto ALiB: todas as respostas

Dados coletados	Número de ocorrências	%
Eu/Mim	312	48,8%
0	245	38,4%
NO	82	12,8%
Total:	639	100

Fonte: Elaborado pelo autor (2023).

Dessas 312 ocorrências validadas, uma ocorrência foi da forma *pr'eu*, 15 ocorrências foram da utilização dos dois pronomes na mesma resposta. Além disso, 202 ocorrências, ou seja, 68,2% foram do pronome oblíquo tônico *mim*, em estrutura tal como a analisada neste estudo, ao passo que 94 ocorrências, isto é, 31,8% foram da forma do caso reto, *eu*, como se ilustra no Gráfico 2.

Gráfico 2 – Uso pronominal na *estrutura para/prá + eu/mim + infinitivo* – dados das capitais do Projeto ALiB: frequência geral das respostas



Fonte: Elaborado pelo autor (2023).

Desse modo, conforme o Gráfico 2, é amplamente perceptível que a forma *mim*, no cômputo geral dos dados, obtém um maior percentual de respostas válidas. Assim,

Talvez queira dizer que estamos presenciando uma mudança na língua que ainda não se completou inteiramente. A construção PARA + MIM + INFINITIVO foi passando das variedades [-cultas] em direção às [+cultas]. Já se insinua na fala de muitos falantes cultos, mas ainda encontra resistências para se incorporar definitivamente às variedades [+cultas]. Estamos assistindo, neste caso, uma briga entre as pressões que a norma-padrão exerce sobre as variedades [+cultas] e as pressões que as variedades [-cultas] exercem sobre as [+cultas] (BAGNO, 2006, p. 209).

Os resultados referentes à estrutura sintática analisada permitem afirmar que, os dados coletados pelas equipes do Projeto ALiB, nos anos 2000, denotam que o pronome *mim* em posição de sujeito é predominante na fala dos informantes, oriundos das 25 capitais elencadas para a composição do *corpus* do referido Projeto.

A partir da observação geral desses valores dispostos no Gráfico 2, dados do Projeto ALiB, em contraste com o Gráfico 1, dados do Projeto NURC, é possível entender que, no panorama do português falado no Brasil, nas amostras de falas analisadas, considerando os aspectos peculiares aos referidos Projetos, nas décadas de 70 e 90, o pronome do caso reto, *eu*, é mais proeminente, com valores percentuais de 67%, ao passo que, nos anos 2000, a preferência, por sua vez, passa a ser do pronome do caso oblíquo, *mim*, com valores percentuais de 68,2%.

7.4 OS CONDICIONADORES PARA A REALIZAÇÃO DO FENÔMENO – PROJETO ALiB

Nesta subseção, serão apresentados alguns aspectos, os mais relevantes, para a obtenção dos resultados, ao utilizar o programa estatístico, no que se referem aos dados do Projeto ALiB, com vistas a delinear a configuração do fenômeno nesse *corpus*.

7.4.1 Alguns comentários sobre as rodadas no programa estatístico

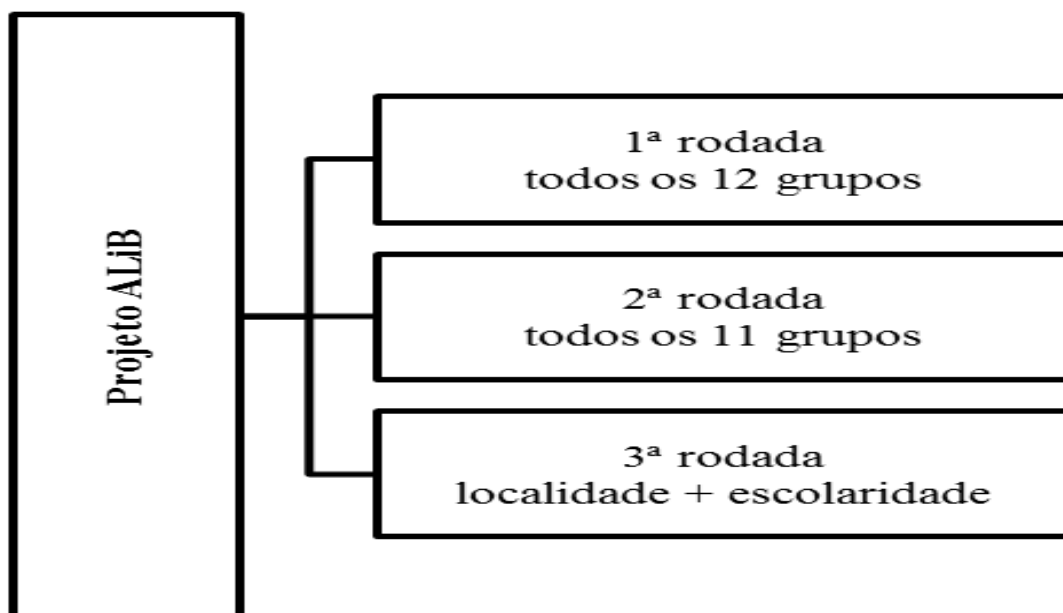
De início, foi realizada uma rodada com todos os dados válidos relativos às elocuições dos 200 inquéritos das 25 capitais brasileiras, dos anos 2000, pertencentes ao banco de dados do Projeto ALiB. As variáveis controladas foram consideradas isoladamente, desse modo, vale lembrar os grupos, conforme seção 6.4:

- (1) Variantes;
- (2) Estilo de registro;
- (3) Natureza semântica do verbo da oração principal;
- (4) Papel semântico do pronome sujeito;
- (5) Função sintática da oração infinitiva;
- (6) Conjugação verbal;
- (7) Forma da preposição;
- (8) Sexo: (Homem e mulher);
- (9) Faixa etária: (faixa I: 18 a 30 anos; faixa II: 50 a 65 anos);
- (10) Escolaridade: (Fundamental incompleto e Universitário);
- (11) Localidade: (25 capitais brasileiras);
- (12) Região Geográfica (Sul, Sudeste, Centro-Oeste, Nordeste e Norte).

Na primeira rodada, dos 12 grupos de fatores analisados, os dois grupos de fatores que apresentaram *nocautes* foram: (i) o papel semântico do pronome sujeito, no que se refere aos papéis de beneficiário e experienciador; e (ii) a função sintática da oração infinitiva, concernente às orações substantivas subjetivas, todos esses foram considerados como estatisticamente irrelevantes na observação da variação em estudo, pois apresentaram unanimidade para a forma *mim*.

Para a segunda rodada, diante dos resultados das rodadas iniciais, a fim de que se entendam os caminhos feitos para a obtenção dos valores percentuais e dos pesos relativos, as rodadas feitas foram esquematizadas na Figura 28.

Figura 28 – Descrição da amostra e das rodadas estatísticas – dados do Projeto ALiB.



Fonte: Elaborado pelo autor (2023).

Salienta-se que, nas duas primeiras rodadas, 1ª e 2ª, os grupos foram considerados isoladamente. E, posteriormente, houve rodada de cruzamento entre escolaridade e diatopia, a 3ª rodada.

7.4.2 As variáveis independentes – extralinguísticas

Nesse subtópico, serão abordados os grupos de fatores extralinguísticos. A ordem da exposição se deu da seguinte maneira: inicialmente, os três grupos, desse conjunto, considerados relevantes para o programa estatístico, a região geográfica, a localidade e a escolaridade. Em seguida, os que não foram selecionados pelo programa, a faixa etária e o sexo.

7.4.2.1 A distribuição diatópica – região geográfica

Com o controle da variável região geográfica, intencionou-se analisar se o fenômeno se encontraria estratificado nas diferentes regiões geográficas do país, sendo possível, dessa maneira, identificar as implicações sociais do fenômeno pelo espaço regional.

Na Tabela 11, a partir dos valores globais, percentuais e pesos relativos obtidos, observa-se que, no panorama das regiões geográficas brasileiras, no que tange à escolha pronominal em contexto de *para/pra + eu/mim + infinitivo*, a região Nordeste destaca-se com peso relativo de 0.53, e a região Sudeste, em seguida, com peso relativo de 0.51, quanto à utilização do *eu*, embora todas as regiões tendam ao ponto da neutralidade.

Tabela 11 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme a região geográfica – dados do Projeto ALiB

NORDESTE		
APLIC./TOTAL	%	P.R.
36/70	51.4%	0.53
SUDESTE		
APLIC./TOTAL	%	P.R.
20/46	43.5%	0.51
SUL		
APLIC./TOTAL	%	P.R.
16/49	32.7%	0.49
NORTE		
APLIC./TOTAL	%	P.R.
27/90	30%	0.48
CENTRO-OESTE		
APLIC./TOTAL	%	P.R.
8/59	13.6%	0.46

Input: 0.294 Significância:
0.000

Fonte: Elaborado pelo autor (2023).

A partir dos pesos relativos demonstrados na Tabela 11, percebe-se a distribuição do fenômeno, nas regiões geográficas, ressaltando que os pesos relativos de todas as regiões são bem próximos do ponto neutro, as tornam sem grandes pesos distintivos quanto aos usos dos pronomes *eu* e *mim*. Dessa maneira, não se pode afirmar que, no aspecto espacial, em especial, referente à região geográfica, tendências quanto ao favorecimento e/ou desfavorecimentos dos pronomes em análise.

Nos *corpora* dos dois Projetos, NURC e ALiB, identificam-se algumas situações semelhantes, quanto ao aspecto das regiões geográficas, haja vista que os dados apresentados

por meio do peso relativo estão, sempre, na faixa próxima ao ponto neutro, fato que se comprova ao cotejar as Tabelas 7 (Projeto NURC) e 11 (Projeto ALiB), respectivamente. Desse modo, a região Sudeste apresenta os pesos relativos 0.52 e 0.51, respectivamente. Depois, ainda no cotejo entre as referidas tabelas, o Sul com pesos relativos 0.39 e 0.49. E, por fim, os dados do Nordeste (Projeto NURC), década de 70, não foram favorecedores quanto ao pronome *eu*, no que tange ao peso relativo, 0.45, ao passo que, nos anos 2000, as ocorrências da forma de prestígio tenderam ao maior uso, com 0.53, embora o peso relativo esteja próximo do ponto neutro.

7.4.2.2 A distribuição diatópica – localidade

Por meio do controle dessa variável, a busca foi pela confirmação de uma das hipóteses principais do estudo: o fenômeno tem um comportamento diferente em termos diatópicos. Nesse sentido, acredita-se que os valores obtidos podem demonstrar que a variação pronominal entre *eu* e *mim*, quando na posição de sujeito, na estrutura aqui analisada, permite separar as capitais em grupos e estabelecer, assim, indícios de possíveis áreas dialetais. Os valores com pesos relativos e os percentuais estão dispostos na Tabela 12.

Tabela 12 – O pronome *eu* em posição de sujeito, nas capitais do Brasil – distribuição conforme a localidade – dados do Projeto ALiB

CAPITAL	APLIC./TOTAL	%	PESO RELATIVO
Salvador	8/9	88.9%	0.94
Fortaleza	5/8	62.5%	0.91
Recife	5/8	62.5%	0.88
Belém	6/7	85.7%	0.87
São Paulo	5/8	62.5%	0.85
Aracaju	5/10	50%	0.83
Natal	4/10	40%	0.72
Belo Horizonte	8/16	50%	0.71
Porto Alegre	7/15	46.7%	0.70
Manaus	8/20	40%	0.69
João Pessoa	3/6	50%	0.68
Teresina	3/6	50%	0.68
Rio de Janeiro	6/12	50%	0.63
Maceió	2/5	40%	0.63
Macapá	4/13	30.8%	0.55
Curitiba	7/21	33.3%	0.52
Goiânia	3/15	20%	0.44
Rio Branco	2/17	11.8%	0.36

São Luís	1/8	12.5%	0.24
Porto Velho	4/11	36.4%	0.18
Boa Vista	3/22	13.6%	0.18
Campo Grande	3/26	11.5%	0.17
Florianópolis	2/13	15.4%	0.16
Cuiabá	2/18	11.1%	0.15
Vitória	1/10	10%	0.11

Input: 0.250 *Significância:*

0.032

Fonte: Elaborado pelo autor (2023).

Diante dos dados expostos na Tabela 12, é possível separar as capitais listadas em três grupos, a fim de melhor observar a distribuição, quando referente ao favorecimento do pronome reto *eu* na estrutura *para/prá + mim + infinitivo*, a saber:

i. Capitais cujos pesos relativos demonstram um favorecimento para a utilização do pronome *eu* (entre 0.94 e 0.55) – Salvador, Fortaleza, Recife, Belém, São Paulo, Aracaju, Natal, Belo Horizonte, Porto Alegre, Manaus, João Pessoa, Teresina, Rio de Janeiro, Maceió e Macapá

ii. Capital cujo peso relativo na faixa de intermediária entre os polos extremos para a utilização do pronome *eu* (0.52) – Curitiba, próxima ao ponto neutro.

iii. Capitais cujos pesos relativos apresentados demonstram um desfavorecimento para a utilização do *eu* (entre 0.44 e 0.11) – Goiânia, Rio Branco, São Luís, Porto Velho, Boa Vista, Campo Grande, Florianópolis, Cuiabá e Vitória, configurando-se, assim, como desfavorecedoras.

Diante dos resultados encontrados, registra-se a proeminência de Salvador (peso relativo de 0.94), em seguida, outras duas localidades situadas no Nordeste brasileiro, Fortaleza e Recife (pesos relativos de 0.91 e 0.88), respectivamente. Ainda nesse sentido, quanto às localidades favorecedoras, notam-se Belém (peso relativo de 0.87), São Paulo (peso relativo de 0.85) e Aracaju (peso relativo de 0.83).

Contrapondo-se às essas localidades situadas como áreas favorecedoras, com pesos relativos baixos para os usos do pronome *eu* em estruturas tal como *para/prá + eu + infinitivo*, encontram-se: todas as três capitais do Centro-Oeste; três capitais do Norte, Rio Branco, Porto Velho e Boa Vista; e capitais únicas das regiões Nordeste, São Luís, Sudeste, Vitória, e, por fim, Sul, Florianópolis.

Vale ressaltar que as cinco capitais pertencentes aos dois *corpora* – Porto Alegre, São Paulo, Rio de Janeiro, Salvador e Recife –, nos dados do Projeto ALiB, estão inclusas no grupo de áreas favorecedoras quanto à realização da forma prestigiada, o que pode denotar

que essas grandes metrópoles brasileiras, com dados dos anos 2000, são resistentes às inovações dos papéis pronominais, fato que se assemelha ao encontrado por Mendes (2016), sobre as cidades maiores serem resistentes às formas pronominais inovadoras.

Os dados em valores percentuais foram responsáveis por espelhar a variação espacial, quanto às formas pronominais em estudo, conforme ilustra a Figura 29.

Figura 29 – Carta experimental M02 – Distribuição dos pronomes *eu* e *mim* quanto à localidade – dados do Projeto ALiB.



Fonte: Elaborado pelo autor (2023).

Considerando os aspectos areais, os dados apontam para possíveis indícios de subáreas dialetais, mesmo que de forma elementar, reconhecendo a insuficiência de dados para traçar áreas dialetais, apenas com dados das capitais. Tal intento se baseia em estudos como o de Cardoso (2016) e o de Mota (2016) que, com dados das capitais do *corpus* do Projeto ALiB, cartas lexicais e cartas fonéticas, respectivamente, indicaram possíveis áreas dialetais, que serão confirmadas e/ou refutados com os dados do interior.

Nessa direção, ao observar a Carta experimental M02, algumas considerações são pertinentes. Notam-se três subáreas de predomínio da forma inovadora, o pronome do caso

oblíquo tônico, *mim* – i. a subárea A, com as capitais do Norte, a exceção de Belém; ii. a subárea B, com todas as capitais da região Centro-Oeste; iii. a subárea C, que engloba todas as capitais da região Sul.

Ainda, nas ilações sobre possíveis traçados dialetais, apresenta-se uma extensa área, de São Luís à São Paulo, regiões Nordeste e Sudeste, iv. a subárea D, uma região de definição complexa, pois há capitais que são de predominância do pronome *eu* – Fortaleza, Recife, Salvador e São Paulo; capitais que estão configuradas como áreas neutras – Teresina, João Pessoa, Aracaju, Belo Horizonte e Rio de Janeiro; e capitais que estão como cidades de proeminência do pronome estigmatizado, *mim* – São Luís, Natal, Maceió e Vitória.

Essas três subáreas, A, B e C, formam um dos polos que divide o Brasil em duas grandes áreas, de um lado esse polo de maior concentração da forma inovadora *mim*, de outro, um polo em que não há definição precisa, pois apresenta capitais que tendem ao inovador e as que tendem ao conservadorismo, subárea D, além daquelas em que não há uma definição. Vale ressaltar que, em todas as capitais, notam-se espaços de disputa entre os pronomes *eu* e *mim*, pois estão em processo de variação.

7.4.2.3 A distribuição por escolaridade

A partir da observação da variável nível de escolaridade, buscava-se averiguar como se comporta a variação entre os pronomes *eu* e *mim* entre os falantes menos e mais escolarizados. Sendo assim, seria possível confirmar ou refutar a outra hipótese – os informantes mais escolarizados tendem à utilização maior da forma do caso reto, *eu*, em detrimento da forma do caso oblíquo tônico, *mim*, em estruturas *para/prá + eu/mim + infinitivo*.

Após as rodadas estatísticas, em que todos os grupos de fatores foram considerados de forma isolada, a variável escolaridade foi uma das categorias significativas selecionadas pelo programa, como pode ser visto na Tabela 13:

Tabela 13 – O pronome *eu* em posição de sujeito, nas capitais do Brasil – distribuição conforme a escolaridade – dados do Projeto ALiB

FUNDAMENTAL			UNIVERSITÁRIO		
APLIC./TOTAL	%	P.R.	APLIC./TOTAL	%	P.R.
25/176	14.2%	0.20	82/138	59.4%	0.85

Input: 0.250 Significância: 0.032

Fonte: Elaborado pelo autor (2023).

Diante desses valores, percebe-se que os informantes pertencentes ao nível fundamental e os informantes do nível universitário apresentam frequências distintas quanto à regra de aplicação do *eu* em posição de sujeito em estruturas *para/prá + eu/mim + infinitivo*. Ao considerar os dados estatísticos, observa-se que os informantes escolarizados utilizam a forma pronominal *eu* na estrutura ora analisada com maior frequência, conforme ilustram o valor percentual (59,4%) e o peso relativo (0.85); ao passo que os informantes com menor tempo de educação formal escolar apresentam porcentagem (14,2%) e peso relativo (0.20) menores, ou seja, desfavorecedores à regra de aplicação.

Desse modo, confirma-se uma das hipóteses principais da tese. Ademais, tal constatação vai ao encontro dos resultados exibidos por Santos e Ribeiro (2020), Santos e Paim (2018) e Figueiredo (2007), uma vez que essa variável foi considerada como relevante para a observação do fenômeno em todos esses estudos.

A partir do que se observa no panorama das capitais brasileiras, é possível afirmar que o pronome *mim* é bastante utilizado na fala dos brasileiros. Nesse sentido, como informa Bagno (2011)

[...] Poderíamos então dizer que **para mim + infinitivo** está a meio caminho de se tornar um traço gradual e de, em breve, não provocar mais a rejeição dos falantes urbanos mais letrados, que se queixam de que ela lhes dói no ouvido. É claro que essa dor só é sentida porque alguém, algum dia, provavelmente em sala de aula, frisou o caráter errado da construção – afinal, esses mesmos falantes letrados usam diversas construções não referendadas pela tradição gramatical sem se queixar de dor alguma... (BAGNO, 2011, p. 729).

O agrupamento das variáveis nível de escolaridade e localidade forneceu valores, percentuais e pesos relativos, que demonstram a situação, de forma particular, de cada capital. O ordenamento das capitais se deu por meio dos maiores pesos relativos, quanto ao favorecimento do pronome *eu*, no grupo de informantes com maior tempo de educação formal. Tais resultados estão dispostos na Tabela 14.

Tabela 14 – O pronome *eu* em posição de sujeito, nas capitais do Brasil – distribuição conforme escolaridade – dados do Projeto ALiB

CAPITAL	FUNDAMENTAL			UNIVERSITÁRIO		
	APLIC./TOTAL	%	P.R	APLIC./TOTAL	%	P.R
Salvador	2/3	67	0.46	6/6	100	0.99
Fortaleza	3/6	50	0.33	2/2	100	0.99
Recife	2/5	40	0.25	3/3	100	0.99
Belém	2/3	7	0.38	4/4	100	0.99
São Paulo	1/4	25	0.27	4/4	100	0.99
Aracaju	2/7	29	0.22	3/3	100	0.99
Maceió	0/3	0	0.01	2/2	100	0.99
João Pessoa	0/3	0	0.01	3/3	100	0.99
Manaus	1/13	8	0.14	7/7	100	0.99
Goiânia	0/10	0	0.01	3/5	60	0.98
Belo Horizonte	1/7	12	0.13	7/8	88	0.90
Macapá	0/8	0	0.01	4/5	80	0.90
Campo Grande	1/13	8	0.80	2/13	15	0.89
Cuiabá	0/8	0	0.01	2/10	20	0.89
Porto Alegre	1/8	12	0.32	6/7	86	0.87
Curitiba	2/12	17	0.24	5/9	56	0.80
Rio de Janeiro	1/5	20	0.25	5/7	71	0.73
Natal	2/7	29	0.20	2/3	67	0.57
Teresina	1/3	33	0.20	2/3	67	0.57
Rio Branco	1/14	7	0.19	1/3	33	0.53
Boa Vista	0/13	0	0.01	3/9	33	0.53
São Luís	0/6	0	0.01	1/2	50	0.42
Porto Velho	1/2	50	0.23	3/9	33	0.31
Florianópolis	1/7	12	0.14	1/6	17	0.22
Vitória	0/5	0	0.01	1/5	20	0.21

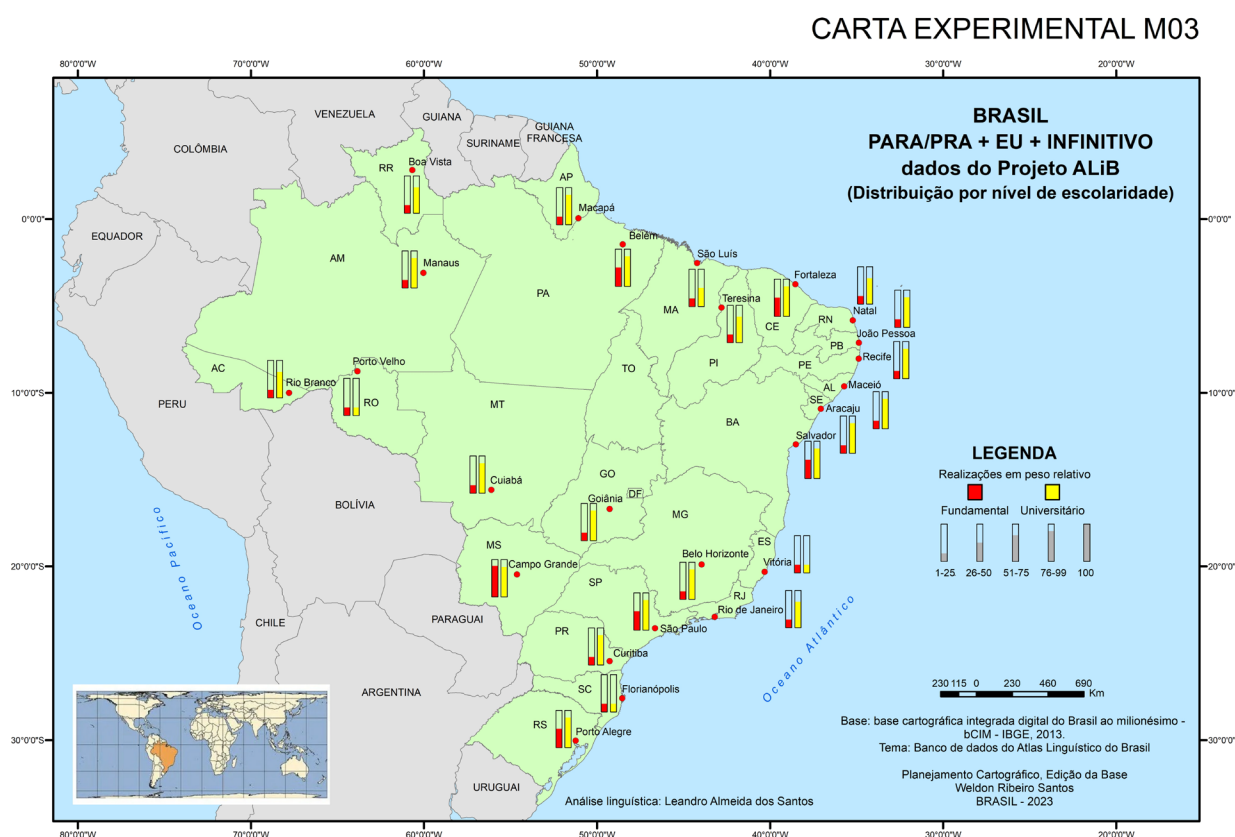
Input: 0.265 *Significância*: 0.034

Fonte: Elaborado pelo autor (2023).

Ressalta-se que, no grupo de capitais em que os informantes com maior tempo de educação formal, ou seja, os informantes universitários, a região Nordeste, em quase sua totalidade, está inclusa, a exceção de São Luís, que, conforme valores percentuais e pesos relativos, dispostos na Tabela 14, pertence ao grupo que desfavorece a utilização da forma do caso reto, *eu*, quer seja entre os informantes com menos ou mais tempo de escolarização. Com peso relativo próximo ao ponto neutro estão Porto Velho, Florianópolis e Vitória. Rio Branco e Boa Vista, respectivamente, com 0.53.

Diatopicamente, de modo a ilustrar os valores em peso relativo dispostos na Tabela 14, a Figura 30 apresenta a situação das 25 capitais em estudo, quanto ao nível de escolaridade.

Figura 30 – Carta experimental M03 – distribuição dos pronomes *eu* quanto à escolaridade – dados do Projeto ALiB.



Fonte: Elaborado pelo autor (2023).

Considerando-se as capitais em que os informantes com nível universitário apresentaram quase de forma categórica o pronome do caso reto, *eu*, estão nove localidades: Salvador; Fortaleza; Recife; Belém; São Paulo; Aracaju; Maceió; João Pessoa e Manaus, com peso relativo de 0.99.

Ainda, nesse sentido, destaca-se a capital Campo Grande, haja vista que, com pesos relativos que demonstram ser favoráveis à utilização do *eu*, tanto na fala de informantes mais e menos escolarizados, tem-se: 0.89 para os informantes com nível universitário e 0.80 para os informantes com nível fundamental. Outro ponto de destaque são capitais que apresentam frequência baixa nos dois níveis de escolaridade, Porto Velho, Florianópolis e Vitória, conforme pesos relativos dispostos na Carta experimental M03.

Ainda, conforme esses valores apresentados na referida carta, podem-se notar reflexos dos aspectos sociohistóricos das localidades. Destaca-se, nesse sentido, àquelas capitais em que o ensino universitário foi implementado no século XIX – Recife, Salvador, Rio de

Janeiro, São Paulo e Porto Alegre, todas com alto índice de utilização da forma de prestígio, a exceção de Belo Horizonte.

Nessa direção, de forma majoritária, as capitais que tiveram implementações do ensino universitário posteriormente, datadas no século XX, o pronome *mim* é predominante. Esses resultados confirmam a hipótese inicial estabelecida, quanto maior tempo de acesso à Educação escolar formal, os informantes tendem à utilização da forma de prestígio, nesse caso, o pronome *eu*.

7.4.2.4 A distribuição por faixa etária

Buscou-se, por meio do controle dessa variável, verificar se a faixa etária exercia influência sobre os usos das duas formas pronominais em estudo. Nesse sentido, podendo demonstrar aspectos sobre a ação do tempo, no que tange ao aspecto diageracional, sobre o fenômeno. Expõem-se os valores encontrados na Tabela 15.

Tabela 15 – O pronome *eu* em posição de sujeito, nas capitais do Brasil – distribuição conforme a faixa etária – dados do Projeto ALiB

FAIXA II (50 – 65 anos)		
APLIC./TOTAL	%	P.R.
51/155	32.9%	0.50
FAIXA I (18 – 30 anos)		
APLIC./TOTAL	%	P.R.
56/159	35.2%	0.49

Input: 0.294 Significância: 0.000

Fonte: Elaborado pelo autor (2023).

Observa-se que, conforme os dados dispostos na Tabela 15, os valores percentuais entre os informantes da faixa I e da faixa II e os valores em peso relativo demonstram que, para a variável faixa etária, os dados apontam para a neutralidade, ou seja, a partir dos pesos relativos 0.50 (faixa II) e 0.49 (faixa I), não se pode afirmar que a faixa etária seja fator de favorecimento quanto ao uso das formas *eu* e *mim*.

Em cotejo com as faixas etárias do Projeto NURC, percebe-se, de modo geral, uma aproximação, pois a forma inovadora (*mim*) possui maior aceitação nas faixas etárias mais novas. Tal constatação também foi apontada pelo estudo de Silva e Nogueira (2018). Ao passo que as escolhas pronominais dos informantes pertencentes às faixas etárias mais velhas, nos *corpora* dos Projetos NURC e ALiB, tendem levemente para a forma *eu*, embora nos dois *corpora*, os pesos relativos apontem para uma direção de neutralidade. Conforme

esses resultados, estaria o PB caminhando no sentido de fortalecimento dos usos da forma inovadora?

7.4.2.5 A distribuição por sexo

A partir da observação da variável sexo, buscava-se averiguar como se dá a estratificação do fenômeno entre os homens e as mulheres.

Após as rodadas estatísticas, em que todos os grupos de fatores foram considerados de forma isolada, a variável sexo não foi considerada como relevante pelo programa. Na Tabela 16, expõem-se os valores obtidos:

Tabela 16 – O pronome *eu* em posição de sujeito, nas capitais do Brasil – distribuição conforme o sexo – dados do Projeto ALiB

MULHER		
APLIC./TOTAL	%	P.R.
59/167	35.3%	0.53
HOMEM		
APLIC./TOTAL	%	P.R.
48/147	32.7%	0.46

Input: 0.294 Significância: 0.000

Fonte: Elaborado pelo autor (2023).

Observa-se que, conforme valores dispostos na Tabela 16, os valores percentuais entre os informantes homens e mulheres são bem próximos, por isso o grupo não foi considerado relevante para o programa estatístico, pois, pelas proximidades dos pesos relativos, há uma proximidade com o ponto de neutralidade. De acordo com Trudgill (1991),

Todos os pesquisadores chegaram à conclusão de que, mesmo levando em conta outras variáveis como idade, a educação e a classe social, as mulheres produzem de modo consistente formas linguísticas mais próximas da linguagem padrão (norma padrão) ou mais prestigiosa que as dos homens, ou então que elas produzem com mais frequência formas desse tipo (TRUDGILL, 1991, p. 78)⁵⁸.

Quando observados os dados encontrados nesta variável com outras pesquisas, notam-se alguns aspectos, a saber: Gomes (2019), observa-se que, na fala dos homens, há predominância do *mim*. Por sua vez, Maia (2016) aponta que as estudantes estão mais

⁵⁸ Todos los investigadores llegaron a la conclusión de que, incluso teniendo en cuenta otras variables como la edad, la educación y la clase social, las mujeres producen consistentemente formas lingüísticas que están más cerca del lenguaje estándar (norma estándar) o más prestigiosas que las de los hombres, o que más a menudo producen tales formas.

propensas tanto ao pronome oblíquo tônico quando ao apagamento – zero anafórico/catafórico, segundo a referida autora.

Vale ressaltar que, historicamente, aos homens e às mulheres foram atribuídos papéis sociais diferenciados, logo, todo o contexto sociohistórico, sociocultural e socioeconômico deve ser levado em consideração, quanto aos resultados que apontem diferenças no modo de utilização da língua.

7.4.3 As variáveis independentes – intralinguísticas

Nesse subtópico, serão abordados os grupos de fatores intralinguísticos. A ordem da exposição se deu da seguinte maneira: primeiramente, o único grupo, desse conjunto, considerado relevante para o programa estatístico, a saber: a forma da preposição. Em seguida, tipo de questionário, a natureza semântica do verbo da oração principal, papel semântico do pronome sujeito, função sintática da oração infinitiva e, por fim, a conjugação verbal.

7.4.3.1 Forma de preposição

Pretendeu-se, a partir do estabelecimento dessa variável, avaliar influência sobre os usos dos pronomes *eu* e *mim*, conforme as formas da preposição, quer seja na forma plena, *para*, ou na forma contracta, *pra*.

Como já notificado no início desta seção, esse grupo de fator foi considerado relevante pelo programa estatístico. Os valores obtidos estão dispostos na Tabela 17.

Tabela 17 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme forma da preposição – dados do Projeto ALiB

PARA		
APLIC./TOTAL	%	P.R.
15/23	65.2%	0.79
PRA		
APLIC./TOTAL	%	P.R.
92/291	31.6%	0.47

Input: 0.250 Significância:0.032

Fonte: Elaborado pelo autor (2023).

Diante desses valores expostos na Tabela 17, percebe-se que, nos contextos em que são iniciados pela forma plena, *para*, a maior probabilidade é da utilização do *eu* em posição

de sujeito, com peso relativo 0.79. Por fim, em contraste, são observados que, nos contextos em que são iniciados pela preposição em sua forma contracta, *pra*, com peso relativo de 0.47, apresenta proximidade com o ponto de neutralidade.

Tais resultados possuem aspectos semelhantes, haja vista que, nos dados do Projeto NURC, a forma plena da preposição é categórica quanto ao uso pronominal de prestígio. Possíveis interpretações podem ser aludidas a fim de ilustrar esses resultados, tais como a forma preposicional plena associa-se em contextos mais formais em que o pronome conservador e canônico é utilizado, ao passo que a forma contracta tende a ser utilizada em situações em que o pronome inovador é também utilizado.

7.4.3.2 Estilo de registro

Procurava-se, por meio da observação dessa variável, averiguar se as diferentes partes do Questionário (2001), utilizado pelas equipes do Projeto ALiB, como instrumento para obtenção dos fenômenos, exerciam influência na escolha pronominal, ou seja, se os informantes alteravam o comportamento linguístico quer sejam nas partes mais monitoradas – QFF, PR, QSL e QMS, na questão específica destinada a obtenção do fenômeno em estudo nessa tese – QMS 023, e, por fim, nas partes consideradas como livres – PM, TDS e LE. Na Tabela 18, encontram-se dispostos os valores obtidos.

Tabela 18 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme estilo de registro – dados do Projeto ALiB

QMS		
APLIC./TOTAL	%	P.R.
76/171	44.4%	0.60
MONITORADO		
APLIC./TOTAL	%	P.R.
20/92	21.7%	0.41
LIVRE		
APLIC./TOTAL	%	P.R.
11/51	21.6%	0.29

Input: 0.293 Significância: 0.001

Fonte: Elaborado pelo autor (2023).

A partir desses valores, nota-se que as partes mais livres do Questionário do Projeto ALiB tendem a desfavorecer o uso do pronome reto *eu*, em posição de sujeito, com peso relativo de 0.29, seguidas das partes monitoradas, com peso relativo de 0.41. Ao passo que,

conforme esperado, na questão QMS 023, alvo de apuração do fenômeno, há indícios fortes de monitoramento por parte dos informantes, haja vista que este fator apresentou um favorecimento à utilização do *eu*, com peso relativo de 0.60.

Tal fato pode ser observado à luz do grau de monitoramento frente à situação estilística da elocução, haja vista que, em muitos momentos, na questão do QMS 023, há informantes que utilizam a forma conservadora, mas, ao longo do inquérito linguístico, alternam os usos entre os pronomes conservador e o inovador. Logo, é possível afirmar que os monitoramentos podem ser determinantes quanto às escolhas linguísticas.

7.4.3.3 Natureza semântica do verbo da oração principal

Buscou-se controlar esse grupo de fatores a fim de observar a relevância do papel semântico exercido pelo verbo que se encontra na oração principal da sentença. Estabeleceram-se como fatores os verbos existenciais; estativos; ação ou mudança de estado; elocutivos; e avaliativos e modais. Os resultados obtidos para esse grupo constam na Tabela 19.

Tabela 19 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme natureza semântica do verbo da oração principal – dados do Projeto ALiB

EXISTENCIAIS		
APLIC./TOTAL	%	P.R.
2/7	28.6%	0.60
AÇÃO/MUDANÇA DE ESTADO		
APLIC./TOTAL	%	P.R.
16/61	26.2%	0.53
ELOCUTIVOS		
APLIC./TOTAL	%	P.R.
79/185	42.7%	0.51
TRANSFERENCIAIS		
APLIC./TOTAL	%	P.R.
4/18	22.2%	0.50
AVALIATIVOS E MODAIS		
APLIC./TOTAL	%	P.R.
5/33	15.2%	0.38
ESTATIVOS		
APLIC./TOTAL	%	P.R.
1/10	10%	0.34

Input: 0.294 Significância:0.000

Fonte: Elaborado pelo autor (2023).

Com base nos valores apresentados na Tabela 19, a situação dos pronomes em estudo, em relação à natureza semântica do verbo da oração principal, demonstrou números e valores percentuais que permitem subdividi-los em três grupos, os que apresentam um comportamento favorável à utilização do pronome reto *eu*, os existenciais, com peso relativo 0.60. Os que apresentam um comportamento que tendem ao ponto neutro: os de ação/mudança de estado, com peso relativo 0.53; os elocutivos, com peso relativo 0.51; e os transferenciais, com peso relativo 0.50. Em contraste, no outro grupo, pertencentes aos desfavorecedores, estão os avaliativos/modais e os estativos, respectivamente, com pesos relativos de 0.38 e 0.34.

Quando comparados aos dados do Projeto NURC, como se nota no item 6.2.3.3, há aspectos semelhantes, como os verbos existenciais pertencerem aos contextos de favorecimento dos usos do pronome *eu*. Os verbos de ação/mudança de estado apresentam divergência, haja vista que nos dados dos anos 70/90, eles estão como desfavorecedores, nos dados dos anos 2000, eles já possuem proximidade com o ponto de neutralidade. E, por fim, os verbos avaliativos/modais e os estativos também são semelhantes, mas no sentido do desfavorecimento nos dois *corpora*.

7.4.3.4 Papel semântico do pronome sujeito

Procurava-se, por meio da observação dessa variável, averiguar se o papel desempenhado pelo pronome sujeito exercia influência na escolha pronominal da oração subordinada.

Durante a primeira rodada estatística, os papéis de beneficiário e experienciador apresentaram *nocauts*, pois as ocorrências catalogadas foram categóricas quanto à utilização do pronome do caso oblíquo, *mim*. Com valores percentuais de 34.1% e peso relativo de 0.49, nota-se que o papel semântico de agente possui proximidade ao ponto neutro, não podendo afirmar, portanto, seu favorecimento quanto à utilização do pronome *eu*, em posição de sujeito.

Se comparados aos resultados encontrados no item 7.2.3.1, a partir dos dados do Projeto NURC, no que concerne ao desfavorecimento do pronome *eu*, o papel semântico beneficiário é observado nos dois *corpora*. Com isso, hipotetiza-se ser o traço beneficiário como favorecedor da forma *mim*, haja vista os contextos de usos e a preposição *para* que antecede o pronome e o próprio verbo. Tais resultados vão na direção contrária aos dados

encontrados por Gomes (2019, p. 87), pois, conforme a autora, “[...] o pronome oblíquo *mim* está perdendo a ideia de beneficiário e passando a expressar agentividade, sendo, inclusive, interpretado exclusivamente como agente em 51% dos dados.”

7.4.3.5 Função sintática da oração infinitiva

A inclusão dessa variável na pesquisa decorreu da necessidade de controlar a relevância da função sintática exercida pela oração infinitiva. Nesse sentido, quatro aspectos, conforme contextos dos usos, foram estabelecidos: as orações substantivas subjetivas; as orações substantivas objetivas; as orações adverbiais finais; as orações completivas nominais.

No entanto, não houve registros de contextos com completivas nominais, além disso, após a primeira rodada, as orações substantivas subjetivas apresentaram apenas uma única ocorrência, tal foi catalogada com o uso do pronome *mim*, conforme resultados dispostos na Tabela 20.

Tabela 20 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme função sintática da oração infinitiva – dados do Projeto ALiB

SUBSTANTIVAS OBJETIVAS		
APLIC./TOTAL	%	P.R.
16/58	27.6%	0.62
ADVERBIAIS FINAIS		
APLIC./TOTAL	%	P.R.
91/256	35.5 %	0.47

Input: 0.294 Significância: 0.000

Fonte: Elaborado pelo autor (2023).

Diante desses valores, vê-se que as orações adverbiais finais, com peso de 0.47, se aproximam do ponto de neutralidade, tendendo ao desfavorecimento. Por fim, em contraste, são observadas as substantivas objetivas, com peso de 0.62, figurando-as como favorecedoras do uso do pronome *eu*.

No cotejo com os dados coletados pelas equipes do Projeto NURC, décadas de 70 e 90, notam-se alguns aspectos: i. as orações subjetivas objetivas são favorecedoras nos dois *corpora*; as orações subjetivas tendem ao favorecimento da forma oblíqua tônica. Por fim, as adverbiais finais apresentam dados díspares, pois, nos *corpora* das décadas 70/90 são contextos favorecedores, no *corpus* dos anos 2000, desfavorecedores.

No cotejo estabelecido por Figueiredo (2007) entre as variantes *mim* e *zero* apontando para que as orações subjetivas e objetivas são favorecedoras do pronome *mim*, e as orações finais e completivas proporcionam o *zero*. Nessa direção, há semelhanças com os resultados expostos por Gomes (2019), pois as orações as orações subjetivas, objetivas e modais tendem ao favorecimento do pronome *mim*, ao passo que as orações finais e completivas nominais tendem ao favorecimento do sujeito nulo.

7.4.3.6 Conjugação Verbal

Com o controle desse grupo de fatores, tencionou-se analisar as possíveis interferências das conjugações verbais. Os resultados obtidos se encontram na Tabela 21.

Tabela 21 – O pronome *mim* em posição de sujeito nas capitais do Brasil – distribuição conforme conjugação verbal – dados do Projeto ALiB

2ª (-ER)		
APLIC./TOTAL	%	P.R.
84/207	40.6%	0.53
1ª (-AR)		
APLIC./TOTAL	%	P.R.
19/87	21.8%	0.50
3ª (-IR)		
APLIC./TOTAL	%	P.R.
2/20	10%	0.17

Input: 0.294 Significância: 0.000

Fonte: Elaborado pelo autor (2023).

Nota-se, por meio dos resultados expostos na Tabela 21, que não há probabilidade de favorecimento quanto à utilização do *eu*, no *corpus* analisado, pois os verbos terminados na segunda e primeira conjugações, com pesos relativos 0.53 e 0.50, estão próximos da neutralidade. E os verbos terminados na terceira conjugação são os que demonstram desfavorecimento, com peso relativo de 0.17.

Quando comparados aos dados do NURC, notam-se similaridades, pois os verbos terminados nas segunda e primeira conjugações são os que obtêm maiores pesos relativos, ao passo que o terminados na terceira conjugação são, sempre, os que apresentam o menor valor.

7.5 OBSERVAÇÕES INICIAIS – AGRUPAMENTO DOS *CORPORA*

Nesta subseção, serão apresentados alguns aspectos para a obtenção dos resultados, ao utilizar o programa estatístico, no que se refere à junção dos *corpora*, com vistas a delinear a configuração do fenômeno nos dados dos dois Projetos e dos fatores intercomparáveis entre eles. Destaca-se que todos os informantes, nesse referido agrupamento, possuem o nível universitário.

A partir dos *corpora* das cinco capitais analisadas, Porto Alegre, São Paulo, Rio de Janeiro, Salvador e Recife, em três tempos diferentes – 70, 90 e 2000 –, com base nos dados do Projeto NURC e do Projeto ALiB, foi obtida uma amostra de 136 ocorrências no cômputo total, das quais 10 ocorrências foram de NO, ou seja, 7,4%. Ao passo que foram contabilizadas 126 ocorrências válidas com a utilização dos pronomes *eu* e *mim*, 92,6%, conforme Tabela 22.

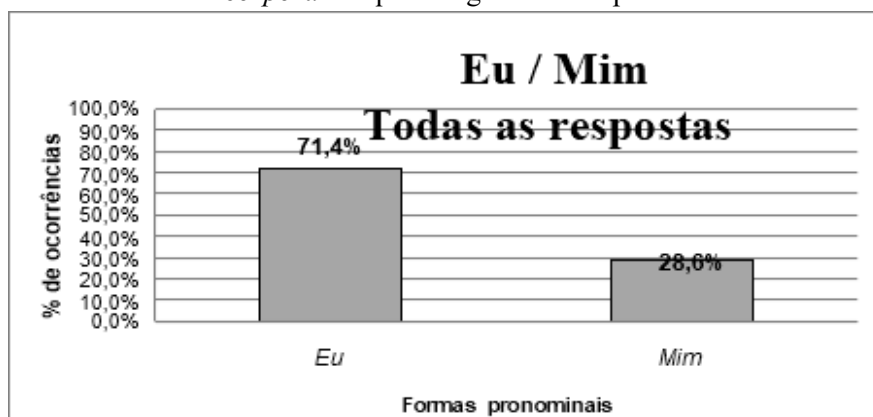
Tabela 22 – Uso pronominal na estrutura *para/prá + eu/mim + infinitivo* – agrupamento dos *corpora*: todas as respostas

Dados coletados	Número de ocorrências	%
Eu/Mim	126	92,6%
NO	10	7,4%
Total:	136	100%

Fonte: Elaborado pelo autor (2023).

Dessas 126 ocorrências válidas, 90 ocorrências, ou seja, 71,4% foram do pronome *eu*, em posição de sujeito, precedido pela preposição *para*, ao passo que 36 ocorrências, isto é, 28,6% foram da forma do pronome *mim*, como se ilustra no Gráfico 3.

Gráfico 3 – Uso pronominal na estrutura *para/prá + eu/mim + infinitivo* – agrupamento dos *corpora*: frequência geral das respostas



Fonte: Elaborado pelo autor (2023).

É notório, em conformidade ao exposto no Gráfico 3, que a forma *eu*, no cômputo geral dos dados, obtém um maior percentual de respostas válidas. Os dados referentes à estrutura sintática analisada permitem afirmar que, nos dados coletados pelas equipes do Projeto NURC, nas décadas de 70 e de 90, e nos dados coletados pelas equipes do Projeto ALiB, nos anos 2000, o pronome *eu* em posição de sujeito é predominante na fala dos informantes com nível universitário, oriundos das cinco capitais coincidentes nos *corpora*, fato que comprova a influência da educação formal nos usos da língua e, nesse caso, nessas escolhas pronominais.

7.5.1 Alguns comentários sobre as rodadas no programa estatístico

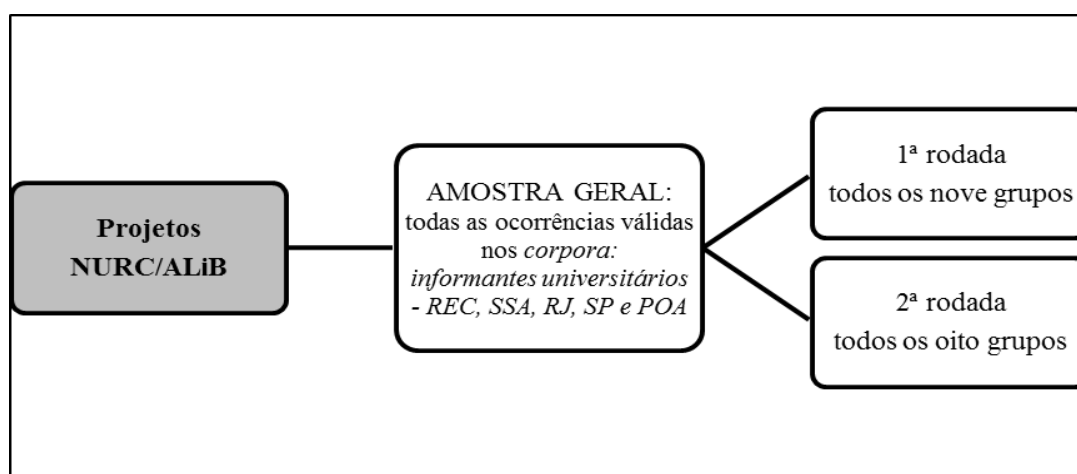
De antemão, foi realizada uma rodada com todos os dados válidos relativos às elocuições dos inquiridos das cinco capitais brasileiras coincidentes nos *corpora*. As variáveis controladas foram consideradas isoladamente, desse modo, vale lembrar os grupos, conforme seção 6.4:

- (1) Variantes;
- (2) Natureza semântica do verbo da oração principal;
- (3) Papel semântico do pronome sujeito;
- (4) Função sintática da oração infinitiva;
- (5) Conjugação verbal;
- (6) Forma da preposição;
- (7) Década: (70/90/2000);
- (8) Localidade: (Porto Alegre; Rio de Janeiro; São Paulo; Salvador e Recife);
- (9) Região Geográfica (Sul; Sudeste e Nordeste).

Na primeira rodada, dos nove grupos de fatores analisados, os grupos que apresentaram *nocautes* foram: (i) natureza semântica do verbo da oração principal; no que tange aos transferenciais; (ii) a forma da preposição, no caso da forma plena *para*; e (iii) a década, no que concerne aos anos 90. Todos esses foram considerados como estatisticamente irrelevantes na observação da variação em estudo, pois apresentaram unanimidade, ou seja, foram categóricos para o uso da forma *eu*.

Para a segunda rodada, um grupo foi descartado, forma da preposição, assim, ficando, apenas, oito grupos, estes foram novamente submetidos ao programa. As rodadas feitas foram esquematizadas na Figura 31, a fim de que se entendam os caminhos feitos para a obtenção dos valores percentuais e dos pesos relativos.

Figura 31 – Descrição das amostras e rodadas estatísticas – dados dos Projetos NURC e ALiB – agrupamento dos *corpora*: informantes universitários.



Fonte: Elaborado pelo autor (2023).

Expõem-se os resultados obtidos, após as rodadas feitas, nas subseções seguintes.

7.5.2 As variáveis independentes – intralinguísticas

Nessa subseção, serão abordados os grupos de fatores intralinguísticos. A ordem da exposição se deu da seguinte maneira: inicialmente, o único grupo considerado relevante para o programa estatístico, a saber: função sintática da oração infinitiva. Posteriormente, os outros grupos, a saber: natureza semântica do verbo da oração principal, papel semântico do pronome sujeito, a conjugação verbal e a forma da preposição.

7.5.2.1 Função sintática da oração infinitiva

Buscou-se controlar a relevância da função sintática exercida pela oração infinitiva. Nesse sentido, após as rodadas, os resultados obtidos seguem expostos na Tabela 23.

Tabela 23 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme função sintática da oração infinitiva – agrupamento dos *corpora*

ABVERBIAIS FINAIS		
APLIC./TOTAL	%	P.R.
64/75	85.3%	0.60
SUBSTANTIVAS OBJETIVAS		
APLIC./TOTAL	%	P.R.
12/15	80%	0.50
SUBSTANTIVAS SUBJETIVAS		
APLIC./TOTAL	%	P.R.
2/5	40%	0.14
COMPLETIVAS NOMINAIS		
APLIC./TOTAL	%	P.R.
1/7	14.3%	0.04

Input: 0.795 *Significância:* 0.186

Fonte: Elaborado pelo autor (2023).

Diante desses valores, vê-se que as orações adverbiais finais, com peso relativo de 0.60, figurando entre as favorecedoras quanto ao pronome *eu*. As substantivas objetivas, com 0.50, exercendo um papel de neutralidade. Ao passo que as orações subjetivas e as completivas nominais são mais susceptíveis ao desfavorecimento da forma pronominal do caso reto, com pesos relativos, respectivos de 0.14 e 0.04.

Em relação aos contextos oracionais apresentados a partir dos dados isolados dos referidos Projetos em cotejo com o agrupamento dos dados dos *corpora*, os contextos, predominantemente, favorecedores ao pronome *eu* são os das orações objetivas. Por sua vez, as orações subjetivas, nessa direção, são desfavorecedoras.

7.5.2.2 Natureza semântica do verbo da oração principal

A fim de observar a relevância do papel semântico exercido pelo verbo que se encontra na oração principal da sentença, buscou-se controlar esse grupo de fatores. Expõem-se os resultados obtidos na Tabela 24.

Tabela 24 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme natureza semântica do verbo da oração principal – agrupamento dos *corpora*

EXISTENCIAIS		
APLIC./TOTAL	%	P.R.
7/9	77.8%	0.75
AÇÃO / MUDANÇA DE ESTADO		
APLIC./TOTAL	%	P.R.
28/36	77.8%	0.49
ELOCUTIVOS		
APLIC./TOTAL	%	P.R.
16/22	72.7%	0.38
AVALIATIVOS E MODAIS		
APLIC./TOTAL	%	P.R.
8/11	72.7%	0.38
ESTATIVOS		
APLIC./TOTAL	%	P.R.
10/13	76.9%	0.34

Input: 0.811 *Significância:*0.000

Fonte: Elaborado pelo autor (2023).

Com base nos valores apresentados na Tabela 24, a situação do fenômeno em estudo, em relação à natureza semântica do verbo da oração principal, demonstrou-se números e valores percentuais bem distintos, permitindo subdividi-los em três grupos, os que apresentam um comportamento favorável à utilização do pronome reto *eu*, os existenciais, com peso relativo 0.75. Em contraste, no outro grupo, estão os ação/mudança de estado, com peso relativo 0.49, próximo ao ponto neutro; e, por fim, os elocutivos, os avaliativos/modais e os estativos, respectivamente, com pesos relativos de 0.38; 0.38 e 0.34, desfavorecedores.

Quando comparadas as três frentes, *corpora* do Projeto NURC, *corpus* do ALiB, conforme resultados espelhados nos itens, respectivamente, 7.2.3.3 e 7.4.3.3, o contexto favorecedor à regra de aplicação são os existenciais, pois os pesos relativos indicam isso.

7.5.2.3 Papel semântico do pronome sujeito

Por meio da observação dessa variável, buscou-se averiguar se o papel desempenhado pelo pronome sujeito exerce influência na escolha pronominal da oração subordinada. Na Tabela 25, encontram-se dispostos os valores obtidos.

Tabela 25 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme papel semântico do pronome sujeito – agrupamento dos *corpora*

AGENTE		
APLIC./TOTAL	%	P.R.
67/82	81.7%	0.56
EXPERIENCIADOR		
APLIC./TOTAL	%	P.R.
5/7	71.4%	0.42
BENEFICIÁRIO		
APLIC./TOTAL	%	P.R.
7/13	53.8%	0.18

Input: 0.803 Significância: 0.000

Fonte: Elaborado pelo autor (2023).

A partir desses valores, nota-se que o papel semântico de agente tende levemente ao favorecimento quanto à utilização do pronome *eu*, em posição de sujeito, com peso relativo de 0.56. Ao passo que os papéis semânticos de experienciador e beneficiário atuam no sentido de desfavorecimento, com pesos relativos de 0.42 e 0.18, respectivamente. No que tange ao papel semântico do pronome sujeito, esses resultados apontam que o papel de beneficiário é desfavorecedor nos dados do NURC, do ALiB e no cotejo desses *corpora*.

7.5.2.4 Conjugação Verbal

Com o controle desse grupo de fatores, buscou-se analisar as possíveis interferências das conjugações verbais nas escolhas pronominais para os contextos em estudo. Os resultados obtidos se encontram na Tabela 26.

Tabela 26 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme conjugação verbal – agrupamento dos *corpora*

2ª (-ER)		
APLIC./TOTAL	%	P.R.
36/45	80%	0.55
1ª (-AR)		
APLIC./TOTAL	%	P.R.
30/41	73.2%	0.45
3ª (-IR)		
APLIC./TOTAL	%	P.R.
13/16	81.2%	0.45

Input: 0.806 Significância: = 0.000

Fonte: Elaborado pelo autor (2023).

Observa-se, por meio dos resultados expostos na Tabela 26, que a probabilidade de leve favorecimento quanto à utilização do *eu*, nos *corpora* analisados, é dos verbos terminados na segunda conjugação, com peso relativo 0.55. E os verbos terminados na terceira e na primeira conjugações são os que demonstram um aparente desfavorecimento, com pesos relativos de 0.45 e 0.45. Nesse sentido, confirma-se que é maior a probabilidade do *eu* ser utilizado com verbos da 2ª conjugação, tal assertiva é comprovada com os resultados encontrados com dados do Projeto NURC, isoladamente.

7.5.2.5 Forma de preposição

A partir do estabelecimento dessa variável, buscou-se avaliar influência sobre os usos dos pronomes *eu* e *mim*, conforme as formas da preposição.

Como já mencionado no início desta seção, esse grupo de fatores não foi considerado relevante pelo programa estatístico, uma vez que, na primeira rodada, as três ocorrências documentadas da forma plena, *para*, são categóricas quanto a utilização do *eu*. No entanto, na forma contracta, notaram-se 79/99 ocorrências da forma *eu*, o que corresponde ao total de 79.8%. Para o estudo de Carvalho e Almeida (2013), a forma mais utilizada é a forma *pra*, o que leva a inferir que a forma contracta vem sendo bastante utilizada antes do pronome *eu*.

7.5.3 As variáveis dependentes – extralinguísticas

Nessa subseção, serão abordados os grupos de fatores de ordem extralinguísticos. A ordem da exposição se deu da seguinte maneira: a década; em seguida, a região geográfica e a localidade. Ressalta-se que nenhum grupo, desse conjunto, foi considerado relevante para o programa estatístico.

7.5.3.1 A distribuição por década

A partir do controle dessa variável, tencionou-se demonstrar se os usos dos pronomes *eu* e *mim*, catalogados em três tempos distintos, em duas décadas, 70 e 90, que são dados do Projeto NURC, e nos anos 2000, que são dados do Projeto ALiB, sofriam interferências relacionadas ao tempo.

Nos anos 90, as 13 ocorrências catalogadas são categóricas quanto à utilização do *eu*. No entanto, nos anos 70 e nos anos 2000, foram catalogados dados referentes à variação dessas formas pronominais, os resultados seguem expostos na Tabela 27.

Tabela 27 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme a década – agrupamento dos *corpora*

2000		
APLIC./TOTAL	%	P.R.
20/26	76.9%	0.48
70		
APLIC./TOTAL	%	P.R.
48/63	76.2%	0.46

Input: 0.811 Significância: 0.000

Fonte: Elaborado pelo autor (2023).

Nota-se, por meio dos resultados expostos na Tabela 27, que não há probabilidade de favorecimento quanto à utilização do *eu*, nos *corpora* analisados, referentes aos dados mais recentes, coletados nos anos 2000, pelas equipes do Projeto ALiB, e os da década de 70, referentes ao do Projeto NURC, pois demonstram, com pesos relativos de 0.48 e 0.46, respectivamente, proximidade ao ponto neutro Logo, conforme esses resultados, afirma-se que há grande probabilidade de utilização das duas formas pronominais, ou seja, elas estão em variação no PB.

7.5.3.2 A distribuição diatópica – região geográfica

Com o controle da variável região geográfica, buscou-se analisar o comportamento do fenômeno nas diferentes regiões geográficas brasileiras. Na Tabela 28, a partir dos valores globais, percentuais e pesos relativos obtidos, observa-se o panorama das regiões geográficas brasileiras, no que tange à escolha pronominal em contexto de *para/pra + eu/mim + infinitivo*.

Tabela 28 – O pronome *eu* em posição de sujeito nas capitais do Brasil – distribuição conforme a região geográfica – agrupamento dos *corpora*

SUDESTE		
APLIC./TOTAL	%	P.R.
55/68	80.9%	0.50
NORDESTE		
APLIC./TOTAL	%	P.R.
17/22	77.3%	0.49
SUL		
APLIC./TOTAL	%	P.R.
7/12	58.3%	0.46

Input: 0.811 *Significância:*

0.000

Fonte: Elaborado pelo autor (2023).

A partir dos pesos relativos expostos na Tabela 28, percebe-se, nas três regiões geográficas, a região Sul, com peso de 0.46; e a região Nordeste, com peso de 0.49, por fim, a região Sudeste, com peso de 0.50, uma neutralidade, não sendo possível classificá-las em regiões geográficas que tendem ao favorecimento ao pronome *eu*.

7.5.3.3 A distribuição diatópica – localidade

Por meio do controle dessa variável, buscou-se avaliar o comportamento do fenômeno em estudo nas cinco capitais brasileiras, coincidentes nos dois *corpora*. Os resultados seguem expostos na Tabela 29.

Tabela 29 – O pronome *eu* em posição de sujeito, nas capitais do Brasil – distribuição conforme a localidade – agrupamento dos *corpora*

CAPITAL	APLIC./TOTAL	%	PESO RELATIVO
São Paulo	8/9	88.9%	0.69
Rio de Janeiro	47/59	79.7%	0.53
Salvador	12/15	80%	0.51
Recife	5/7	71.4%	0.38
Porto Alegre	7/12	58.3%	0.25

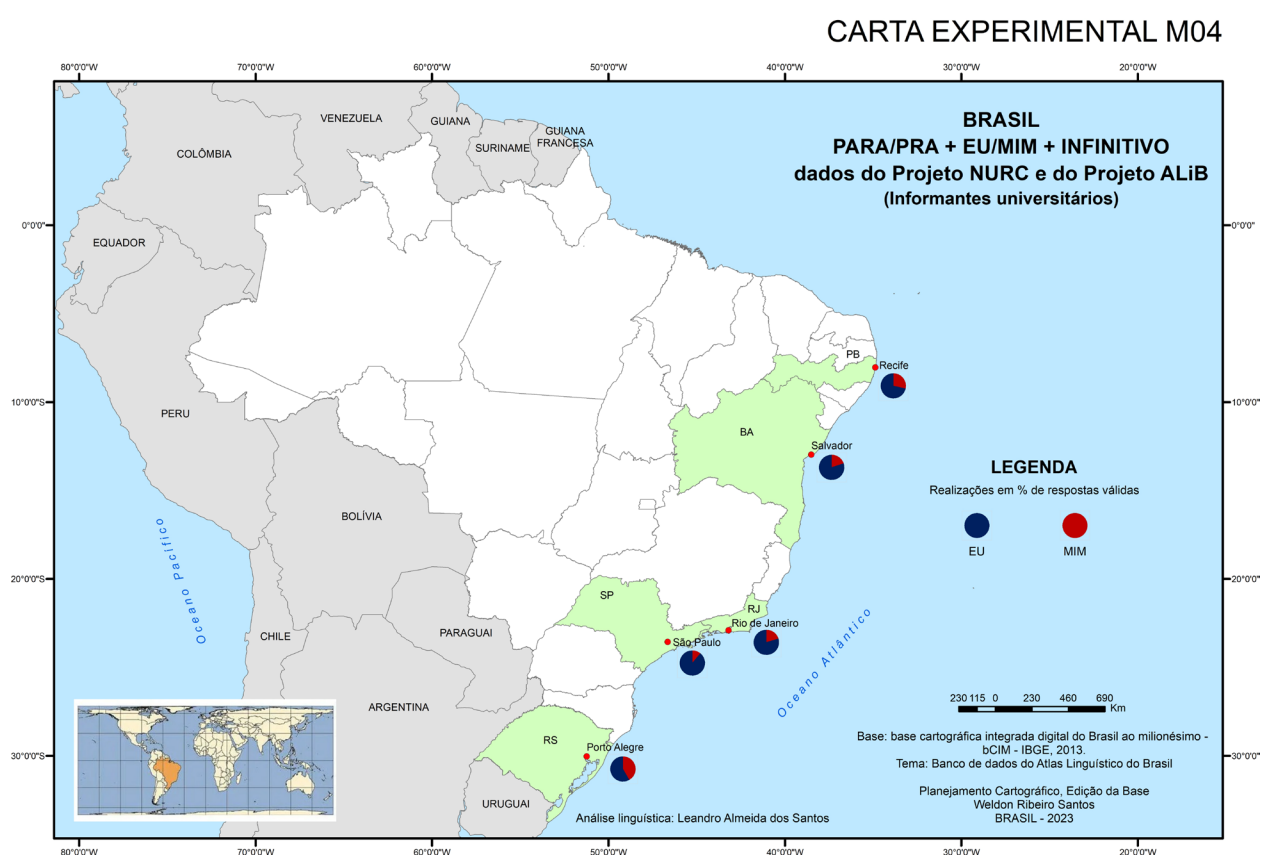
Input: 0.811 *Significância:* 0.000

Fonte: Elaborado pelo autor (2023).

As capitais Recife e Porto Alegre, nas amostras analisadas, são desfavorecedoras ao uso do pronome oblíquo reto *eu*. Em contraste, a capital sudestina, São Paulo, com peso

relativo de 0.69, apresenta maior inclinação na escolha desse pronome. Por fim, Rio de Janeiro e Salvador, quanto à escolha pronominal, estão propensas à uma leve tendência a escolha do *eu* na estrutura em análise, com pesos relativos de 0.53 e 0.51, respectivamente, embora estejam bem próximas do ponto de neutralidade. Nessa direção, a Carta experimental M04 demonstra a variação espacial no agrupamento dos dados desses *corpora*.

Figura 32 – Carta experimental M04 – Distribuição dos pronomes *eu* e *mim* quanto à localidade – dados do Projeto NURC e do Projeto ALiB – informantes universitários



Fonte: Elaborado pelo autor (2023).

Diante dos valores percentuais apresentados na Figura 32⁵⁹, em contraste com as configurações espaciais dessas cinco capitais, isoladamente, nos *corpora* do Projeto NURC e do Projeto ALiB, alguns casos específicos são notáveis, por exemplo:

i. nos anos 70, a capital Porto Alegre apresenta inclinação para usos da forma *mim*, fato que se mantém nos anos 2000, no agrupamento dos *corpora*, com dados exclusivos dos universitários, a forma *eu* figura como levemente majoritária;

ii. Rio de Janeiro, também, nesse sentido, apresenta dados divergentes, pois, nos *corpora* das décadas de 70/90 e no agrupamento dos *corpora* dos Projetos, apenas com informantes universitários, há predominância da forma de prestígio, *eu*. Nos dados do Projeto ALiB, o percentual indica equilíbrio das duas formas pronominais em estudo;

iii. Recife, Salvador e São Paulo são áreas de utilização do pronome de prestígio linguístico, o pronome do caso reto *eu*, quando se notam os valores percentuais. Um fato a observar é que, das cinco capitais coincidentes nos *corpora*, elas três são as mais antigas, quando se observa a data de fundação.

7.6 METALINGUAGEM NOS/DOS DADOS

Nessa subseção, serão apresentados alguns aspectos sobre os dados coletados e que, de certa maneira, podem servir como mote para olhar os dados pronominais expostos em toda a seção 7, bem como fornecer caminhos para o tratamento interpretativo e qualitativo de pesquisas dialetológicas, na dimensão diarreferencial, com vistas a ressaltar as percepções sobre a metalinguagem que envolvem as situações e os contextos das entrevistas sobre a língua falada e, também, sobre as não obtenções dos dados.

7.6.1 As percepções metalinguísticas

Em muitos contextos, a pressão social sobre a língua e sobre determinados usos são marcados por monitoramentos, *tabus* e, sobretudo, estigmas, muitas vezes, fomentados no seio e no imaginário social por instituições, por exemplo, a família e a escola, haja vista que

O falante está em constante processo de avaliação da língua, seja de forma consciente ou não. No processo da variação, embora tenham o mesmo valor de verdade ou representacional, às formas linguísticas diferentes podem ser atribuídas avaliações ou valorações sociais igualmente diferentes, o que se dá por conta das pressões sociais que operam constantemente sobre a língua, “não de algum ponto remoto no passado, mas como uma força social imanente agindo no presente vivo” (LABOV, 2008, p. 21) (FREITAG et al, 2016, p. 139).

Como exposto na seção introdutória desta tese, o fenômeno ora explorado, as escolhas pronominais entre o *eu* e o *mim* para ocupar a posição de sujeito, precedidos pela preposição

para, em estruturas tais como *para/prá + eu/mim + infinitivo*, são alvos de atenção por parte da sociedade.

Assim, a utilização do pronome oblíquo tônico *mim* em posição de sujeito, na estrutura tal como a anteriormente mencionada, em muitas comunidades de fala, torna-se motivo para falas preconceituosas e ofensivas, no que tange ao lugar de quem a pronunciou. Muitas vezes, associações pejorativas são feitas com a língua indígena e, também, há uma categorização equivocada, colocando-a no lugar do “feio”, do “inculto” e do “erro”. No entanto, conforme se observa no *corpus* do Projeto ALiB

Os dados referentes à estrutura sintática analisada permitem afirmar que, contrariando o imaginário popular, o qual difunde que *mim* em posição de sujeito é predominante na fala rural, indígena e/ou de baixa escolaridade, a maioria dos brasileiros, informantes das capitais, opta pela forma *mim* (SANTOS; RIBEIRO, 2020, p. 181).

Ao visar a íntima ligação entre língua, cultura e sociedade, haja vista não haver língua dissociada do social, alguns recortes de contextos de falas no *corpus* do Projeto ALiB⁶⁰ foram retirados, a fim de que possam servir de subsídios para o tratamento qualitativo/interpretativo dos dados, no que concerne à metalinguagem envolvida nos inquéritos analisados, avaliando, dessa maneira, a variação diarreferencial, conforme postulações de Thun (2000), as crenças e as atitudes linguísticas também. Seguem expostos, a seguir, os exemplos ilustrativos.

(29) INQ.- Vamos dizer que alguém peça para você fazer uma tarefa. Só que a outra tarefa, só que a outra pessoa acha que a tarefa era para ela.

INF.- Huhum.

INQ.- Ai você diz: “Bom essa tarefa é na verdade para...”

INF.- Para outra pessoa (...)

INQ.- Não é você que tem que fazer quem tem que fazer é você.

INF.- Mas na verdade é para ela é *para mim fazer*? Então?

INQ.-É

INF.- *Para eu fazer* (risos).

INQ.- Por que você consertou? O que você falou?

INF.- É que *mim* não pode ficar no meio da frase (risos) parece índio falando. (risos).

INQ.- Tá bom.

(Inf. 06 – 220: Curitiba, mulher, faixa I, universitário – QMS 023).

⁶⁰ Nos *corpora* do Projeto NURC não foram encontradas marcas linguísticas para tais análises.

(30) INQ.- Alguém pede pra você é... fechar uma porta, aí você não fecha a porta, aí a pessoa vai lá pra fechar, você fala não deixa essa porta é pra... quem fechar?

INF.- Pra *eu fechar*.

INQ.- Humhum.

INF.- Mim Andréia, mim fazê (risos). Ai, me dá uma agonia quando alguém fala assim.

INQ.- (risos). Todo mundo fala assim?

INF.- Costumam falá *pra mim fechá*, mais eu acho feio.

INQ.- É...

INF.- Parece índio, mim Andréia, mim conversa.

INQ.- Na escola eles batem muito nisso né, pro vestibular né?

INF.- É.

(Inf. 06 – 006: Manaus, mulher, faixa I, universitário – QMS 023).

Observa-se, nos dois contextos situacionais das entrevistas, a manifestação de um pensamento fomentado no imaginário coletivo, o de que o fenômeno ora estudado na tese tem origem na fala indígena, tais falas podem ser identificadas como preconceito linguístico étnico-racial, em referência aos indígenas. Por vezes, os usuários dessa forma pronominal, nos mais diversos contextos de usos da língua, são alvo de correções e zombarias. Sobre essa questão, Bagno (2011) comenta:

Não vamos apelar para argumentos preconceituosos que atribuem a sintaxe **para mim + infinitivo** à ignorância ou ao descaso dos brasileiros pela língua – não vamos, portanto, afirmar, como fez certa ocasião um jornalista, que “só índio fala pra mim fazer”. Os índios brasileiros que falam português aprenderam a língua na forma das variedades rurais mais estigmatizadas, de modo que eles não têm “culpa” nenhuma na ocorrência dessa sintaxe – como também não cabe responsabilizar os falantes rurais e rurbanos pelo surgimento dela. Muito pelo contrário, o processo de surgimento dessa sintaxe revela o alto poder de processamento da língua pelo falante comum (BAGNO, 2011, p. 729).

Nos exemplos seguintes, (31) e (32), notam-se outros contextos de falas e outra forma de categorização do fenômeno, o lugar do “erro”, por sua vez, categorizados como preconceito linguístico auto-metalinguístico.

(31) INF. – *Pra mim escrevê. É pra eu...*

CIR. – Ele não fala *pra mim*, ele sempre fala *pra eu*, ele corrige quem fala *mim*.

Sempre observa, ele sempre fala *pra eu fazê* e ainda critica quem fala *mim*.

INQ. – Mais o normal que as pessoas falam aí...

INF. – É *pra mim*, mais a gente fala *pra eu*, o correto é *pra eu escrevê*, mas você está sempre, o dia-
adia...

(Inf. 07 – 190: Vitória, homem, faixa II, universitário – QMS 023).

(32) INF. – Ah, porque assim, eu falo é *pra mim fazer*, mas eu sei que tá errado.

INQ. – Não mais como você fala?

INF. – É *pra mim*.

INQ. – Pra mim fazer?

INF. – Ahan.

(Inf. 08 – 021: Porto Velho, mulher, faixa II, universitário – QMS 023).

Observam-se, no conjunto dos exemplos que ilustram os estigmas que cercam o fenômeno, aspectos que podem aludir às reflexões sobre a língua, tais como: i. as falas recortadas são de informantes universitários, o que denota a influência dos meios formais de Educação, enquanto agentes que moldam os usos linguísticos e as escolhas pronominais dos informantes. Esses resultados coadunam com os encontrados na pesquisa de Maia (2016), a partir da percepção de estudantes universitários sobre o fenômeno; ii. a propensão das mulheres à correção e ao monitoramento sobre os usos linguísticos, haja vista a maioria dos exemplos citados (29), (30) e (32).

Nos exemplos seguintes, (33) e (34), notam-se os usos das duas formas pronominais, no total, foram 15 ocorrências, uma vez que “[...] quando indagadas sobre quais dentre várias formas são características de sua própria fala, as respostas das pessoas refletem a forma que elas acreditam gozar de prestígio ou ser a ‘correta’, mais do que a forma que elas realmente empregam” (LABOV, 2008 [1972], p. 248). Algumas hipóteses podem ser levantadas, sobretudo, aquelas que sugerem uma hipercorreção e/ou um monitoramento linguístico, devido ao fenômeno receber uma avaliação social bem marcada e de forma pejorativa.

(33) INQ.- Eh, Fabiano, por exemplo, alguém pede para você fazer uma tarefa, um trabalho. Mas outra pessoa acha que a tarefa é para ela. Então você diz: essa tarefa, na verdade, é para... INF.- *Mim*.

INQ.- Fazer.

INF.- Fazê. *Pra mim fazê, pra eu fazê*.

INQ.- Pra eu fazer, né?

(Inf. 01 – 041: Fortaleza, homem, faixa I, fundamental – QMS 023).

(34) INQ.- Alguém pede para você fazer uma tarefa, um trabalho. Mas outra pessoa acha que a tarefa é para ela. Então você diz: essa tarefa, na verdade, é para... INF. – *Pra eu fazê, pra mim fazê*.

(Inf. 03 – 053: Natal, homem, faixa II, fundamental – QMS 023).

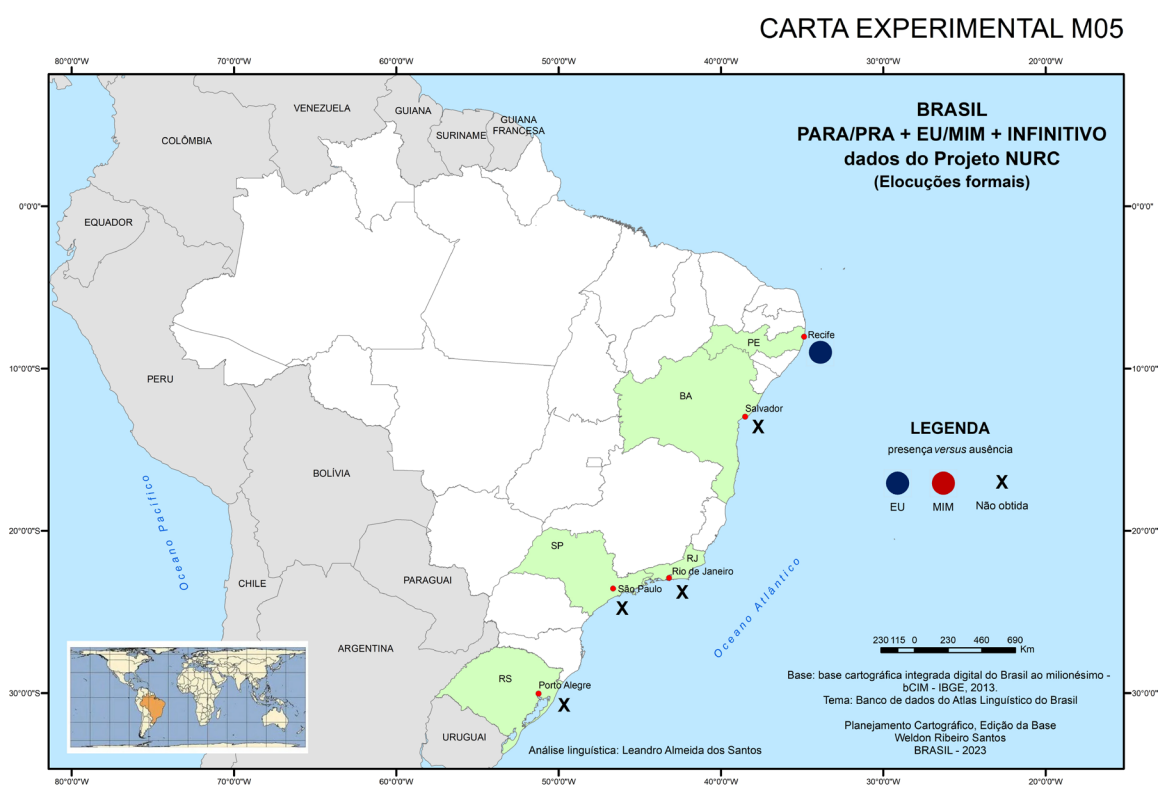
Nos dois últimos referidos exemplos, observam-se as duas formas pronominais sendo utilizadas, o que pode indicar, conforme os dados evidenciam, uma variação estrita entre esses pronomes. Tais exemplos ilustram os dados encontrados por Santos e Paim (2018), quanto apontam que 15,60% dos informantes realizaram as duas formas pronominais, *eu/mim*. Ademais, esses dados trazem à tona os celeumas difundidos no imaginário coletivo sobre “erro”, no que tange aos usos linguísticos. Essa última hipótese, por exemplo, pode exercer influência sobre a não obtenção das respostas, em especial, as Não Obtidas – NO –, casos que serão expostos pormenorizados no item 7.6.2.

7.6.2 As respostas Não Obtidas

Aqui, serão tratados os dados que, por algum motivo, foram agrupados em respostas Não Obtidas – NO. Como exposto na subseção 2.4, o não dado sugere caminhos diversos para interpretação, dessa maneira, separá-los e categorizá-los são ações necessárias para as pesquisas sobre a língua falada. Para além dessas, tentativas de interpretação, bem como, quando possível, cartografiação são salutares, com vistas ao maior aproveitamento na exegese dos dados.

Para os dados do Projeto NURC, nas gravações feitas em situações formais de usos da língua, nas Elocuções Formais (EF), espera-se um maior monitoramento e, conseqüentemente, alguns fenômenos em variação são refreados. Nesse sentido, a Figura 33, ilustra a representação da única forma catalogada nesse tipo de entrevista e, nas outras localidades, o não dado.

Figura 33 – Carta experimental M05 – Distribuição dos pronomes *eu* e *mim* quanto às Elocuções Formais – dados do Projeto NURC.



Fonte: Elaborado pelo autor (2023).

Na Carta experimental M05, a partir dos dados coletados pelas equipes do Projeto NURC, foi possível perceber que, nas Elocuções formais, somente em Recife foi registrada uma única ocorrência, conforme exemplo (5a). Nas outras capitais, de modo a facilitar a leitura da carta pelos consulentes, seguiram-se as orientações disponíveis em Santos (2016/2020) para a cartografia do não dado. Assim, contrastando diferentes formas de cartografar na história da geolinguística brasileira, conforme apontado na subseção 2.4.

Para os dados do Projeto ALiB, na Tabela 30, encontram-se a sistematização do agrupamento sobre o dado NO, a fim de fornecer caminhos para facilitar a leitura dos dados, seguem as siglas: SE – as respostas fornecidas com os pronomes em estudo, mas sem a estrutura, apenas como forma de preenchimento da formulação elaborada pelo inquiridor; NO⁶¹ – as respostas que, mesmo após diversas tentativas, reformulações, não foram alcançadas durante a realização dos inquéritos; e ADV – advérbios inseridos na estrutura em análise – *para/prá + eu/mim + infinitivo* – mais precisamente entre o pronome e o verbo.

⁶¹ Pelo que foi observado, não foram catalogados dados categorizados como NS (Não Sabe) e NL (Não Lembra).

Tabela 30 – Uso pronominal na estrutura *para/prá + eu/mim + infinitivo* – dados das capitais do Projeto ALiB: respostas não obtidas

Dados coletados	Número de ocorrências	%
SE	58	70,7 %
NO	15	18,3 %
ADV	09	11%
Total:	82	100%

Fonte: Elaborado pelo autor (2023).

Assim, apresentam-se inferências, ainda sobre o não dado, desta feita a esquematização das respostas apresentadas, mas que só funcionaram como uma espécie de preenchimento do questionamento feito no QMS 023 – exemplos (35) e (36), ou somente com um dos pronomes sem o verbo – exemplo (37), ou seja, respostas que não foram consideradas como válidas no cômputo geral, pois não seguiram a estrutura analisada – *para/prá + eu/mim + infinitivo*.

(35) INF: – Para *mim*.

(Inf. 01 – 034: Teresina, homem, faixa I, fundamental – QMS 023).

(36) INQ: - Alguém pede pra você fazer uma tarefa, mas outra pessoa acha que a tarefa é pra ela. Então você diz: “Não, olha essa tarefa na verdade é para... fazer... é para...” INF.- *Mim*.

INQ.- Fazer.

INF.- *Mim fazê*.

(Inf. 01 – 021: Porto Velho, homem, faixa I, fundamental – QMS 023).

(37) INQ: - Alguém pede pra o senhor varrer aqui a calçada, aí o senhor demora um pouquinho vem a sua irmã que vai varrer, você diz assim: “Não, não essa calçada é pra...” quem varrer? “É pra...”

INF.- *Mim, né?*

(Inf. 03 – 002: Macapá, homem, faixa II, fundamental – QMS 023).

Outra parte dos dados, 15 ocorrências, foram de respostas NO, não obtidas. Os informantes não deram respostas para a QMS 023, questão destinada à apuração da variação entre os pronomes em estudo nesta tese, também nas outras partes desses inquéritos não foram encontradas estruturas com a variação buscada.

Ao observar os dados pelo prisma diatópico, 12 dessas são oriundas da região Nordeste, duas do Norte e uma do Sul. Ao observar pelo prisma social, 10 ocorrências foram catalogadas na fala dos informantes homens; no que tange à escolaridade, os informantes com o nível universitário são responsáveis por 11 ocorrências; e, por fim, no que concerne à

faixa etária, os informantes com a faixa II apresentaram, levemente, um uso maior, 8 ocorrências.

Por fim, nos contextos em que foram inseridos advérbios – *não, num e também* –, na estrutura em análise, – *para/pra + eu/mim + não, num e também + infinitivo* – foram contabilizadas nove ocorrências, conforme sinalizado na Tabela 30. Nesse sentido, perfazendo um total de 11%. O exemplo (38) ilustra uma dessas ocorrências.

(38) INQ.– E de outros lugares assim do Brasil mais distante? Falam diferente?

INF. – Não... só em Campos, só longe mehmó.

INQ.– Aqui dentro do estado mesmo?

INF. – Do estado não... é, é dentro do estado porque o resto...

INQ.– Como é que você viajou tanto? Você ainda não me contou essa história.

INF. – Tem que contá a história que eu viajei também? (...) Aí pra minha mãe me mandá saí fora e (eles reclamá), *pra mim num morrê*, eu tive que viajá (inint) é, já que ia pagá né?

(...) (Inf. 01 – 202: Rio de Janeiro, homem, faixa I, fundamental – PM 04).

Para os exemplos ilustrativos (35), (36), (37), as reflexões são direcionadas ao instrumento de pesquisa, o Questionário ALiB (2001), bem como para à formulação da QMS 023, pois, da maneira como o questionamento é aplicado, alguns informantes entendem que basta um simples preenchimento da lacuna deixada na situação hipotética para que a pergunta seja respondida. Seguem os dois contextos que ilustram a formulação do QMS 023:

1. Alguém pede para você / o (a) senhor (a) fazer uma tarefa. Mas outra pessoa acha que a tarefa era para ela. Então, você / o (a) senhor (a) diz: Essa tarefa, na verdade, é para ____ fazer.
2. [Vamos supor que você / o (a) senhor (a) precisa distribuir tarefas por três pessoas. Por exemplo: a limpeza é para alguém fazer, a roupa é para outra pessoa lavar e a terceira tarefa (...) é para você / o (a) senhor (a). Como você / o (a) senhor (a) diria? (COMITÊ NACIONAL... 2001, p. 40) [grifos nossos].

Nesse sentido, sugere-se que, para pesquisas futuras, no que tange à busca da variação entre os pronomes *eu* e *mim* em posição de sujeito, outras formas de obtenção sejam testadas e formuladas, a fim de que toda estrutura em análise seja evidenciada nas elocuições. Para os contextos hodiernos, há necessidade de questionários que sejam mais dinâmicos e menos densos. Logo, muitos instrumentos utilizados anteriormente, hoje, carecem de reformulações, haja vista as modificações pelas quais a sociedade vem passando, sobretudo, pela dinamicidade que envolve os contextos contemporâneos.

7.7 SÍNTESE DA SEÇÃO

Esta seção, 7, tem por meta expor a descrição e a análise dos dados da pesquisa, com vistas a ofertar detalhes e pormenores de como o fenômeno estudado está configurado em dois *corpora* distintos de língua falada, coletado em três tempos décadas de 70 e 90, pelas equipes do Projeto NURC, e nos anos 2000, pelas equipes do Projeto ALiB.

Nas duas subseções iniciais, 7.1 e 7.2, abordam-se os aspectos referentes aos dados do Projeto NURC. A partir das análises descritiva e quantitativa, aponta-se, nesses *corpora*, há uma relativa variação entre os pronomes *eu* e *mim*, em posição de sujeito, precedidos pela preposição *para*. Os valores percentuais demonstram que, nesses *corpora*, o pronome pessoal do caso reto – *eu* – é mais proeminente, com 67%. Após as análises estatísticas, conforme o *Goldvarb X*, alguns fatores foram elencados como favorecedores à utilização da forma prestigiada, o pronome *eu*, a saber: a função sintática da oração infinitiva.

Em seguida, nas subseções 7.3 e 7.4, explicitam-se os procedimentos adotados concernentes aos dados do Projeto ALiB. A partir das análises descritiva e quantitativa, aponta-se, nesse *corpus*, também, uma relativa variação entre os pronomes *eu* e *mim*, em posição de sujeito, precedidos pela preposição *para*. Todavia, o pronome oblíquo tônico é mais proeminente, com 68,2%. Nesse sentido, após análises estatísticas, alguns fatores foram elencados como favorecedores à utilização da forma *eu*, a saber: forma da preposição; escolaridade; localidade; e, por fim, região geográfica.

Na subseção 7.5, a partir do agrupamento estabelecido com os dois *corpora*, são expostos os procedimentos quanto ao levantamento e tratamento dos dados, bem os resultados estatísticos obtidos, por meio do *Goldvarb X*, além dos valores percentuais. Dessa maneira, observa-se que, ao comparar os dados do Projeto NURC com os dados do Projeto ALiB, que a função sintática da oração infinitiva atua com favorecimento da forma pronominal *eu*.

Após essas exposições, na subseção 7.6, demonstra-se uma análise qualitativa dos dados catalogados, no que tange às percepções metalinguísticas que, incluem-se, também, a análise das respostas Não Obtidas, NO, buscando caminhos interpretativos para tais fenômenos que, de certo modo, revelam aspectos que moldam as escolhas pronominais dos informantes das capitais brasileiras analisadas nos *corpora*.

Com vista a sistematizar as discussões desta seção, no tocante aos resultados obtidos, o Quadro 31 é responsável por sumarizar os fatores favorecedores nos *corpora* analisados.

Quadro 31 – Resumo dos resultados quanto ao pronome *eu* – dados dos *corpora* analisados.

RESULTADOS PARA O PRONOME DO CASO RETO <i>EU</i>				
VARIÁVEIS		Aspectos favorecedores		
		Dados do Projeto NURC	Dados do Projeto ALiB	Nos dois corpora
DIATÓPICAS	Região geográfica	-	-	-
	Localidade	São Paulo e Rio de Janeiro.	Salvador, Fortaleza, Recife, Belém, São Paulo, Aracaju, Natal, Belo Horizonte, Porto Alegre, Manaus, João Pessoa Teresina, Rio de Janeiro, Maceió e Macapá.	São Paulo.
LINGUÍSTICAS	Natureza semântica do verbo da oração principal	Existenciais, Transferenciais e Elocutivos.	Existenciais.	Existenciais.
	Papel semântico do pronome sujeito	Experienciador.	-	Agente.
	Função sintática da oração infinitiva	Substantivas objetivas e as Adverbiais finais.	Substantivas objetivas.	Adverbiais finais.
	Conjugação verbal	2 ^a -er	-	2 ^a -er
	Forma da preposição	Pra - Para foi categórico.	Para	Pra - Para foi categórico.
SOCIAIS	Sexo	-	-	-
	Faixa etária	Faixa II.	-	-
	Escolaridade	-	Universitário.	-
	Década	70 - 90 foi categórico	-	-
ESTILÍSTICAS	Estilo de registro	D2	QMS	-

Fonte: Elaborado pelo autor (2023).

Portanto, é notório que, conforme exposto nesta seção, os resultados apontam para a interferência de diversos fatores, quanto ao favorecimento do pronome *eu* na estrutura ora analisada, quer sejam diatópicos, linguísticos, sociais e estilísticos. Nessa direção, mais uma vez, frisa-se a relevância dos pressupostos teóricos que consideram a língua em seus contextos de usos, sendo passível de variação e mudança, mediante a fatores internos e externos.

8 ENCRUZILHADAS FINAIS

O que aqui dissemos, numa abordagem que deve ser tida como inicial, longe fica de poder ser considerado conclusivo, muito menos definitivo...

(ROLLEMBERG, 1994, p. 109).

Nesta tese, a observação empírica se voltou para os usos dos pronomes *eu* e *mim* na posição de sujeito, precedidos pela preposição *para*, em duas décadas 70 e 90 – dados do Projeto NURC –, e nos anos 2000 – dados do Projeto ALiB –, oriundos das capitais brasileiras, conforme critérios metodológicos desses dois grandes empreendimentos de pesquisa sobre a língua falada no Brasil. Nesta seção, aporta-se em encruzilhadas que espelham as trajetórias feitas pelo fenômeno, as trilhas percorridas na pesquisa empreendida, como também as possibilidades para os caminhos vindouros.

É prudente afirmar que as conclusões esboçadas nesta tese não são definitivas, visto que, em se tratando de língua falada e dos constantes rearranjos do sistema pronominal do PB, em especial, dos pronomes pessoais, seria inconsequente trazer afirmações categóricas e fixas. Ressalta-se que, desse modo, o estudo sobre a realidade dos pronomes *eu* e *mim*, mapeados na presente pesquisa, está firmado em bases científicas sólidas e confiáveis, embora possua um cunho provisório, como todo o fazer da ciência.

Estudar a língua, tomando-a como um fenômeno dinâmico, mutável, diversificado, sujeito a variação e a mudança linguísticas, é uma tarefa, ao mesmo tempo, árdua e prazerosa, pois inúmeros fenômenos, categorizados em um *locus* de estigma, quando sujeitos à análise pormenorizada e ao rigor científico, revelam, de fato, como estão estruturados. Além disso, as raízes pelas quais recebem tal avaliação social vêm à tona. Para essa compreensão, podem ser considerados os aportes teóricos e metodológicos da Dialectologia e da Sociolinguística Variacionista, vide seção 2, utilizados como aportes teóricos fundamentais nesta tese.

Nesse sentido, no que tange ao Português Brasileiro – PB –, nota-se que, por razões históricas, sociais e culturais, oriundas desde o Latim, perpassando pelos primeiros contatos do colonizador europeu com os povos autóctones até aos acontecimentos mais atuais, há uma variedade de português que apresenta características plurais. Por seu turno, essas retratam a grande extensão de um país com dimensões continentais, bem como as diversas situações dos contatos linguísticos existentes na formação sociohistórica, sociodialetal e

sociolinguística brasileiras, principalmente, refletidas no quadro pronominal, que demonstra uma verdadeira rosa dos ventos, no que tange ao *continuum* de normas.

Quanto à atualização do quadro pronominal brasileiro, notam-se direcionamentos que ajudam na compreensão no modo como a língua vem se configurando, pois, considerando o percurso histórico traçado na seção 3, a reorganização do quadro pronominal do PB reflete aspectos do latim, sobretudo, da variedade vulgar – em alguns contextos, o dativo desempenhava papel de nominativo –, para além dos aspectos atinentes às diversas situações de contato e do multilinguismo generalizado – o português europeu, as línguas gerais indígenas, e o português geral brasileiro, representado pelos povos africanos escravizados e os afrodescendentes –, dos ciclos econômicos – propiciadores, sobretudo, dos movimentos de colonização e povoamento –, dentre outros aspectos.

Conforme sistematizado na seção 4, algumas pesquisas, sob diversas perspectivas teóricas e metodológicas, têm se debruçado sobre esse fenômeno tanto na variedade do PB quanto nas variedades do PE e PA. Tal fato demonstra a produtividade quanto à notoriedade desse fator linguístico que comporta faces e caminhos de interpretações diversos, a depender da concepção que fundamenta o estudo. Portanto, a revisão de literatura realizada nesta tese possui relevância ímpar, haja vista a subsidiar os caminhos de pesquisas futuras que adotarão como objeto de análise a estrutura *para/prá + eu/mim + infinitivo*.

Nesse sentido, quanto à consulta feita nas gramáticas e a importância destas para a construção do respeito à diversidade de normas, do combate ao preconceito linguístico, à implantação de um olhar decolonial sobre os usos linguísticos, bem como à aceitação dos falantes, mediante as variantes que utilizam, algumas formulações são salutares, a saber:

i. Os dados evidenciados nesta tese coincidem com as proposições teóricas mencionadas nas três gramáticas históricas consultadas – Nunes (1945), Said Ali (1964) e Coutinho (1976); as duas gramáticas descritivas – Neves (2000) e Castilho (2016) – também admitem uma mobilidade maior entre as funções sintáticas exercidas pelos pronomes pessoais; entre duas gramáticas pedagógicas/escolares – Bagno (2001) e Cereja e Cochar (2013) –, pois, apenas, em Bechara (2010) encontram-se orientações mais rígidas e distantes da situação de reorganização do quadro pronominal brasileiro, retratando, assim como a única gramática normativa/tradicional, Lima (2018), aspectos mais conservadores, quanto às funções e os papéis sintáticos dos pronomes *eu* e *mim*.

ii. Nesse sentido, por vezes, as perspectivas mais conservadoras tornam-se, conforme história do ensino de língua portuguesa no Brasil, as que gozam de privilégio, ocupando lugar nas salas de aulas, conforme Soares (2002). Não rechaçando essas

perspectivas, até porque elas retratam um tempo pretérito, no que tange ao quadro pronominal do PB, e como os pronomes eram categorizados, sobretudo, tendo seus papéis e flexões estanques. Todavia, acredita-se que, com as contribuições das diversas vertentes teóricas da Linguística, outras perspectivas de gramáticas devam ser contempladas no ensino de língua materna, em que a diversidade de normas esteja representada, a partir do contexto comunicativo do usuário da língua, quer seja na modalidade escrita, quer seja na modalidade oral.

No que é atinente às perspectivas das vertentes teóricas da Linguística que se debruçaram sobre o fenômeno ora estudado, validam-se as diversas contribuições trazidas por elas, à luz dos princípios teóricos que as fundamentam. Defende-se, nesse sentido, que cada estudo ajuda a elucidar possíveis fatos, assim, colaborando para o progresso da ciência e para o entendimento sobre o fenômeno. Embora o sincretismo teórico seja interessante, nesta tese, optou-se por fazer um diálogo mais profícuo com aqueles estudos que utilizaram os mesmos fundamentos que orientaram as análises aqui empreendidas.

Sumarizando a comparação dos resultados desses antecessores com os alcançados nesta tese, ressaltam-se que:

i. Em Monteiro (1994); em Carvalho e Almeida, (2013); em Novais (2018), em Vaz (2020); nos *corpora* do Projeto NURC 70/90, assim como no agrupamento dos *corpora*, Projeto NURC e do Projeto ALiB, com informantes de nível universitário, a forma pronominal canônica – *eu* – foi mais proeminente, isso denota que, nos respectivos *corpora*, alguns informantes optam pela forma de prestígio, principalmente quando esses possuem maior tempo de contato com a Educação formal. No entanto, independente do tempo de escolarização, há áreas e contextos que tendem a favorecer essas escolhas pronominais.

ii. A forma pronominal inovadora – *mim* – figura com maior propensão de uso em: Figueredo (2007); Maia (2016); Silva e Nogueira (2018); Gomes (2019); Soares e Milane (2022), também fôra registrada nas obras clássicas dos estudos dialetais brasileiros: Amaral (1920); Marroquim (1934); e Nascentes (1953), além dos resultados catalogados pelos trabalhos de Santos e Paim (2018); Martins (2019); e Santos e Ribeiro (2020) e nos resultados desta tese, no *corpus* das capitais do Projeto ALiB.

iii. Notam-se diferentes estruturas de organização nos quadros pronominais expostos por Lucchesi (2008); e Lucchesi e Mendes (2009), quando retrata a flexão de caso dos pronomes pessoais na norma culta do PB, no português afro-brasileiro, bem como no português falado em Cabo Verde e em Santo Tomé. Desse modo, evidenciando aspectos importantes, tais como: o PB apresenta um quadro em variação; nas variedades crioulas, o

quadro evidencia flexões sem marcas de valor referencial; ao passo que as variedades cultas tendem a manter as flexões em suas posições consagradas. Pena-Ferreira (2015), em uma perspectiva mais ampla, demonstra aproximações entre o PB, o PE e o PA, haja vista a possibilidade de o fenômeno ser encontrado nessas variedades, não sendo peculiar a somente uma delas, tanto na língua falada quanto na língua escrita.

Com isso, certamente, urge que novos quadros pronominais sejam esboçados, tal como um dos apresentados por Mendes (2016), com vistas a retratar os diversos rearranjos nos usos dos pronomes pessoais tanto na língua falada, nas variedades do PB, de modo a repensar as afirmações encontradas em Monteiro (1994), quando se referem ao *eu* como única forma pronominal a ocupar a posição de sujeito, nas variedades cultas.

Quanto aos aspectos da eleição dos *corpora*, suas respectivas histórias, os frutos e os estágios atuais do Projeto NURC e do Projeto ALiB, evidenciados na seção 5, cabe ressaltar o desafio que é empreender uma investigação utilizando dois marcos importantes para os estudos sobre a língua falada dos brasileiros nativos, por possuírem naturezas e propostas metodológicas diferentes, bem como terem sido construídos em momentos peculiares da história social, cultural e linguística do Brasil. Utilizá-los, pois, constitui-se uma tarefa muito árdua, porém, bastante prazerosa e proveitosa, do ponto de vista científico.

A partir dessas considerações, frisa-se que, metodologicamente, algumas perspectivas, certamente, seriam revistas, conforme concepção do momento atual, a exemplo do qualificador *culta*, naquele período, atribuído aos informantes com nível universitário, pertencentes às capitais alvo do Projeto NURC. No que tange ao Projeto ALiB, talvez, fosse necessário repensar o questionário – levando em consideração a otimização do tempo empregado na documentação linguística em uma sociedade cada vez mais dinâmica e a reformulação de algumas questões, conforme mencionado na subseção 5.2.

Ainda sobre esses dois Projetos, o NURC com 54 anos e o ALiB com 27 anos, cabe destacar que eles se configuram como referências para o estudo sobre o português falado no Brasil. Desse modo, ressaltam-se dois aspectos: i. as periodizações esboçadas nesta tese são importantes, com vistas a registrar as fases e os momentos desses dois grandes Projetos; ii. os atuais estágios, em fase de informatização desses *corpora*, a fim de salvaguardar amostras do português falado em épocas distintas, 70/90/2000, bem como para fornecer melhor acesso aos dados coletados pelas equipes, tornando as pesquisas linguísticas democraticamente mais difundidas, com vistas à popularização do labor da Dialectologia, da Sociolinguística e, conseqüentemente, da Linguística.

Nesse sentido, sobre o futuro das pesquisas sobre a língua falada e a constituição, manutenção e preservação dos acervos de dados orais, destacam-se as ações programáticas, sugeridas como frentes de pesquisas, a fim de fomentar os estudos sobre o português falado no Brasil, são elas: i. autoavaliação - preservação - adequação; ii. informatização e digitalização do *corpus*; iii. aproximação do ensino de língua portuguesa; iv. divulgação das pesquisas; e v. formação de novos pesquisadores, a partir de um letramento sociodialetoológico, conforme subseção 5.2.1.

A partir da eleição dos *corpora*, estabeleceu-se a metodologia desta tese, descrita na seção 6. Vale mencionar o desafio em tentar harmonizar aspectos distintos e tentar extrair dados que pudessem evidenciar as nuances do fenômeno em sincronias diferentes do PB. Dessa forma, os caminhos foram tratá-los de forma isolada, inicialmente, e juntá-los, quando os aspectos fossem intercomparáveis. Posteriormente, seguindo critérios e rigores científicos e, assim, preservando a fidedignidade dos dados.

Diante das análises, seção 7, e a partir dos fundamentos teóricos principais adotados, notou-se que o fenômeno em questão apresenta um comportamento diferenciado, quando visto sob fatores intra e extralinguísticos. Desse modo, verificou-se que, quanto aos aspectos areais, nos *corpora* analisados, os pronomes *eu* e *mim* em posição de sujeito, precedidos pela preposição *para*, demonstram um país diverso, com áreas complexas e áreas delimitadas, quanto aos espaços dialetais. Os dados espelham, sobretudo, a interface dos aspectos linguísticos, sociohistóricos, espaciais e sociais, sobretudo, revelando uma verdadeira diversidade linguística existente nas capitais brasileiras.

Nesse sentido, primeiramente, a distribuição das variantes comprovou que, nos *corpora* do Projeto NURC, quando considerados isoladamente, a partir da fala de informantes com nível universitário, o pronome do caso reto *eu* foi encontrado de forma proeminente, conforme percentual de 67% das ocorrências.

Diatopicamente, conforme demonstra a Carta experimental M01, Recife, Salvador, Rio de Janeiro e São Paulo, na amostra analisada, são favorecedoras quanto à aplicação da regra do pronome do caso reto *eu*, ao passo que Porto Alegre foi considerada como desfavorável. No que é atinente à história social, nota-se que as capitais mais antigas tendem à utilização da forma de prestígio, em detrimento da capital mais recente, que opta pela forma que goza de desprestígio social e avaliação negativa, de modo geral, na norma linguística brasileira.

Posteriormente, as análises estatísticas realizadas pelo *Goldvarb X* apontaram que, nos *corpora* das décadas de 70 e 90, a única variável selecionada e relevante para o uso da forma

conservadora – o pronome *eu* – foi: a função sintática da oração infinitiva, com destaque para os contextos oracionais favorecedores: as orações substantivas objetivas e as orações adverbiais finais.

Em se tratando dos dados do Projeto ALiB, a distribuição das variantes comprovou que, a partir da fala de 200 informantes, oriundos das 25 capitais brasileiras, o pronome do caso oblíquo *mim* foi encontrado de forma predominante, conforme percentual de 68,2% das ocorrências válidas.

Diatopicamente, *a priori*, na amostra analisada, há capitais mais favoráveis quanto à utilização do *eu*, tais como: Salvador, Fortaleza, Recife, Belém, São Paulo, Aracaju, Natal, Belo Horizonte, Porto Alegre, Manaus, João Pessoa, Teresina, Rio de Janeiro, Maceió e Macapá; há àquela que ocupa posição intermediária, embora tendam levemente à forma conservadora, Curitiba. Por sua vez, ainda no que tange à realidade, Goiânia, Rio Branco, São Luís, Porto Velho, Boa Vista, Campo Grande, Florianópolis, Cuiabá e Vitória são consideradas como desfavorecedoras quanto a aplicação da regra do pronome do caso reto *eu*, conforme Carta experimental M02.

As análises estatísticas, *a posteriori*, apontaram que as variáveis selecionadas e relevantes para o uso do pronome *eu* foram: região geográfica, localidade, escolaridade, tipo de questionário e forma da preposição. De acordo com os resultados obtidos, ressalta-se que, nas regiões geográficas, não há limites definidos quanto aos usos dos dois pronomes, pois há uma diversidade.

Além disso, o fator escolaridade, confirmou uma das hipóteses levantadas no início da pesquisa ora empreendida, pois há comportamentos linguísticos diferenciados quando observadas as falas dos informantes do nível fundamental e no nível universitário, conforme valores percentuais dispostos, de modo geral, na Tabela 13 e, de modo particular, em cada capital, na Carta experimental M03. Nesse sentido, ressalta-se que, conforme os dados, quanto maior o tempo de escolarização formal, os informantes tendem ao uso da forma de prestígio.

No eixo intralinguístico, ainda sobre os dados do Projeto ALiB, destacou-se a forma da preposição. Assim, os pesos relativos apontados na Tabela 17 revelaram que os contextos usuais em que os informantes utilizam a forma preposicionada plena, *para*, são os mais favorecedores para serem preenchidos pelo pronome *eu*. Ao passo que os contextos com a forma contracta, *pra*, são desfavorecedores, embora grande quantidade de ocorrências válidas computadas.

No agrupamento dos aspectos coincidentes e intercomparáveis dos *corpora*, Projeto NURC e Projeto ALiB, notou-se que a distribuição das variantes em estudo encontra-se dessa maneira: o pronome *eu*, forma de prestígio, foi encontrado majoritariamente, conforme percentual de 71,4% das ocorrências, nas cinco capitais: Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre. Portanto, todas as cinco localidades, conforme Carta experimental M04, são proeminentes quanto à prestigiada. Por sua vez, a variante *mim* possui 28,6% das ocorrências. Ressalta-se que, nesse agrupamento, todos os dados são de informantes com nível universitário dessas referidas capitais.

Logo, com base nos resultados dispostos nas cartas diatópicas M01, M02 e M04, afirma-se que, quanto ao século e a data de fundação, as capitais Recife, Salvador e São Paulo, nas décadas de 70/90 e nos anos 2000, tendem à conservação do uso predominante da forma prestigiada, o pronome *eu*. Ao passo que, no Rio de Janeiro, essa forma pronominal é favorecida estatisticamente na estrutura em análise *para/prá + eu/mim + infinitivo* nos dados dos anos 70/90 e no agrupamento dos *corpora*, mas apresenta uma distribuição equânime entre os pronomes nos anos 2000, se considerados os valores percentuais. Por fim, a capital mais recente, Porto Alegre, quando comparadas às anteriormente citadas, evidencia uma mudança de escolha pronominal, pois, antigamente, nas décadas de 70 e nos anos 2000, a forma majoritária nos usos é o pronome *mim*, tendência que muda, quanto são observados somente os dados, nos *corpora*, dos informantes com nível universitário completo.

Após as rodadas estatísticas, no agrupamento dos *corpora*, as análises demonstram que, apenas, uma variável foi selecionada e relevante: a função sintática da oração infinitiva, com destaque às orações adverbiais finais, como contextos linguísticos favoráveis à regra de aplicação, o pronome do caso reto *eu*.

Tais resultados, após as análises estatísticas e interpretação dos dados, reforçam a hipótese principal da tese em que aponta para um comportamento diverso e diferenciado do fenômeno sob o ponto de vista espacial, oriundo de frutos dos processos sociais e históricos pretéritos, recentes e em curso nas capitais brasileiras.

Embora não se possa aventar um traçado de isoglossas delimitadoras de áreas dialetais, devido às limitações do estudo, alguns aspectos areais carecem de ilações, a saber: há subáreas que poderão ser confirmadas/refutadas com os dados das localidades do interior, pois as regiões geográficas Norte, Centro-Oeste e Sul possuem índices favoráveis aos usos do *mim*, enquanto uma extensa área, de São Luís a São Paulo, apresenta aspectos indefinidos quanto à caracterização dialetal. Necessita-se de mais dados e das análises sobre as localidades do interior, a fim de que se possam ter afirmações mais próximas do real, no que

tange ao traçado de áreas dialetais, colaborando com a agenda dos anseios dos dialetólogos brasileiros.

Ainda no sentido da confirmação das hipóteses, quando possível, nos dados do Projeto ALiB, a variável social escolaridade evidenciou a força exercida pelos meios de educação formal, tais como a escola e a universidade, haja vista que o fenômeno é refreado, quando os informantes possuem mais tempo de escolarização e, conseqüentemente, um nível de estudo maior. Ainda, consoante a tais dados, sob o aspecto escolaridade, foi possível notar que, predominantemente, as capitais que tiveram acesso ao ensino universitário no século XIX – Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre, todas possuem frequência elevada para o pronome *eu*, a exceção de Belo Horizonte. À luz dos dados agrupados dos *corpora*, a utilização do *eu* é elevada, ressaltando que, apenas, são dados orais de informantes com maior tempo de escolarização, das cinco capitais coincidentes nos *corpora*.

No que tange à faixa etária, aludindo às variáveis sociais, embora essa variável não tenha sido considerada como estatisticamente relevante, nos *corpora* do Projeto NURC, nota-se que a faixa II (36 – 55 anos) é àquela que possui bastante significação quanto aos usos do pronome do caso reto. Por sua vez, no *corpus* do Projeto ALiB, quanto à faixa etária, não há dados que comprovem o favorecimento. Assim, compreende-se que, para esse fenômeno, o recorte diageracional indica possíveis caminhos de interpretação, faixas etárias mais elevadas tendem ao conservadorismo e as faixas etárias mais novas tendem à inovação.

Vale mencionar outra variável social controlada, a década. Como foram dados coletados em décadas diferentes, 70 e 90 – Projeto NURC, e nos anos 2000 – Projeto ALiB, considerando o fator tempo como passível de exercer influência sobre as escolhas pronominais. Essa variável não foi considerada como estatisticamente relevante, haja vista que os dados da década de 70 e dos anos 2000 são desfavorecedores quanto à forma do caso reto. E, nos anos 90, essa forma é categórica.

Ainda sobre as variáveis sociais controladas, a variação diassexual, apenas controlada no *corpus* do Projeto ALiB, não foi considerada como significativa estatisticamente pelo *Goldvarb X*. No entanto, vale ressaltar os resultados obtidos, a saber: a fala dos homens foi considerada com a maior probabilidade de encontrar a variante inovadora, portanto, os usos linguísticos masculinos são favorecedores para os contextos de utilização do pronome *mim*. Ao passo que, na fala das mulheres, encontram-se frequências majoritárias de usos da forma *eu*.

A respeito das variáveis intralinguísticas, notaram-se aspectos diversos – a função sintática da oração infinitiva e forma da preposição. A função sintática da oração infinitiva

figura como uma variável fortemente favorecedora para a utilização do pronome do caso reto na estrutura em análise, conforme dados do Projeto NURC quanto na junção dos *corpora*, com dados intercomparáveis entre eles. Por seu turno, a forma da preposição indica possíveis interpretações, pois, nos dados do Projeto NURC e na junção dos *corpora*, os resultados computados com a forma pronominal *para* são categóricos quanto ao pronome *eu*, embora existam registros de ocorrências com a forma contracta, *pra*, nos dados do Projeto ALiB, a forma plena da preposição possui frequência elevada, portanto, favorecedora aos usos da forma pronominal de prestígio.

Os dados exibidos demonstraram, a partir da subseção 7.6, que a metalinguagem expressa nas falas dos informantes e também espelhada pelos dados categorizados como NO *Não Obtido*, que o fenômeno em estudo é alvo de avaliação social negativa, muitas vezes, atribuindo-o às origens de preconceito e desvalorização, permitindo uma abordagem da variação diarreferencial e também sob o viés das crenças e das atitudes linguísticas. Dessa maneira, há, cada vez mais, uma necessidade de estudos empíricos que desmistifiquem alguns estigmas linguísticos fomentados no seio social que, certamente, ferem e classificam falantes pelas variantes que utilizam. Por exemplo, o pronome *mim*, em posição de sujeito, conforme o senso comum, atribuído aos indígenas.

Ainda sobre o Não Dado e pesquisas que cataloguem poucos dados dialetais e sociolinguísticos, conforme subseção 2.4, nota-se que, historicamente, tais fenômenos não vêm recebendo um tratamento de modo a revelar importantes facetas e colaborar para o entendimento e para a compreensão dos dados linguísticos validados e, quantitativamente, tratados. Por isso, sugere-se que tal posicionamento seja modificado, sobretudo quando nos momentos da exegese e da cartografia, com vistas a permitir e a facilitar a leitura das cartas, conforme proposições esboçadas na referida subseção.

Nesse sentido, contemporaneamente, faz-se necessário que os instrumentos de pesquisa, sobretudo, no que tange ao questionário, possuam um caráter mais prático, sem perder a qualidade para a apuração dos fenômenos da língua. Por exemplo, a QMS 023, como discutida no item 6.6.2, induz um preenchimento pronominal, a partir das formulações hipotéticas, o que a torna uma situação mecanizada, por vezes, capaz de enviesar a obtenção do fenômeno buscado e a contabilização das respostas coletadas.

A cartografia da tese merece algumas considerações, uma vez que, na união entre áreas distintas, por meio do aspecto interdisciplinar da Dialectologia e, do método, geolinguístico, com a Geografia e as Ciências da Computação, foi possível elaborar três cartas temáticas que representam as variantes no espaço geográfico, simbolizando a variação

diatópica, uma carta com dados da variação social, apontando diferentes frequências de usos da forma pronominal *eu*, entre informantes com nível fundamental e nível universitário, e, por fim, uma carta qualitativa, referente à importância do não dado, representando a variação diarreferencial, por meio dos dados das Elocuções formais.

Cabe ressaltar que todas as cinco cartas possuem QR-code, que, quando escaneado, são direcionados à leitura e à análise feitas e hospedadas na plataforma digital aberta ME-QR. Com isso, visa-se aproximar os resultados das pesquisas sobre a língua falada com os aspectos das TDICs, a fim de popularizar a Linguística, ofertando fácil acesso a todos e colaborando com o ensino que valorize a diversidade de usos e caminhe na contramão do preconceito linguístico. Além de instaurar uma nova perspectiva de interfaces para os produtos geolinguísticos com forma e face abrazeiradas, dessa maneira, tornando a geolinguística do Brasil a pioneira nesse quesito.

Mediante às delimitações precisas desta pesquisa, empreendida por meio da tese de doutoramento, compreende-se que, ao analisar o fenômeno em variação entre as formas pronominais *eu* e *mim*, em posição de sujeito, em orações infinitivas, precedidas pela preposição *para*, pôde-se traçar os perfis dialetal e sociolinguístico, baseando-se em falas de brasileiros, datadas em três períodos diferentes de usos da língua, 70/90/2000, a partir dos dados pertencentes ao Projeto NURC e Projeto ALiB, importantes documentações sobre o PB falado.

Ademais, tais perfis são moldados por aspectos diatópicos – região geográfica e localidade – e sociais – principalmente, escolaridade –, para além dos aspectos linguísticos – sobretudo, a função sintática da oração infinitiva e a forma da preposição que antecede os pronomes, como elementos centrais na ótica sobre o fenômeno. Por fim, também se notam aspectos sociohistóricos, como por exemplo, o dativo desempenhando função de sujeito, herança da língua latina, documentada nas variedades do português, o século e a data de fundação, como também o surgimento do ensino universitário nas capitais.

Sugere-se, à propósito, a continuação da investigação sobre o fenômeno, mas, agora, explorando outras frentes, tais como os dados das localidades do interior, quanto aos dados do Projeto ALiB, observando a sociohistória, percurso de fundação e aspectos sobre a escolarização, haja vista que a união desses são importantes pilares para entendimento do quadro pronominal, no que concerne à língua falada no Brasil. Desse modo, fazendo contrapontos entre os resultados aqui expostos, com dados de capitais. Além disso, outros pronomes devem ser investigados, tal como o estudo de Pinho (2012), com as questões do QMS 027 e 028.

Conclui-se que, estando em franco desenvolvimento, os estudos dialetais, por meio da geolinguística pluridimensional, periodizados na 5ª fase, conforme subseção 2.1, possam ser frutíferos em todos os níveis de análise da língua, sobretudo, no morfossintático, a fim de que muitos caminhos sejam trilhados em prol de uma compreensão mais aprofundada dos fenômenos desse nível, além de contribuir para a agenda de pesquisas sobre a estrutura *para/prá + eu/mim + infinitivo*, adicionando esses resultados sobre as capitais aos das localidades do interior do Brasil. Desse modo, portanto, permitindo estabelecer parâmetros mais próximos de definição tanto a estratificação do fenômeno quanto a formulações mais condizentes com a real situação dos pronomes no PB, em relação ao quadro pronominal.

REFERÊNCIAS

ABAURRE, Maria Bernadete Marques; RODRIGUES, Angela C. S. (orgs.) **Gramática do Português Falado**. Campinas: Editora da Unicamp, Fapesp, vol. VIII — Novos estudos descritivos. 2022.

AGUILERA, Vanderci de Andrade. Léxico e áreas dialetais: o que podem demonstrar os dados do ALIB. **ANAIS - VI Congresso Internacional da Abralín / Dermeval da Hora (org.)**. - João Pessoa: Ideia, 2009. v. 2, p.4219-4233.

AGUILERA, Vanderci de Andrade. Crenças e atitudes lingüísticas: o que dizem os falantes das capitais brasileiras. **Estudos Lingüísticos**, n. 37 (2), São Paulo, p. 105-112, 2008a.

AGUILERA, Vanderci de Andrade. Crenças e atitudes lingüísticas: quem fala a língua brasileira? In: RONCARATI, C. N.; ABRAÇADO, J. (orgs.). **Português brasileiro II: contato lingüístico, heterogeneidade e história**. Niterói: EDUFF, 2008b, p. 311-333.

AGUILERA, Vanderci de Andrade. A metodologia e sua aplicação no campo. In: CARDOSO, Suzana et al. **Atlas Linguístico do Brasil (Introdução)**. Londrina: EDUEL, v.1, 2014a, p. 95-111.

AGUILERA, Vanderci de Andrade. (org.) **A geolingüística no Brasil: caminhos e perspectivas**. Londrina: Editora da UEL, 1998.

AGUILERA, Vanderci de Andrade. (org.) **A Geolingüística no Brasil: trilhas seguidas, caminhos a percorrer**. Londrina: EDUEL, 2005.

AGUILERA, Vanderci de Andrade; MILANI, Gleidy Aparecida; MOTA, Jacyra Andrade. (orgs.) **Documentos 1**. Salvador: EDUFBA, 2004.

AGUILERA, Vanderci de Andrade; ROMANO, Valter Pereira. (orgs.) **A geolingüística no Brasil: caminhos percorridos, horizontes alcançados**. Londrina: EDUEL, 2017.

AGUILERA, Vanderci de Andrade; ALTINO, Fabiane Cristina; ISQUERDO, Aparecida Negri. (Orgs.) **Atlas lingüístico do Brasil: descrevendo a língua e formando novos pesquisadores**. 1 CD- ROM: il. v. 1, 2009.

AGUILERA, Vanderci de Andrade; SILVA, Hélen Cristina da; MORATTO, Juliana. Aqui falam diferente? Um estudo metalingüístico com dados do Atlas Linguístico do Brasil. In: RAMOS, Conceição de Maria de Araujo; ALTINO, Fabiane Cristina; PAIM, Marcela Moura Torres. (Orgs.) **Documentos 8**. Projeto Atlas Linguístico do Brasil. O ALiB e suas mulheres: linhas de saudade e homenagem. Cascavel (PR): Edunioeste, 2023. p.187-217.

AGUILERA, Vanderci de Andrade.; YiDA, Vanessa. Projeto ALiB: uma análise das respostas e das não-respostas de informantes das capitais. **SIGNUM**, Estudos Linguísticos, Londrina, n.11/2, p. 15-31, dez. 2008.

ALMEIDA, Napoleão Mendes de. **Gramática Latina**: curso único e completo. 27. ed. São Paulo: Saraiva, 1995.

ALMEIDA, Marcos. **Latim para Todos**. Aracaju: Info Graphics, 2011.

ALVAR, Manuel. **Estructuralismo, geografia linguística y dialectología actual**. Madrid: Gredos, 1969.

ALVES, Diocles Igor Castro Pires. **As vogais médias pretônicas em Minas Gerais nos dados do ALiB**. 2022. 280f. Tese (Doutorado em Língua e Cultura) – Universidade Federal da Bahia, Salvador, 2022.

AMARAL, Amadeu. **O Dialeto Caipira**. 3.ed. São Paulo: HUCITEC, 1955.

ARAGÃO, Maria do Socorro. S. de; BEZERRA DE MENEZES, Cleuza. **Atlas Lingüístico da Paraíba**. Brasília: UFPB; CNPq, Coordenação Editorial, 1984.

BAGNO, Marcos. **A língua de Eulália**: novela sociolinguística. 15ª ed. São Paulo: Contexto, 2006.

BAGNO, Marcos. **Gramática pedagógica do português brasileiro**. São Paulo: Parábola Editorial, 2011.

BAGNO, Marcos. Cinquenta anos do NURC: das grandes realizações às novas esperanças. In: OLIVEIRA JR., Miguel (org.) **NURC 50 anos: 1969-2019**. 1ª ed. São Paulo: Parábola, 2019, p. 9-18.

BAGNO, Marcos (Org.). **Lingüística da Norma**. 2. ed. São Paulo: Loyla, 2002.

BAGNO, Marcos. **Não é errado falar assim!** em defesa do português brasileiro. São Paulo: Parábola Editorial, 2009.

BALBI, Adrien. **Atlas ethnographique du Globe, ou classification des peuples anciens et modernes d'après leurs langues... et suivi du tableau phisique, moral et politique des cinq parties du monde...** Paris: Chez Rey et. Gravier, 1826.

BANDEIRA, Manuel. **Seleta em prosa e verso**. Org: Emanuel de Moras. 4. ed. Rio de Janeiro: José Olympio, 1986.

BARBADINHO NETO, Raimundo (Org.). **Estudos filológicos**: volume dedicado à memória de Antenor Nascentes. Rio de Janeiro: Academia Brasileira de Letras, 2003. v. I. (Coleção Antônio de Morais Silva, Estudos de Língua Portuguesa).

BECHARA, Evanildo. **Moderna Gramática Portuguesa**. 37. ed. ampliada e atualizada conforme novo acordo Ortográfico. Rio de Janeiro: Nova Fronteira, 2009.

BECHARA, Evanildo. **Gramática escolar da língua portuguesa**. 2. ed. ampliada e atualizada pelo novo acordo Ortográfico. Rio de Janeiro: Nova Fronteira, 2010.

BENVENISTE, Émile. A natureza dos pronomes. *In*: BENVENISTE, Émile. **Problemas de linguística geral I**. Tradução Maria da Glória Novak e Maria Luisa Néri. 5. ed. Campinas: Pontes, 2005. p. 277-283.

BORTONI-RICARDO, Stella Maris. **Nos chegemos na escola, e agora?** Sociolinguística e educação. 2 ed. São Paulo: Parábola Editorial, 2005.

BRANDÃO, Silvia Figueiredo. **A geografia lingüística no Brasil**. São Paulo: Ática, 1991.

BRASIL. **Decreto nº 7.387, de 9 de dezembro de 2010**. Institui o Inventário Nacional da Diversidade Linguística e dá outras providências. Brasília, DF, 2010. Disponível em: http://portal.iphan.gov.br/uploads/legislacao/Decreto_n_7.387_de_9_de_dezembro_de_2010.pdf Acesso em 18 maio 2021.

BRASIL. **Parâmetros curriculares nacionais**. Língua Portuguesa. Brasília: MEC/SEF, 1997.

BRASIL. **Base Nacional Curricular Comum**. Disponível em: <http://basenacionalcomum.mec.gov.br/>>. Acesso em: 24 dez. 2017. <

BRASIL. Portaria n.º 536, de 26 de maio de 1952. **Diário Oficial da União**, Rio de Janeiro, 22 maio 1952.

BRASIL. Decreto n.º 30.643, de 20 de março de 1952. Institui o Centro de Pesquisas da Casa de Rui Barbosa e dispõe sobre seu funcionamento. **Diário Oficial da União**, Rio de Janeiro, 20 mar. 1952. Seção 1. p.4665.

BYBEE, J. Mechanisms of change in grammaticalization: the role of frequency. *In*: Jandra, R.; Bybee, Joan. 2010. **Language, usage and cognition**. Cambridge: Cambridge.

CABRAL, A. S. A. C. et al. Atlas sonoro das línguas indígenas do Brasil (ALISB): um projeto em andamento. *In*: RAZK, A.; OLIVEIRA, M. B.; LIMA, A. F. (org.). **Estudos geossociolinguísticos do português brasileiro**. Campinas: Pontes, 2017. p. 137-156.

CALVET, Louis-Jean. **Sociolinguística: uma introdução crítica**. Tradução Marcos Marcionilo. São Paulo: Parábola Editorial, 2002.

CALLOU, Dinah Maria Isensee. (Org.) **A linguagem falada culta na cidade do Rio de Janeiro**: materiais para seu estudo. Rio de Janeiro. Serviço Gráfico/UFRJ, v. 1, Elocuções formais, 1991.

CALLOU, Dinah Maria Isensee; LOPES, Célia Regina. (Orgs.) **A linguagem falada culta na cidade do Rio de Janeiro**: materiais para seu estudo. Rio de Janeiro. Faculdade de Letras/UFRJ, v. 3, Diálogo entre dois informantes, 1994.

CALLOU, Dinah. Quando a Dialectologia e a Sociolinguística se encontram. *In*: **Estudos Linguísticos e Literários**, n. 41, Salvador, Programa de Pós-Graduação em Língua e Cultura, Programa de Pós-Graduação em Literatura e Cultura da Universidade Federal da Bahia, p. 2948, jan./jun. 2010.

CALLOU, Dinah; BARBOSA, Afrânio; LOPES, Célia. O português do Brasil: Polarização Sociolinguística. In: MOTA, Jacira; CARDOSO, Suzana; MATTOS E SILVA, Rosa Virgínia (Orgs.). **Quinhentos anos de história Linguística do Brasil**. Salvador: Secretaria da Cultura e do Turismo do Estado da Bahia, 2006. p. 257-291.

CÂMARA JR., Joaquim Mattoso. **Estrutura da Língua Portuguesa**. 36. ed. Petrópolis: Vozes, [1978] 2004.

CANÇADO, Márcia. **Manual de Semântica: noções básicas e exercícios**. Belo Horizonte, Editora UFMG, 2008, 183p.

CARDOSO, Suzana et al. **Atlas Linguístico do Brasil**. v. 1 (Introdução). Londrina: EDUEL, 2014a.

CARDOSO, Suzana et al. **Atlas Linguístico do Brasil**. v. 2 (cartas linguísticas). Londrina: EDUEL, 2014b.

CARDOSO, Suzana Alice Marcelino. **Geolinguística: tradição e modernidade**. São Paulo: Parábola, 2010.

CARDOSO, Suzana Alice Marcelino. Tinha Nascentes razão? (Considerações sobre a divisão dialetal do Brasil). **Estudos: Linguísticos e Literários**, Salvador, n. 5, p. 47-59, 1986.

CARDOSO, Suzana. Língua: meio de opressão ou de socialização? In: FERREIRA, Carlota; MOTA, Jacyra; ANDRADE, Nadja; SILVA, Myrian; SILVA, Rosa Virgínia; CARDOSO, Suzana; ROLLEMBERG, Vera; FREITAS, Judith (Orgs.). **Diversidade do Português do Brasil** - estudos de dialectologia rural e outros. 2ª ed. revista. Salvador: Centro Editorial e Didático da UFBA; 1994. p. 229-233.

CARDOSO, Suzana. As vogais médias pretônicas no Brasil: uma visão diatópica. In: AGUILERA, Vanderci (org). **Português no Brasil: estudos fonéticos e fonológicos**. Londrina: Editora da UEL, 1999, p. 95-108.

CARDOSO, Suzana Alice Marcelino. A Dialectologia no Brasil: perspectivas. **DELTA**, São Paulo, v. 15, n. especial, p. 233-255, 1999.

CARDOSO, Suzana Alice Marcelino. Que dimensões outras, que não a diatópica, interessam aos atlas linguísticos? In: CONGRÈS INTERNATIONAL DE LINGUISTIQUE ET PHILOGIE ROMANES, 22., 1998, Bruxelas. **Actes...**Tübingen: Niemeyer, 2000a. v. 3, p. 411-416.

CARDOSO, Suzana Alice M. Da elaboração do questionário à formulação da pergunta: o “efeito gatilho” nas pesquisas geolinguísticas. In: VI Congresso Nacional da ABRALIN, 2009, João Pessoa. **Atas do VI Congresso Nacional da ABRALIN**. João Pessoa: Associação Brasileira de Linguística, 2009. p. 4506-4510.

CARDOSO, Suzana Alice M. Áreas dialetais do português brasileiro: o que dizem as primeiras cartas do ALiB. In: AGUILERA, Vanderci de Andrade; DOIRON, Maranúbia Barbosa. (orgs.) **Estudos Geossociolinguísticos brasileiros e europeus: uma homenagem a Michel Contini**. Londrina: Eduel, 2016, p.33-48.

CARDOSO, Suzana Alice M; MOTA, Jacyra Andrade; MATTOS E SILVA, Rosa Virgínia. (Orgs.). **Quinhentos anos de história Linguística do Brasil**. Salvador: Secretaria de Cultura e Turismo do Estado da Bahia, 2006.

CARDOSO, Suzana Alice M; MOTA, Jacyra A; PAIM, Marcela M. T.; RIBEIRO, Silvana S. C. (Orgs.). **Documentos 6: Projeto Atlas Linguístico do Brasil história e memórias**. Salvador: Quarteto, 2016.

CARDOSO, Suzana Alice M; MOTA, Jacyra A; PAIM, Marcela M. T.; RIBEIRO, Silvana S. C; TELES, Ana Regina T. F. (Orgs.). **Documentos 7: ALiB: 20 anos**. Salvador: Quarteto, 2017.

CARVALHO, Dannel da Silva. **A estrutura interna dos pronomes pessoais em português brasileiro**. 2008. 154 f. Tese (Doutorado em Letras) – Universidade Federal de Alagoas, Maceió, 2008.

CARVALHO, Jeany; ALMEIDA, Norma Lúcia F. de. **Variação entre as formas “preposição + eu e preposição + mim” no Português popular de Feira de Santana**. Disponível em: <<https://livrozilla.com/doc/1261407/varia%C3%A7%C3%A3o-entre-as-formas-%E2%80%9Cpreposi%C3%A7%C3%A3o---eu-e-preposi%C3%A7%C3%A3o---->>> Acesso em 16 nov. 2021.

CASTILHO, A. T. de (org.) **Gramática do Português Falado**. Campinas: Editora da Unicamp, Fapesp, vol. I — A ordem. 1991.

CASTILHO, Ataliba T. de; PRETI, Dino. (Orgs.). **A linguagem falada culta na cidade de São Paulo: materiais para seu estudo**. São Paulo: T.A Queiroz, v. 1, Elocuções formais. 1986.

CASTILHO, Ataliba T. de; PRETI, Dino. (Orgs.). **A linguagem falada culta na cidade de São Paulo: materiais para seu estudo**. São Paulo: T.A Queiroz, v. 2, Diálogo entre dois informantes. 1986-1987.

CASTILHO, Ataliba T. de. Políticas linguísticas no Brasil: o caso do Português Brasileiro. **Revista PUCP**, São Paulo, p. 271-297, 2001.

CASTILHO, Ataliba T. de. **Nova gramática do português brasileiro**. 1. ed., 4ª reimpressão. São Paulo: Contexto, 2016.

CASTILHO, Ataliba T. Rumos da dialetologia portuguesa. **ALFA**, n. 18/19, p. 115-153, 1972-1973. Disponível em: <http://seer.fclar.unesp.br/alfa/article/viewFile/3512/3285>. Acesso em: 12 ago. 2015.

CEGALLA, Domingos Paschoal. **Novíssima Gramática da Língua Portuguesa**. 48. ed. rev. São Paulo: Companhia Editora Nacional, 2008.

CEREJA, William Roberto; COCHAR, Thereza. **Gramática reflexiva: texto, semântica e interação**. Volume único/ William Roberto Cereja, Thereza Cochar Magalhães. 4. ed. reform. São Paulo: Atual, 2013.

CHAMBERS, J. K.; TRUDGILL, Peter. **La dialectología**. Madrid: Visor Libros, 1994.

CLARO, Daniela Barreiro; PAIM, Marcela Moura Torres; JESUS; Luis Emanuel Neves de. O sistema do Projeto Atlas Linguístico do Brasil: análise linguística automatizada. In: **Diadorim**, Rio de Janeiro, vol. 23, n. 1, p. 273-287, jan.-jun. 2021.

CLARO, Daniela Barreiro; TELLES, Ana Regina Silva; RIBEIRO; Silvana Soares Costa. Desafios do desenvolvimento do ALiBWeb: um sistema web para o Projeto ALiB. In: **Gallæcia**. Estudos de lingüística portuguesa e galega. Universidade de Santiago de Compostela, 2017. DOI <http://dx.doi.org/10.15304/cc.2017.1080.13>, pp. 299-306.

COMITÊ NACIONAL DO PROJETO ALiB: **Atlas Linguístico do Brasil**: Questionários. Londrina: UEL, 2001.

COSERIU, Eugenio. **Sincronia, Diacronia e História**. El problema Del Cambio linguístico. Madrid: Gredos, 1988.

COSTA, Celiane Sousa. **Variação e territorialização linguísticas**: um estudo geolinguístico da diversidade lexical em comunidades quilombolas do Baixo Amazonas. 2019. Tese (Doutorado) - Universidade Federal do Pará, Instituto de Letras e Comunicação, Belém, 2019. Programa de Pós-Graduação em Letras.

COUTINHO, Ismael de Lima. **Gramática Histórica**. 7. ed. rev. Rio de Janeiro: Ao Livro Técnico, 1976.

CUNHA, Celso e CINTRA, Luís Felipe Lindley. **Nova gramática do português contemporâneo**. 3. ed. Rio de Janeiro: Nova Fronteira, 2001 [1985].

DIAS, Marcelo Pires. **Atlas geossociolinguístico quilombola do Nordeste do Pará (AGQUINPA)**. 2017. 2 v. Tese (Doutorado) - Universidade Federal do Pará, Instituto de Letras e Comunicação, Belém, 2017. Programa de Pós-Graduação em Letras.

FARACO, Carlos Alberto. **Norma culta brasileira**: desatando alguns nós. São Paulo: Parábola Editorial, 2008, 200p.

FARACO, Carlos Alberto; MOURA, Francisco M. de; MARUXO, José H. Jr. **Gramática**. 20. ed. 8ª impressão. São Paulo: Editora Ática. 2016.

FARACO, Carlos Alberto. **História sociopolítica da língua portuguesa**. São Paulo: Parábola Editorial, 2016, 400p.

FERREIRA, Carlota; FREITAS, Judith; MOTA, Jacyra Andrade; ANDRADE, Nadja; CARDOSO, Suzana Alice Marcelino.; ROLLEMBERG, Vera.; ROSSI, Nelson. **Atlas Lingüístico de Sergipe**. Salvador: UFBA / Fundação Estadual de Cultura de Sergipe, 1987. Salvador: EDUFBA, 2005.

FERREIRA, Carlota; CARDOSO, Suzana Alice Marcelino. **A Dialectologia no Brasil**. São Paulo: Contexto, 1994.

FERREIRA, Carlota; MOTA, Jacyra; ANDRADE, Nadja; SILVA, Myrian; SILVA, Rosa Virgínia; CARDOSO, Suzana; ROLLEMBERG, Vera; FREITAS, Judith (Orgs.). **Diversidade do Português do Brasil** - estudos de dialectologia rural e outros. 2ª ed. revista. Salvador: Centro Editorial e Didático da UFBA; 1994.

FERREIRA, Carlota; MOTA, Jacyra; ANDRADE; ROLLEMBERG, Vera. Sergipe e Bahia: algumas diferenças lexicais. In: FERREIRA, Carlota; MOTA, Jacyra; ANDRADE, Nadja; SILVA, Myrian; SILVA, Rosa Virgínia; CARDOSO, Suzana; ROLLEMBERG, Vera;

FREITAS, Judith (Orgs.). **Diversidade do Português do Brasil** - estudos de dialectologia rural e outros. 2ª ed. revista. Salvador: Centro Editorial e Didático da UFBA; 1994.

FERREIRA, Carlota; CARDOSO, Suzana Alice Marcelino. Um panorama da dialectologia no Brasil. **Revista Internacional de Língua Portuguesa**, Lisboa, n. 14, p. 91-105, 1995.

FERREIRA, Carlota. Atlas Prévio dos Falares Baianos: alguns aspectos metodológicos. In: AGUILERA, Vanderci de Andrade. (Org.). **A geolinguística no Brasil**: caminhos e perspectivas. Londrina: UEL, 1998. p. 15-30.

FERREIRA, Manuela de Barros; CARRILHO, Enertina; LOBO, Maria; SARAMAGO, João; CRUZ, Luísa Segura da. Variação linguística: perspectiva dialectológica. In: FARIA, Isabel Hub; PEDRO, Emília Ribeiro; DUARTE, Inês; GOUVEIA, Carlos A. M. (orgs) **Introdução à Linguística Geral e Portuguesa**. 1996, p. 479-502.

FIGUEIREDO, João Ricardo Melo. **Variação e Mudança no Uso do Sujeito de Primeira Pessoa do Singular em Orações Infinitivas Iniciadas por “para” na Fala Carioca**. 2007. 121 f. Dissertação (Mestrado em Linguística) – Rio de Janeiro: UFRJ, 2007.

FIGUEIREDO JR., Selmo Ribeiro. **Atlas linguístico pluridimensional do português paulista**: níveis semântico-lexical e fonético-fonológico do vernáculo da região do Médio Tietê. Orientadores: Manoel M. Santiago-Almeida & Harald Thun. 2018. 2.120 p. Tese (Doutorado em Letras) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2019.

FIGUEIREDO JR., Selmo Ribeiro; MARGOTTI, Felício Wessling; ALMEIDA, Manoel Mourivaldo Santiago; NASCIMENTO; Juscelino Francisco do. Metodologia Geolinguística: agentes em geral e técnicas de inquérito semântico-lexical em específico. In: **Cad. Est. Ling.**, Campinas, v.63, p. 1-16, e021010, 2021.

FREITAG, Raquel M. Ko. Sociolinguística no/do Brasil. In: **Cadernos de Estudos Linguísticos**. v. 58.3, p. 445-460, set./dez. 2016.

FREITAG, Raquel M. Ko. NURC, um banco de dados sociolinguístico. In: OLIVEIRA JR, Miguel. **NURC 50 anos** – 1969 - 2019. São Paulo: Parábola, 2019, p.125-134.

FREITAG, Raquel M. Ko; SANTANA, Cristiane Conceição de; ANDRADE, Thais Regina Conceição de; SOUSA, Valéria Santos. Avaliação e variação linguística: estereótipos, marcadores e indicadores em uma comunidade escolar. In: FREITAG, Raquel M. Ko; SEVERO; Cristine Gorski; GORSKI, Edair Maria. (orgs) **Sociolinguística e política linguística**: olhares contemporâneos. São Paulo: Blucher, 2016, p.139-159.

GILLIÉRON, J. & EDMONT, E. (1902-1910) **Atlas Linguistique de la France (ALF)**, 35 fasc. Paris: Honoré Champion.

GOMES, A. B. S. **Tinha só um dia pra mim fazer a inscrição**. Um estudo sociolinguístico da variação pronominal em orações infinitivas iniciadas por para. 2019. 145f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista, São José do Rio Preto, 2019.

GUY, Gregory. A questão da crioulização no português do Brasil. In: ZILLES, Ana Maria Staht. **Estudos de variação lingüística no Brasil e no Cone Sul**. Porto Alegre: EDUFRGS, 2005.

GUY, Gregory; ZILLES, Ana. **Sociolinguística Quantitativa: instrumental de análise**. São Paulo: Parábola, 2007.

HILGERT, José. (Org.). **A linguagem falada culta na cidade de Porto Alegre: materiais para o seu estudo**. Passo Fundo: Ediupf; Porto Alegre: Ed. Universidade/UFRGS, v.1, Diálogos entre informante e documentador. 1997.

HILGERT, José. (Org.). **A linguagem falada culta na cidade de Porto Alegre: elocuições formais**. Porto Alegre: Editora da UFRGS, 2007.

HILGERT, José. (Org.). **A linguagem falada culta na cidade de Porto Alegre: diálogos entre dois informantes**. Florianópolis: Insular, 2009.

ILARI, Rodolfo.; BASSO, Renato. **O português da gente: a língua que estudamos a língua que falamos**. São Paulo: Contexto: 2006.

ISQUERDO, Aparecida Negri. Capitais brasileiras: um olhar para a história da cidade e a história do nome. In: CARDOSO, Suzana et al. **Atlas Linguístico do Brasil**. v. 2 (Cartas Linguísticas). Londrina: EDUEL, 2014b, p. 11-26.

ISQUERDO, Aparecida Negri; ALTINO, Fabiane Cristina; AGUILERA, Vanderci de Andrade. (Orgs.) **Atlas linguístico do Brasil: descrevendo a língua e formando novos pesquisadores**. 1 CD- ROM: il. v.2, 2012.

ISQUERDO, Aparecida Negri; ALTINO, Fabiane Cristina; AGUILERA, Vanderci de Andrade. (Orgs.) **Atlas linguístico do Brasil: descrevendo a língua e formando novos pesquisadores**. 1 CD- ROM: il. v. 3, 2014.

ISQUERDO, Aparecida Negri; ALTINO, Fabiane Cristina; AGUILERA, Vanderci de Andrade. (Orgs.) **Atlas linguístico do Brasil: descrevendo a língua e formando novos pesquisadores**. 1 CD- ROM: il. v. 4, 2016.

ISQUERDO, Aparecida Negri; TELES; Ana Regina Torres Ferreira. A rede de pontos. In: CARDOSO, Suzana et al. **Atlas Linguístico do Brasil**. v. 1 (Introdução). Londrina: EDUEL, 2014^a, p. 37-78.

JABERG, Karl; JUD, Jakob. **Sprach- und Sachatlas Italiens und derSüdschweiz, I-VIII**. Zofingen: Rieger, 1928-1940.

KOCH, Walter. O questionário morfossintático. In: AGUILERA, Vanderci de Andrade; MILANI, Gleidy Aparecida; MOTA, Jacyra Andrade. (orgs.) **Documentos 1**. Salvador: EDUFBA, 2004, p.71-75.

LABOV, William. **Modelos sociolingüísticos**. Tradução por José Miguel Marinas Herreras. Madrid: Cátedra, 1983.

LABOV, William. **Padrões Sociolingüísticos**. Tradução por Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008 [1972].

LAWRENCE, Helen; TAGLIAMONTE, Sali. **Goldvarb2001: A Multivariate Analysis Application for Windows**. Nova York: Department of Language and Linguistic Science University of York, 2001.

LEITE, Yonne; CALLOU, Dinah. **Como falam os brasileiros**. 4. ed. Rio de Janeiro: Zahar, 2010.

LEMLE, Miriam; NARO, Antony J. **Competências básicas do português**. Rio de Janeiro, MOBREAL, Fundação Ford, 1999, 151p.

LIMA, Carlos Henrique da Rocha. **Gramática normativa da língua portuguesa**; curso médio; prefácio de Serafim da Silva Neto. 26. ed. Rio de Janeiro: José Olympio, 1985.

LIMA, Fabio Ronne de Santana; COSTA, Daniela de Souza Silva. Dialetologia e Sociolingüística educacional em interface: uma proposta para o ensino de língua portuguesa a partir de dados do Atlas Linguístico do Brasil. **Diadorim**, Rio de Janeiro, vol. 24, número 2, p. 709-730, 2022.

LIMA, Rocha. **Gramática Normativa da Língua Portuguesa**. 54. ed. Rio de Janeiro: José Olympio, 2018.

LINS, Alex Batista. Três hipóteses e alguns caminhos para melhor compreender o processo constitutivo do português brasileiro. In: OLIVEIRA, K., CUNHA E SOUZA, H. F.; SOLEDADE, J., (Orgs.). **Do português arcaico ao português brasileiro: outras histórias** [online]. Salvador: EDUFBA, 2009. p. 271-296.

LOPES, Célia Regina; CUNHA, C. de S. **Pronomes pessoais: a pesquisa sociolingüística e a atualização da gramática**. I Congresso Internacional da Abralín. Salvador: UFBA, 1994.

LUCCHESI, Dante. **Língua e Sociedade Partidas: a polarização sociolingüística do Brasil**. Contexto: Loyola, 2015.

LUCCHESI, Dante. As duas vertentes da história sociolingüística do Brasil. **DELTA**, São Paulo, v. 17, n. 1, p. 97-132, 2001.

LUCCHESI, Dante. Norma lingüística e realidade social. In: BAGNO, Marcos (Org.). **Lingüística da Norma**. 2. ed. São Paulo: Loyola, 2002. p. 63-90.

LUCCHESI, Dante. Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão do conjunto. In: RONCARATI, Claudia; ABRAÇADO, Jussara (Orgs.). **Português brasileiro II: contato linguístico, heterogeneidade e história**. Rio de Janeiro: Letras, 2008. p. 367-389.

LUCCHESI, Dante. **Sistema, Mudança e Linguagem: um percurso na história da lingüística moderna**. São Paulo: Parábola, 2004.

LUCCHESI, Dante; MENDES, Elisângela dos Passos. A flexão de caso dos pronomes pessoais. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza. (Orgs.) **O português afrobrasileiro**. - Salvador: EDUFBA, 2009, p. 471-488.

MAIA, LÍlian de Sant,Anna. **Uso variável de “eu/mim” e “mim/me” em textos escritos de anos finais do ensino fundamental: a influência de fatores extralingüísticos**. 2016. 86f. Dissertação (Mestrado Profissional em Letras) – Universidade Estadual de Santa Cruz, Ilhéus, 2016.

MARCUSCHI, Luiz Antônio. **Produção textual, análise de gêneros e compreensão**. São Paulo: Parábola Editorial, 2008.

MARROQUIM, Mario. **A língua do Nordeste: Alagoas e Pernambuco**. 1. ed. Maceió: EDUFAL, 1934.

MARTINS, Mariana Spagnolo. Os Pronomes pessoais (eu e mim) nas capitais brasileiras a partir dos dados do projeto ALiB. **Revista Primeira Escrita**, Aquidauana, n. 6, p. 134-144, 2019.

MARTINET, André. **De las estepas a los océanos: el indoeuropeo y los “indoeuropeos”**. Madrid: Gredos, 1997.

MATTOS E SILVA, Rosa Virgínia. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola, 2008.

MATTOS E SILVA, Rosa Virgínia. **O Português Arcaico: Morfologia e Sintaxe**. 2. ed. São Paulo: Contexto, 2001.

MATTOS E SILVA, Rosa Virgínia. **Contradições no ensino de Português: a língua que se fala x a língua que se ensina**. São Paulo: Contexto, 1995.

MAURER Jr., Theodoro Henrique. **O problema do latim vulgar**. Rio de Janeiro: Acadêmica, 1959.

MEDINA MORALES, Francisca. Problemas metodológicos de La sociolingüística histórica. In: **Forma y Función**, enero-diciembre, n. 18. Universidad Nacional de Colombia. Bogotá, D.C., 2005, p. 115-137.

MELLO, Heliana; RASO, Tommaso. Corpora de fala: Panorama e inovações recentes. In: OLIVEIRA JR., Miguel Oliveira (org.) **NURC 50 anos: 1969-2019**. 1ª ed. São Paulo: Parábola, 2019, p. 135-167.

MENDES, Elisângela dos Passos. **A flexão de caso pronominal no *continuum* do português popular da Bahia**. 2016. 270 f. Tese (Doutorado em Letras) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2016.

MIOTO, Carlos; SILVA, Maria Cristina Figueiredo; LOPES, Ruth Elizabeth Vasconcellos. **Novo Manual de Sintaxe**. 3 ed. Florianópolis: Insular, 2007.

MONTEIRO, José Lemos. **Pronomes pessoais: subsídios para uma gramática do português do Brasil**. Fortaleza: UFC, 1994.

MOTA, Jacyra Andrade. Estrela cadente nos atlas regionais brasileiros. **Revista do GELNE**, Fortaleza, n. 1, v.2, p. 25-31, 1999.

MOTA, Jacyra Andrade. O Projeto NURC no Brasil: 50 anos de história na Bahia. In: OLIVEIRA JR., Miguel Oliveira (org.) **NURC 50 anos: 1969-2019**. 1ª ed. São Paulo: Parábola, 2019, p. 55-79.

MOTA, Jacyra Andrade. Por onde caminha a palatalização do português brasileiro. In: AGUILERA, Vanderci de Andrade; DOIRON, Maranúbia Barbosa. (orgs.) **Estudos Geossociolinguísticos brasileiros e europeus: uma homenagem a Michel Contini**. Londrina: Eduel, 2016, p.49-65.

MOTA, Jacyra Andrade; CARDOSO, Suzana A. M. (Orgs.). **Documentos 2: Projeto Atlas linguístico do Brasil**. Salvador: Quarteto, 2006.

MOTA, Jacyra Andrade. Áreas dialetais brasileiras. In: CARDOSO, Suzana Alice M.; MOTA, Jacyra Andrade; MATTOS E SILVA, Rosa Virgínia. (Orgs.). **Quinhentos anos de história Linguística do Brasil**. Salvador: Secretaria de Cultura e Turismo do Estado da Bahia, 2006. p. 319-357.

MOTA, Jacyra; ROLLEMBERG, Vera. (Orgs.). **A Linguagem Falada Culta na Cidade de Salvador**. Materiais para seu estudo. Salvador: Instituto de Letras da UFBA, vol. 1, Diálogos entre Informante e Documentador. 1994.

MOTA, Jacyra; ROLLEMBERG, Vera. (Orgs.). **A Linguagem Falada Culta na Cidade de Salvador**. Materiais para seu estudo. Salvador: Instituto de Letras da UFBA, vol. 2, Elocuções formais. 2006.

MOTA, Jacyra; PAIM, Marcela; ROLLEMBERG, Vera (Orgs.). **A Linguagem Falada Culta na Cidade de Salvador**. Materiais para seu estudo (década de 90). Salvador: Instituto de Letras da UFBA, vol. 3, Diálogos entre dois informantes. Salvador: Vento Leste, 2011.

MOTA, Jacyra Andrade; SILVA, Amanda dos Reis. Sociolinguística, uma disciplina histórica: retrospectivas, desenvolvimento e aplicações. **Macabéa** – Revista Eletrônica do Netlli. Cariri: PB. v. 8., n. 2., jul.-dez. 2019, p. 11-35.

MOTA, Jacyra Andrade; RIBEIRO, Silvana Soares Costa; OLIVEIRA, Josane Moreira. (orgs.). **Atlas linguístico do Brasil**. vol. 3:comentários às cartas linguísticas 1. Londrina: EDUEL, 2023.

NARO, Anthony Julius; SCHERRE, Maria Marta Pereira. **Origens do português brasileiro**. São Paulo: Parábola, 2007.

NASCENTES, Antenor. **O linguajar carioca em 1922**. Rio de Janeiro: Süsskind de Mendonça, 1922.

NASCENTES, Antenor. **O linguajar carioca**. 2. ed. Completamente refundida. Rio de Janeiro: Organização Simões, 1953.

NASCENTES, Antenor. **Bases para a elaboração do atlas lingüístico do Brasil**. Rio de Janeiro: MEC; Casa de Rui Barbosa, v. 1, 1958; v. 2, 1961.

NASCENTES, Antenor. Divisão dialectológica do território brasileiro. **Revista Brasileira de Geografia**, Rio de Janeiro, v. 17, n. 2, p. 213-219, abr./jun. 1955.

NASCENTES, Antenor. Études dialectologiques du Brésil. **ORBIS** - Bulletin International de Documentat ion Linguistique, Louvain, t. 2, n. 2, p. 438-444, 1953.

NEVES, Maria Helena de Moura. **Gramática de usos de português**. 2. ed. São Paulo: Editora Unesp, 2000.

NOVAIS, V. S. **Uso dos pronomes “mim” e “eu” na posição de sujeito em frases infinitivas iniciadas pela preposição para**: o que pensam os falantes universitários do Sertão Alagoano. 2018. 68f. Trabalho de Conclusão de Curso (Graduação em Letras) – Universidade Federal de Alagoas, Delmiro Gouveia, 2018.

NUNES, Dr. José Joaquim. **Compêndio de Gramática Histórica Portuguesa**. 8. ed. Lisboa: Livraria Clássica Editora, 1945.

OLIVEIRA, Gilvan Muller de. Plurilinguismo no Brasil: repressão e resistência. **Synergies Brésil**, nº 7, 2009, p.19-26.

OLIVEIRA, Josane Moreira de; MOTA, Jacyra Andrade; RIBEIRO, Silvana Soares Costa. (Orgs). **A Linguagem Falada Culta na Cidade de Salvador**. Materiais para seu estudo (década de 90). Salvador: Instituto de Letras da UFBA, vol. IV, Diálogos entre Informante e Documentador e Elocuções Formais. Salvador: Editora da UEFS, 2023. (no prelo).

OLIVEIRA, Josane Moreira de; PAIM, Marcela Moura Torres; RIBEIRO, Silvana Soares Costa. A importância do Atlas Lingüístico do Brasil para o ensino de português. **Revista Tabuleiro de Letras** (PPGEL, Salvador, online), v. 12, n. 3, dez. 2018.

OLIVEIRA, Josane Moreira de; GONÇALVES, Clézio Roberto; MOTA, Jacyra Andrade. Sociolinguística e Dialectologia: novas abordagem de pesquisa e ensino. **Diadorim**, Rio de Janeiro, vol. 23, n. 1, p. 288-303, jan.-jun. 2021.

OLIVEIRA JR., Miguel. O Projeto NURC Digital. In: OLIVEIRA JR., Miguel Oliveira (org.) **NURC 50 anos: 1969-2019**. 1ª ed. São Paulo: Parábola, 2019, p. 169-193.

OLIVEIRA JR., Miguel. NURC Digital Um protocolo para a digitalização, anotação, arquivamento e disseminação do material do Projeto da Norma Urbana Linguística Culta

(NURC). **CHIMERA**: Revista De Corpus De Lenguas Romances Y Estudios Lingüísticos, 3(2), 2016, 149–174. <https://doi.org/10.15366/chimera2016.3.2.004>.

PENA-FERREIRA, Ediene. Diferentes usos dos pronomes pessoais: inovação do Português Brasileiro ou tendência da Língua Portuguesa? *In: Atas do V Simpósio Mundial de Estudos de Língua Portuguesa*. Lecce, Itália, 2015.

PEREIRA, Maria Angela Botelho; RONCARATI, Cláudia Nívia. O caso do sujeito em orações infinitivas introduzidas por “para” no português do Rio. São Paulo: **Delta**, v. 9, n. 1, 1993.

PEREIRA, Eduardo Carlos. **Grammatica expositiva**. Curso superior. 20. ed. São Paulo: Companhia Editora Nacional, 1918 [1907]).

PERINI, Mário A. **Gramática descritiva do português**. 4. ed. 8ª impressão. São Paulo: Editora Ática, 2005.

PETER, Margarida Maria Taddoni. Línguas Africanas no Brasil. *In: CARDOSO, Suzana Alice M.; MOTA, Jacyra Andrade; MATTOS E SILVA, Rosa Virgínia. (Orgs.). Quinhentos anos de história Linguística do Brasil*. Salvador: Secretaria de Cultura e Turismo do Estado da Bahia, 2006. p. 117-142.

PINHO, Antonio José de. **Aspectos da história da língua**: Um estudo diacrônico e sincrônico dos pronomes oblíquos tônicos. 2012. 353 f. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2012.

PORTILHO, Danyelle Almeida Saraiva. **O falar amazônico**: uma análise da proposta de Nascentes (1953) a partir de dados do Projeto ALiB. 2013. 155 f. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal do Mato Grosso do Sul, Campo Grande, 2013.

PRETI, Dino; URBANO, Hudinilson. (Orgs.). **A linguagem falada culta na cidade de São Paulo**: materiais para seu estudo. São Paulo: T.A Queiroz, v. 3, Diálogo entre informante e documentador. 1988.

PRETI, Dino. (Org.) **Análise de textos orais**. 5ª ed. - São Paulo: Humanitas FFLCH/USP, 2001 - (PROJETOS PARALELOS: v.1).

PROJETO ALIB: 20 anos pesquisando a fala dos brasileiros. **Mesa redonda**. Coordenada por Silvana Soares Costa Ribeiro. Salvador. 1 vídeo (1h 29min 10seg). Publicado pela TV UFBA. Disponível em: <https://www.youtube.com/watch?v=QnsnXFRPbMI>. Acesso em: 23 nov. 2020.

RAMOS, Conceição de Maria de Araujo; ALTINO, Fabiane Cristina; PAIM, Marcela Moura Torres. (Orgs.) **Documentos 8**. Projeto Atlas Linguístico do Brasil. O ALiB e suas mulheres: linhas de saudade e homenagem. Cascavel (PR): Edunioeste, 2023. 510p.

RAVIZZA, Padre João. **Gramática Latina**. 9. ed. Niterói: Escolas profissionais Salesianas, 1940.

RIBEIRO, Silvana Soares Costa. **Brinquedos e brincadeiras infantis na área do “Falar Baiano”**. 2012. 752 f. Tese (Doutorado em Letras) – Universidade Federal da Bahia, Salvador, 2012.

RIBEIRO, José; ZÁGARI, Mário Roberto L.; PASSINI, José; GAIO, Antonio Pereira. **Esboço de um atlas lingüístico de Minas Gerais**. Rio de Janeiro: Fundação Casa de Rui Barbosa, 1977.

ROCHA, Ricardo Machado. **Morfossintaxe de Caso nos Pronomes Pessoais do PB/MG atual**. 2010. 133 f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

RODRIGUES, Aryon Dall'Igna. Tupi, tupinambá, línguas gerais e o português do Brasil. *In*: NOOL, Volker; DIETRICH, Wolf (Orgs.). **O português e o tupi no Brasil**. São Paulo: Contexto, 2010. p. 27-47.

RODRIGUES, Aryon Dall'Igna. As línguas gerais sul-americanas. **Papia**, v. 4, n. 2, p. 6-18, 1996.

ROLLEMBERG, Vera. Realizações palatais de /k/ e /g/ em área sergipana. *In*: FERREIRA, Carlota; MOTA, Jacyra; ANDRADE, Nadja; SILVA, Myrian; SILVA, Rosa Virgínia; CARDOSO, Suzana; ROLLEMBERG, Vera; FREITAS, Judith (Orgs.). **Diversidade do Português do Brasil** - estudos de dialectologia rural e outros. 2ª ed. revista. Salvador: Centro Editorial e Didático da UFBA; 1994. p. 107-110.

ROMANO, Valter Pereira. Balanço crítico da Geolinguística brasileira e a proposição de uma divisão. **Entretextos**, Londrina, UEL, v. 13, n. 2, p. 203-242, jul./dez. 2013.

ROMANO, Valter Pereira. **Em busca de falares a partir de áreas lexicais no centro-sul do Brasil**. 2015. 286 f. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2015.

ROMANO, Valter Pereira. Desdobramentos, Desafios E Perspectivas Da Geolinguística Pluridimensional No Brasil. *In*: MOTA, Jacyra Andrade; OLIVEIRA, Josane Moreira; PAIM, Marcela Moura Torres; RIBEIRO, Silvana Soares Costa (Orgs.). **Contribuições de estudos geolinguístico para o Português Brasileiro: uma homenagem a Suzana Cardoso**. 1ed. Salvador: EDUFBA, 2020, v. 1, p. 11-39.

ROSSI, Nelson. **Atlas prévio dos falares baianos**. Rio de Janeiro: INL-MEC, 1963.

ROSSI, Nelson. A Dialectologia. *In*: **Revista ALFA**, n. 11. (Revista da Faculdade de Filosofia, Ciências e Letras de Marília), Marília, SP: FFCL de Marília, 1967, p. 89- 128.

SÁ, Maria da Piedade M. de; CUNHA, Dórias de Arruda C. da; LIMA, Ana Maria; JR., Miguel Oliveira. (orgs). **A linguagem falada culta na cidade do Recife: materiais para seu estudo**. Recife: Ed. Universitária da UFPE, v. 1, Diálogos entre informantes e documentador. 1996.

SÁ, Maria da Piedade M. de; LIMA, Ana Maria; JR., Miguel Oliveira; CUNHA, Dórias de Arruda C. da. (orgs). **A linguagem falada culta na cidade do Recife: materiais para seu**

estudo. Recife: Programa de Pós-Graduação em Letras da UFPE, v. 2, Elocuções formais. 2005.

SÁ, Maria da Piedade M. de; LIMA, Ana Maria; CUNHA, Dórias de Arruda C. da; JR., Miguel Oliveira. (orgs). **A linguagem falada culta na cidade do Recife**: vol. 3, Diálogos entre dois informantes [recurso eletrônico]. Recife: Ed. UFPE, 2017.

SÁ, Edmilson José de. **Atlas Linguístico quilombola do Moxotó-Ipanema Pernambucano (ALQUIMIPE)**. Relatório de Pós-Doutorado, Belém - PA: Universidade Federal do Pará, 2018.

SAID ALI, M. **Gramática Histórica da Língua Portuguesa**. 6. ed. Rio de Janeiro: Edições Melhoramentos, 1964.

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

SANTOS, Grazielle Ferreira da Silva. **O léxico dos jogos e diversões infantis no corpus do Projeto ALiB: visitando o Falar Nordestino**. 2018. 207 f. Dissertação (Mestrado em Língua e Cultura) – Universidade Federal da Bahia, Salvador, 2018.

SANTOS, Grazielle Ferreira da Silva; PAIM, Marcela Moura Torres. Para mim ou para eu fazer? Um retrato da Região Sudeste do Brasil a partir dos dados do Projeto ALiB. **A cor das Letras**, Feira de Santana, v. 19, n. 2, p. 141-150, 2018.

SANTOS, Jessica Tavares dos. **Descrição do pronome na tradição gramatical brasileira do século XX**. 2017. 254 f. Tese (Doutorado em Estudos da Linguagem) Universidade Federal Fluminense, 2017.

SANTOS, Leandro Almeida dos. **Brincando pelos caminhos do Falar Fluminense**. 2016. 197 f. Dissertação (Mestrado em Língua e Cultura) – Universidade Federal da Bahia, Salvador, 2016.

SANTOS, Leandro Almeida dos. Pelos caminhos das áreas dialetais e lexicais do Brasil. **Calentoscópio**, v.8, nº 2, p.44-62, 2020.

SANTOS, Leandro Almeida dos. As não respostas e as fronteiras dialetais brasileiras: possíveis caminhos. **Revista Entrepalavras**, Fortaleza, ano 10, v. 10, n. 3, p. 1-19, set./dez. 2020.

SANTOS, Leandro Almeida dos; RIBEIRO, Silvana Soares Costa. O que dizem os brasileiros: “Pra eu” ou “Pra mim” fazer?: In: MOTA, Jacyra Andrade; OLIVEIRA, Josane Moreira; PAIM, Marcela Moura Torres; RIBEIRO, Silvana Soares Costa (Orgs.). **Contribuições de estudos geolinguístico para o Português Brasileiro: uma homenagem a Suzana Cardoso**. 1ed. Salvador: EDUFBA, 2020, v. 1, p. 167-187.

SANTOS, Leandro Almeida dos; RIBEIRO, Silvana Soares Costa. A mulher e a Dialetoologia brasileira. **Diadorim** (Rio de Janeiro), v. 23, p. 304-326, 2021.

SCHERRE, M. M. P. Respeito linguístico. In: MEDEIROS, R.; SCHERRE, P. **Dicionário: rumo à civilização da religião e ao bem viver.** 2019. Disponível em: <http://sementedossonhos.org/wp/>. Acesso em: 21 mar 2021.

SILVA, Alba Valéria T. A.; SANTOS, A. S.O Léxico no livro didático: *corpus* ALiB. In: SILVEIRA, Jader Luís da; SILVEIRA, Regiane Paula da. (Org.). **Linguística e literatura: cultura, sociedade e história.** 1ed. Formiga (MG): Real Conhecer, 2022, v. 10, p. 105-122.

SILVA, André Pereira da Silva; NOGUEIRA, Vilson Ramos dos Santos. **O uso dos pronomes mim e eu na função de sujeito na comunidade de fala de Jacobina-Ba:** análise variacionista. 2018. 62 f. Trabalho de Conclusão de Curso (Graduação em Letras Vernáculas) – Universidade Estadual da Bahia, Jacobina, 2018.

SILVA, Greize Alves da; AGUILERA, Vanderci de Andrade. Os atlas linguísticos brasileiros e o inquiridor: em busca de uma metodologia adequada. In: **SIGNUM: Est. Ling.,** Londrina, v. 12, n. 1, p. 317-341, jul. 2009.

SILVA, Greize Alves da; ROMANO, Valter Pereira. (Orgs.) **Tendências da Geolinguística brasileira e a nova geração de atlas linguísticos.** São Carlos: Pedro & João Editores, 2022. 400p.

SILVA, Greize Alves da; ROMANO, Valter Pereira. O Atlas Linguístico do Brasil e os Atlas de pequeno domínio: complementações e propósitos. In: SILVA, Greize Alves da; ROMANO, Valter Pereira. (Orgs.) **Tendências da Geolinguística brasileira e a nova geração de atlas linguísticos.** São Carlos: Pedro & João Editores, 2022. p. 17-46.

SILVA, Maria Cristina Figueredo. **A posição sujeito no português brasileiro:** frases finitas e infinitivas. Campinas, SP: Editora da UNICAMP, 1996.

SILVA, Nádia Leticia Pereira. **De amarrar o facão a casa da ruindade:** unidades fraseológicas nas áreas semânticas ciclos da vida e convívio social e comportamento social no interior do Maranhão e Ceará. 2020. 99 f. Dissertação de Mestrado (Graduação em Letras Vernáculas) – Programa de Pós-Graduação em Letras/ cch. Universidade Federal do Maranhão, São Luís, 2020.

SILVA, Rosa Virgínia. Dizem que vai mal o vernáculo no Brasil (Reflexão sobre o português e seu ensino). In: FERREIRA, Carlota; MOTA, Jacyra; ANDRADE, Nadja; SILVA, Myrian; SILVA, Rosa Virgínia; CARDOSO, Suzana; ROLLEMBERG, Vera; FREITAS, Judith (Orgs.). **Diversidade do Português do Brasil** - estudos de dialectologia rural e outros. 2ª ed. revista. Salvador: Centro Editorial e Didático da UFBA; 1994. p. 215-227.

SILVA-CORVALÁN, Carmen. **Sociolingüística: Teoría y Análisis.** Madrid: Editorial Alhambra, 1989.

SILVA NETO, Serafim da. **Guia para estudos dialectológicos.** Belém: Conselho Nacional de Pesquisas; Instituto Nacional de Pesquisas da Amazônia, 1957.

SOARES, Magda. Português na escola - história de uma disciplina curricular. In: BAGNO, Marcos (Org.). **Linguística da Norma.** 2. ed. São Paulo: Loyola, 2002. p. 155-177.

SOARES, Thaynara Cardoso; MILANI, Sebastião Elias. O uso do pronome oblíquo mim exercendo a função de sujeito em dados do NURC/RJ. In: **Revista Porto das Letras**, v. 08, nº 1, 2022, p. 167-177.

SOUZA, Ana Rita Carvalho de. **As estrelas no corpus do Projeto ALiB: revisitando o Falar Amazônico**. 2021. Dissertação. (Mestrado em Língua e Cultura). Universidade Federal da Bahia, Salvador. 2021.

TARALLO, Fernando. **A pesquisa Sociolinguística**. São Paulo: Ática, 2001.

TELES, Ana Regina Ferreira. **Cartografia e Georreferenciamento na Geolinguística: revisão e atualização das regiões dialetais e da rede de pontos para a elaboração do Atlas Linguístico do Brasil formuladas por Antenor Nascentes**. 2018. Tese (Doutorado em Língua e Cultura). Universidade Federal da Bahia, Salvador. 2018.

TELES, Ana Regina Ferreira; RIBEIRO; Silvana Soares. A Cartografia dos dados. In: CARDOSO, Suzana et al. **Atlas Linguístico do Brasil**. v. 1 (Introdução). Londrina: EDUEL, 2014a, p. 113-133.

THUN, H. La géographie linguistique romane à la fin du XX siècle. In: RAENDONCK, D. V. et all. (Orgs.). **Actes du XXII Congrès International de Linguistique e Philologie Romanes**. Bruxelles, 1998, p. 367-409.

THUN, Harald; FORTE, Carlos; ELIZAINCÍN, Adolfo. **Atlas Lingüístico Diatópico y Diastrático del Uruguay (ADDU)**. Fasc. A.1. Kiel: Westensee, 2000.

THUN, Harold. A dialetologia no Rio da Prata. In: ZILLES, Ana Maria S. (Org.). **Estudos da variação lingüística no Brasil e no Cone Sul**. Porto Alegre: UFRGS, 2005.

THUN, Harald. Pluridimensional Cartography. In: LAMELI, Alfred; KEHREIN, Roland; RABANUS, Christian (eds.). **Language mapping**. Berlin: de Gruyter Mouton, 2010, p. 506523.

TORRENT, T. T. **“O Homem vai botar uma casa para MIM morar” - Uma Abordagem Sociocognitivista e Diacrônica da Construção de Dativo com Infinitivo**. 75 f. 2005. Dissertação (Mestrado em Linguística) – Faculdade de Letras da Universidade Federal de Juiz de Fora, Juiz de Fora, 2005.

TRUDGILL, Petter. Sexo e prestígio linguístico. In: AEBISCHER, Verena; FOREL, Claire. (orgs.) **Falas femininas, falas masculinas?** Tradução de Celene M. Cruz. São Paulo: Brasiliense, 1991, P. 77-101.

VAZ, Júlia Tôres Ferreira. **Uma análise do uso do pronome eu/mim como sujeito em orações infinitivas no contexto de grupos virtuais de jovens brasileiros**. 15f. 2020. Trabalho de Conclusão de Curso (Graduação em Letras). Universidade de Brasília, Salvador. 2020.

VÄÄNÄNEN, Veikko. **Introducción al latín vulgar**. Madrid: Gredos, 1968.

VITORIO, Elyne S.L.A. **A língua usada no sertão alagoano**: constituição da amostra. Trabalho apresentado no III Estudos da Linguagem do Sertão. Delmiro Gouveia, jun. 2017.

VOTRE, S. J. A relevância da variável escolaridade. In: MOLLICA, M, C. & BRAGA, M. L. (orgs.) **Introdução à Sociolinguística**: o tratamento da variação. São Paulo: Contexto, 2007.

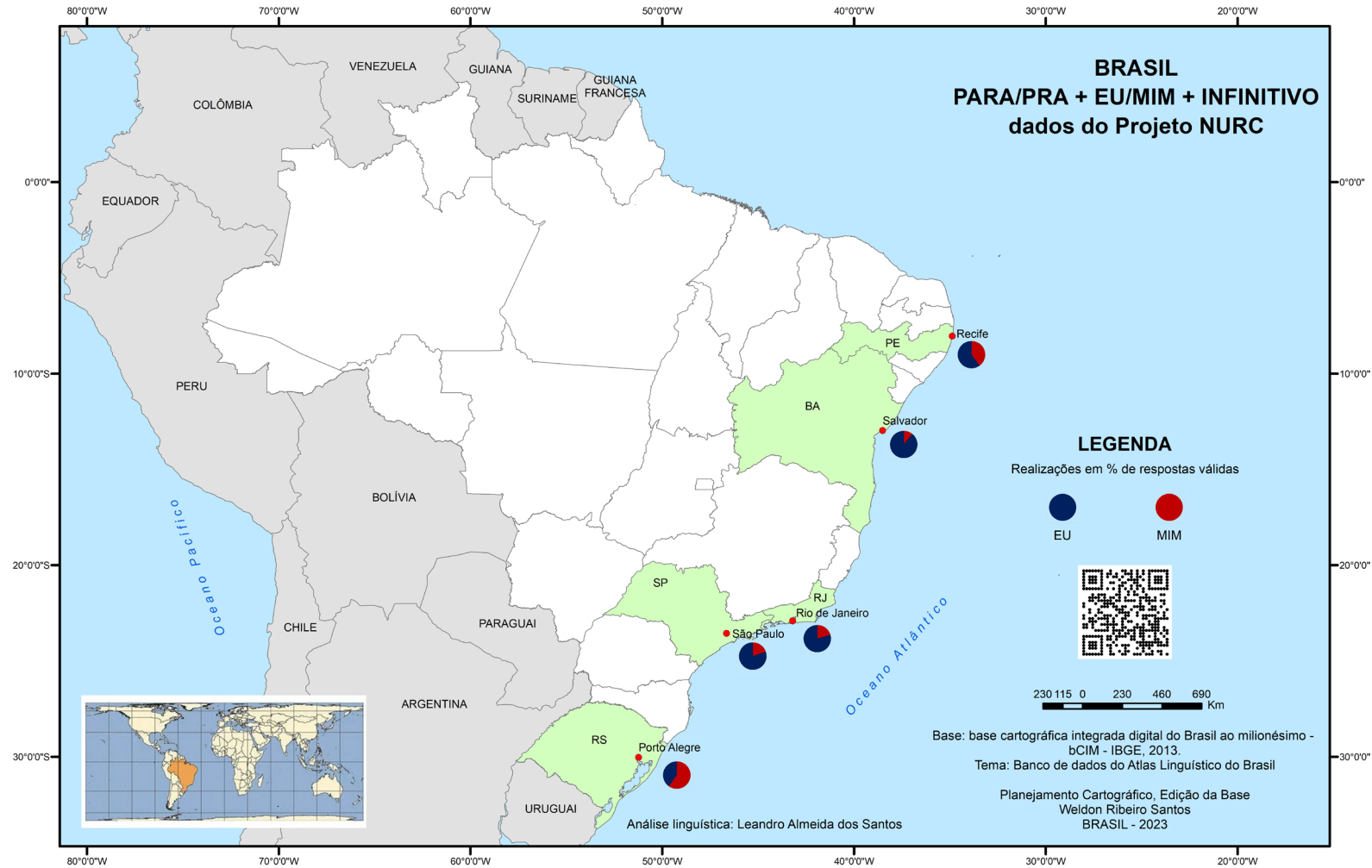
WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. **Fundamentos empíricos para uma teoria da mudança lingüística**. Tradução de Marcos Bagno. São Paulo: Parábola, 2006 [1968].

ZÁGARI, Mário Roberto. Os falares mineiros: esboço de um Atlas Linguístico de Minas Gerais. In: AGUILERA, Vanderci de Andrade (org.). **A geolinguística no Brasil**: caminhos e perspectivas. Londrina: Ed. UEL, 2005 [1998]. p. 31-77.

ZIAMANDANIS, Claire. La falta de respuesta y el desconocimiento en Puerto Rico: silenciar también es contestar. In: MORALES, Amparo; CARDONA, Julia; LÓPEZ MORALES, Humberto; FORASTIERI, Eduardo (eds.). **Estudios de lingüística hispánica**. San Juan, Puerto Rico: Editorial de La Universidad de Puerto Rico, 1999. p. 656-664.

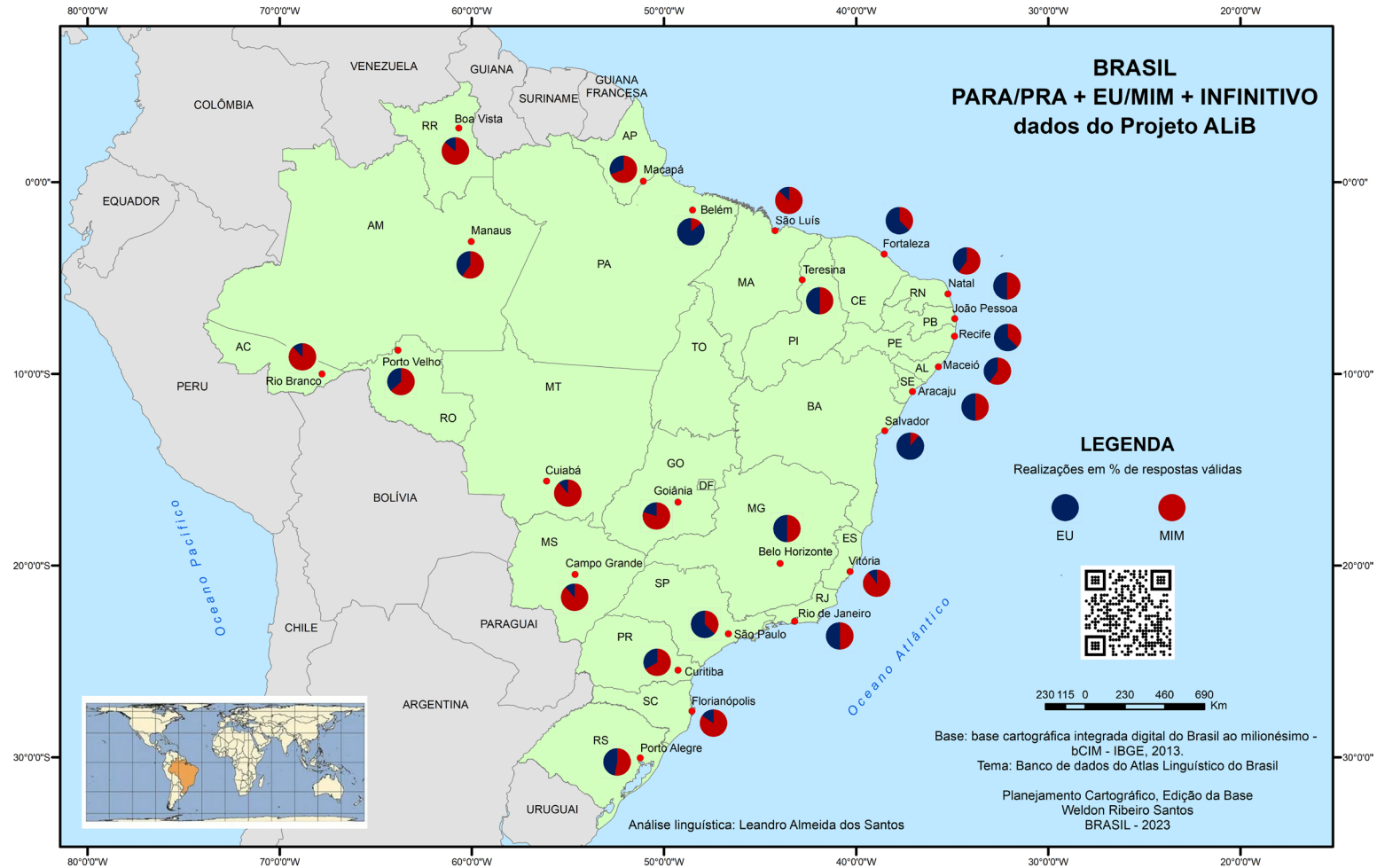
APÊNDICE A

CARTA EXPERIMENTAL M01



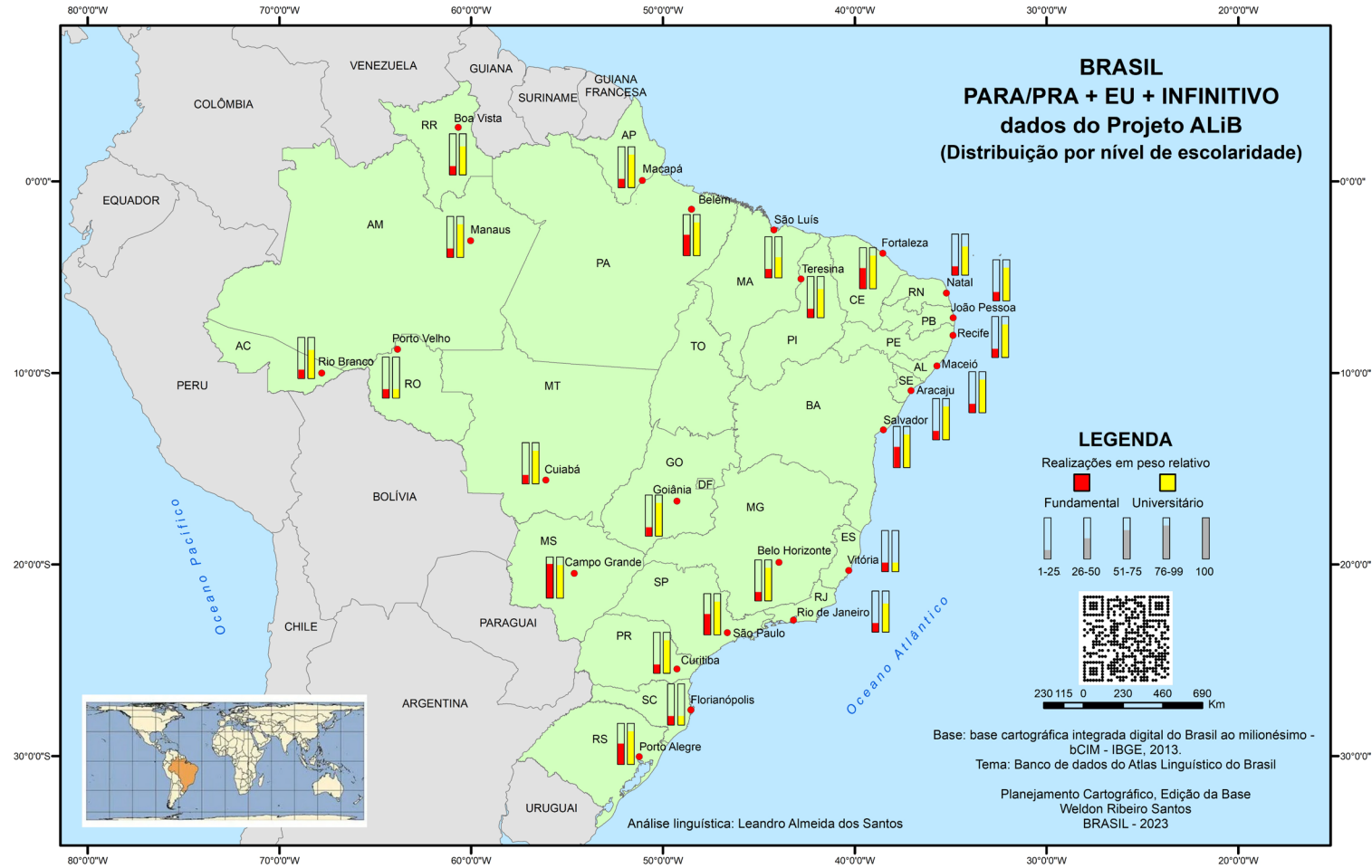
APÊNDICE B

CARTA EXPERIMENTAL M02



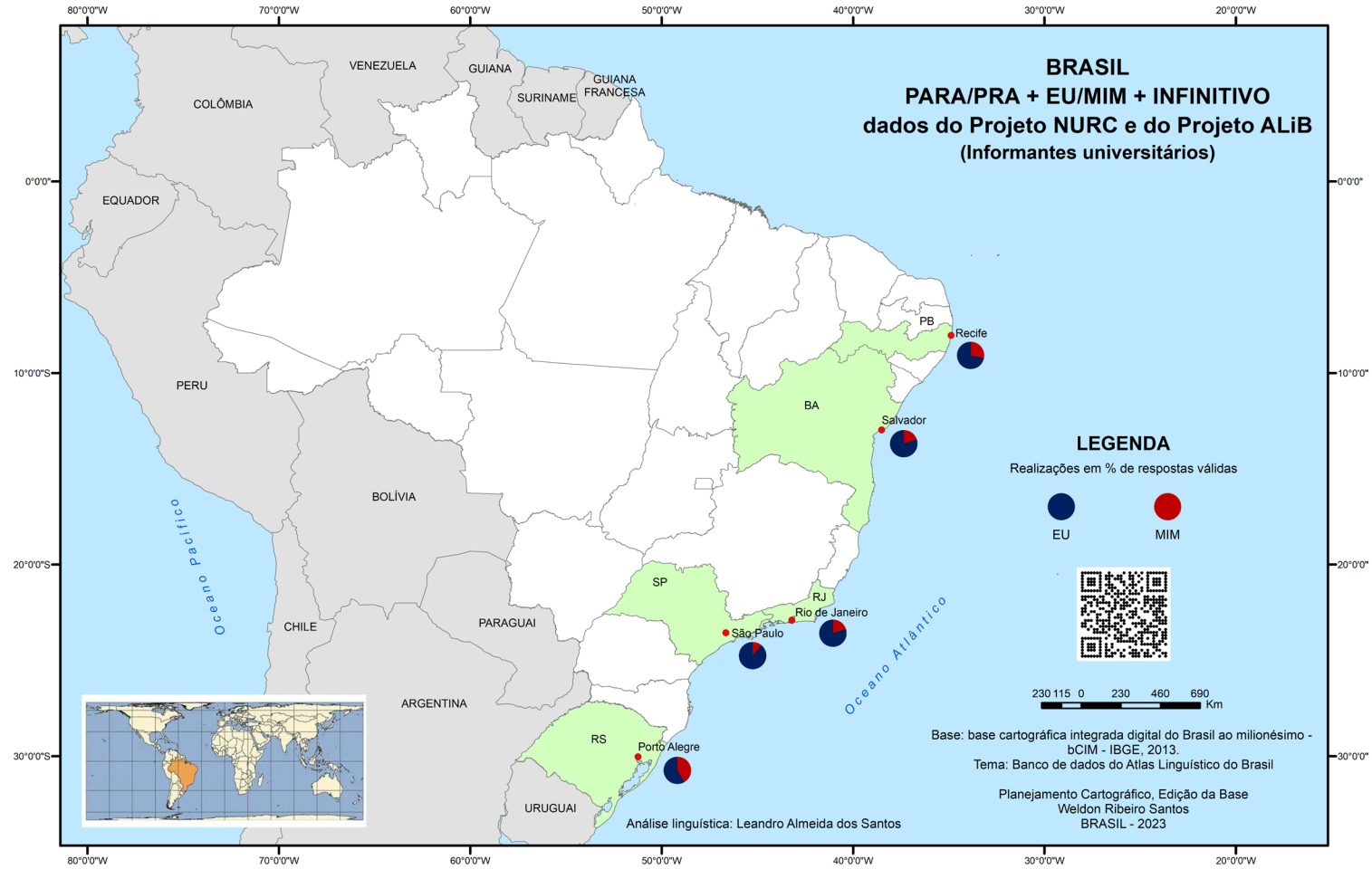
APÊNDICE C

CARTA EXPERIMENTAL M03



APÊNDICE D

CARTA EXPERIMENTAL M04



APÊNDICE E

CARTA EXPERIMENTAL M05

