



UNIVERSIDADE FEDERAL DA BAHIA

Instituto de Letras

Programa de Pós-Graduação em Língua e Cultura



LUDINALVA SANTOS DO AMOR DIVINO

***TU E VOCÊ EM CINCO ESTADOS DO NORDESTE A PARTIR
DOS DADOS DO PROJETO ATLAS LINGUÍSTICO DO
BRASIL: UM ESTUDO VARIACIONISTA***

Salvador
2020

LUDINALVA SANTOS DO AMOR DIVINO

***TU E VOCÊ EM CINCO ESTADOS DO NORDESTE A PARTIR DOS
DADOS DO PROJETO ATLAS LINGUÍSTICO DO BRASIL: UM
ESTUDO VARIACIONISTA***

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

Orientadora: Profa. Dra. Josane Moreira Oliveira

Salvador
2020

FICHA CATALOGRÁFICA

LUDINALVA SANTOS DO AMOR DIVINO

TU E VOCÊ EM CINCO ESTADOS DO NORDESTE A PARTIR DOS DADOS DO PROJETO ATLAS LINGUÍSTICO DO BRASIL: UM ESTUDO VARIACIONISTA

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

Aprovada em: 18 de dezembro de 2020

BANCA EXAMINADORA



Profa. Dra. Josane Moreira Oliveira – Orientadora
Universidade Federal da Bahia – UFBA

Profa. Dra. Norma Lúcia Fernandes de Almeida – Examinadora externa
Universidade Estadual de Feira de Santana – UEFS

Prof. Dr. Clézio Roberto Gonçalves – Examinador externo
Universidade Federal de Ouro Preto – UFOP

Profa. Dra. Marcela Moura Torres Paim – Examinadora interna
Universidade Federal da Bahia – UFBA/Universidade Federal Rural de Pernambuco –
UFRPE

Prof. Dr. Danniel da Silva Carvalho – Examinador interno
Universidade Federal da Bahia – UFBA

À minha mãe, Maria Dinalva Santos do Amor Divino, e à memória do
meu pai, Luiz Bispo do Amor Divino, dedico esta conquista.

AGRADECIMENTOS

A Deus, em primeiro lugar, por ter me dado forças e condições para empreender esta tarefa e finalizá-la. Por se fazer presente nos momentos de pouca inspiração.

A minha família, pelo respeito aos períodos de ausência, em especial a minha mãe, pelas inúmeras orações e pelo incentivo.

A meu pai (*in memoriam*), por ter plantado em mim a garra de lutar pelos sonhos e tornar-me cada vez melhor.

A minha orientadora, Josane Moreira de Oliveira, por ter me aceitado em um momento bastante frágil, pela paciência nos encontros de orientação, pelas contribuições e auxílio, pois não é fácil ser professora em exercício e doutoranda; enfim, por me orientar com sabedoria e humanidade. Quanta admiração!

À professora Suzana Alice Marcelino Cardoso (*in memoriam*), pessoa de suma importância na minha vida acadêmica, pela sua paciência e generosidade, as quais me levaram a mudar de tema e pesquisar algo com que já tinha grande afinidade.

A Natalício, pelo auxílio nas primeiras rodadas com o pacote estatístico.

Agradecimento especial a Neila Maria, pela parceria em congressos, pelo carinho e incentivo na minha vida acadêmica.

A Lucival, pelos ajustes necessários no texto, no período de qualificação.

A Marcelo Marques, pela paciência, torcida e incentivo.

À equipe gestora do CETEP, especialmente a Juvenal Barreto, pela compreensão e apoio não somente ao meu trabalho enquanto regente bem como ao doutorado.

À secretaria da Pós-Graduação, Cristiane, sempre tão prestativa, humana e solícita com as questões administrativas referentes à Pós-Graduação.

Ao Projeto ALiB, que forneceu um rico material para análise, pelo dedicado trabalho de coleta e organização dos dados.

Às professoras Norma Almeida e Marcela Paim, por aceitarem participar da banca de qualificação, além da contribuição inestimável que deram nas respectivas análises e arguições.

A Shirley, pela colaboração, sugestões e revisão de dois capítulos.

Às colegas Patrícia e Eloísa, pelas conversas e parcerias no período de cumprimento dos créditos.

Aos colegas Givaldo e Geisa, por serem corteses em disponibilizar seus projetos.

A Fabrício, pelo incentivo até o doutorado e pela amizade, sempre presente, mesmo em momentos de distância física.

À professora Ana Regina Teles (*in memoriam*), pela boa vontade e perfeição na execução das cartas linguísticas.

A Evany, companheira de lutas, pelo carinho e energia positiva direcionada.

À aluna Isabela, pela grande ajuda na audição de inquéritos e seleção preliminar dos dados.

A Joana, por ter com quem partilhar as tensões de ser doutoranda, fazendo-me mais perseverante.

Às meninas do Projeto ALiB-BA, Ana Rita e Greice, pela paciência nas pendências de inquéritos e disponibilidade de tempo para me auxiliarem no fornecimento de informações sobre o Projeto ALiB sempre que se fez necessário.

À colega Sandra Prudêncio, pelo manuseio para arquivamento dos inquéritos da Região Nordeste e ajuda na recuperação de inquéritos com problemas de audição.

A professora Célia Telles, pela compreensão em virtude das limitações de horários no final da aula, por conta do retorno para casa, exatamente 320Km a serem percorridos.

A Luizinho, por me auxiliar nos ajustes no computador quando apresentava problemas técnicos que, por vezes, surgiam e exigiam resoluções rápidas.

A Cláudia, pelo carinho e pela confecção das tabelas e gráficos iniciais.

A todos os colegas e professores do Programa de Pós-Graduação em Língua e Cultura, pela convivência e estímulo constante.

Enfim, a todos que contribuíram de alguma forma para que este sonho pudesse ser realizado.

Meu muitíssimo obrigada!

*“Os que confiam no Senhor serão como o monte Sião,
que não se abala, mas permanece para sempre”.*

(Salmos125:1)

RESUMO

Dada a necessidade de ampliar os estudos dialetais e sociolinguísticos sobre a realidade linguística do Nordeste brasileiro, procuramos fazer uma fotografia geossociolinguística do português falado em cinco Estados dessa região: Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e Sergipe, no que concerne ao uso das formas de tratamento *tu* e *você*. Nosso objetivo geral é investigar o uso das formas de tratamento *tu* e *você* nos Estados mencionados, a partir da análise de um extrato do *corpus* do ALiB, enfocando todos os questionários (questões de prosódia, questionário semântico-lexical, questionário morfossintático, temas para discurso semidirigidos e questões metalinguísticas), exceto o Questionário Fonético-Fonológico. Como objetivos específicos, visamos a: (i) analisar a relevância das variáveis sociais – localidade, faixa etária, sexo, escolaridade – no comportamento dos falantes com relação ao uso das formas *tu* e *você*; (ii) analisar a relevância das variáveis linguísticas – parte do inquérito, tipo de referência, função sintática da variante, tempo verbal e paralelismo – na ocorrência das formas *tu* e *você*; (iii) representar, em cartas linguísticas, o uso das formas em estudo, com vistas a delinear a realidade dialetal nos Estados nordestinos analisados; (iv) observar como as formas *tu* e *você* estão distribuídas, em cada localidade, entre os informantes. A nossa amostra conta com a análise das falas de 140 informantes, compreendendo as capitais dos cinco Estados e 25 cidades do interior, totalizando 30 localidades. Os informantes foram distribuídos entre ambos os sexos, duas faixas etárias (18 a 30 anos e 50 a 65 anos) e dois graus de escolaridade (até o ensino fundamental II e universitário). O nível universitário é controlado somente nas capitais. A hipótese central do trabalho pauta-se no pressuposto de que os falantes da região Nordeste que residem no interior utilizam, preferencialmente, o pronome *tu* e os moradores das capitais dão preferência à utilização do *você* nas suas interlocuções. O estudo foi feito segundo os pressupostos teóricos da Dialetologia Pluridimensional Contemporânea e da Sociolinguística Variacionista, que prioriza os princípios gerais para o estudo da mudança linguística descritos por Weinreich, Labov e Herzog (2006 [1968]), ou seja, a variação é inerente à língua, é sistemática e condicionada por fatores intra- e extralinguísticos. A análise dos dados e os resultados de cálculos estatísticos foram feitos com o auxílio do pacote GoldVarb X. No tocante à interpretação dos dados, realizaram-se duas análises: uma geral, incluindo todos os Estados estudados e outra de cada Estado, separadamente. Na análise geral, obtivemos um total de 1.995 ocorrências de pronomes de segunda pessoa, sendo 1.769 contextos com o emprego de *você* (88,7%) e 226 usos de *tu* (11,3%). Mesmo com um número reduzido do pronome *tu* na amostra investigada, focamos a nossa análise em elucidar quais os contextos em que esse pronome ainda perdura. A rodada binária do GoldVarb X apontou como variáveis relevantes no emprego de *tu*: referênciação, parte do inquérito, localidade e escolaridade. Na descrição da análise por Estado, as variáveis selecionadas como significativas foram: em Pernambuco – referênciação, parte do inquérito, faixa etária e localidade; na Paraíba – parte do inquérito, faixa etária e localidade; em Alagoas – sexo e tipo de questionário; no Rio Grande do Norte e em Sergipe não houve variáveis selecionadas.

Palavras-chave: Formas de Tratamento; Pronomes; Variação Linguística; Dialetologia; Sociolinguística.

ABSTRACT

Given the need to expand dialectal and sociolinguistics studies on the linguistic reality of Northeast Brazilian, we tried to take a geo-sociolinguistic photograph of the Portuguese spoken in five States in that region: Rio Grande do Norte, Paraíba, Pernambuco, Alagoas and Sergipe in relation the use of these treatment forms *tu* and *você*. Our general objective is to investigate these uses at the mentioned states from the analysis of extract of ALiB corpus, focusing on all questionnaires (prosody questions, semantic-lexical questionnaire, morphosyntactic questionnaire, topics for semi-directed speech and metalinguistic questions), except the Phonetic-Phonological Questionnaire. As specific objectives, we aim do it at: (i) to analyze the relevance of social variables – location, age, gender, education – in the behavior of speakers about the use of *tu* and *você* forms; (ii) to analyze the relevance of language variation – part of this inquiry, type of reference, syntactic function of the variant, verbal time and parallelism – in the occurrence of the *tu* and *você* forms; (iii) to represent the use of the forms under study in order to delineate the dialectal reality in the analyzed northeastern states; (iv) to observe how the *tu* and *você* forms are distributed, in each location, among informants. Our sample includes date analysis of 140 informants, it achieves the capitals of five states and 25 towns, they total 30 places. The informants were distributed between both sexes, two age groups (18 to 30 years old, and 50 to 65 years old) and two levels of education (up to elementary school II and university). The university level is controlled only in the capitals. The central hypothesis of the work is based on the assumption about the speakers from Northeast region that live in the downtown prefer the pronoun *tu*, and residents of the capitals prefer the use of *você*. The study was carried out according to the theoretical assumptions of Contemporary multidimensional Dialectology and Variationist Sociolinguistics, which prioritizes the general principles for the study of language change described by Weinreich, Labov and Herzog (2006 [1968]), that is the variation is inherent to the language, it is systematic and it's conditioned by interlinguistic and extralinguistic factors. The analysis of the data and the results of statistical calculations were carried out with the aid of the GoldVarb X package. Regarding the interpretation of the data, two analyzes were carried out: a general one, including all the states studied and the other from each state separately. In the general analysis, we obtained a total of 1,995 occurrences of second-person pronouns, being 1,769 contexts with the use of *você* (88.7%) and 226 uses of *tu* (11.3%). Even with a small number of *tu* pronoun in the sample investigated, we focus our analysis on clarify the contexts in which this pronoun still persists. The GoldVarb X binary round pointed out as relevant variables in the use of *tu*: referencing, part of the inquiry, location and schooling. In the description of the analysis by state, the variables selected as significant were: in Pernambuco – in referencing, part of the inquiry, age group and location; in Paraíba - part of the inquiry, age group and location; in Alagoas – gender and type of questionnaire; in Rio Grande do Norte and Sergipe there were no variables selected.

Keywords: Ways of Treatment; Pronouns; Linguistic Variation; Dialectology; Sociolinguistic.

LISTA DE ABREVIATURAS E SIGLAS

| | |
|-------------------|---|
| AFEBG | Atlas Fonético do Entorno da Baía de Guanabara |
| ALAC | Atlas Etnolinguístico do Acre |
| ALAM | Atlas Linguístico do Amazonas |
| ALAP | Atlas Linguístico do Estado do Amapá |
| ALECE | Atlas Linguístico do Estado do Ceará |
| ALEPG | Atlas linguístico-Etnográfico de Portugal e da Galiza |
| ALERS | Atlas Linguístico-Etnográfico da Região Sul do Brasil |
| ALESP | Atlas Linguístico do Estado de São Paulo |
| ALF | Atlas Linguistique de la France |
| ALiB | Atlas Linguístico do Brasil |
| ALiMA | Atlas Linguístico do Maranhão |
| ALiMAT | Atlas Linguístico do Mato Grosso |
| ALiPA | Atlas Geo-Sociolinguístico do Pará |
| ALiPI | Atlas Linguístico do Piauí |
| ALiR | Atlas Linguistique Roman |
| ALiRN | Atlas Linguístico do Rio Grande do Norte |
| ALiSon-Rio | Atlas Linguístico Sonoro do Estado do Rio de Janeiro |
| ALISPA | Atlas Linguístico Sonoro do Pará |
| ALMS | Atlas Linguístico do Mato Grosso do Sul |
| ALPB | Atlas Linguístico da Paraíba |
| ALPR | Atlas Linguístico do Paraná |
| ALS | Atlas Linguístico de Sergipe |
| APES | Atlas Linguístico do Espírito Santo |
| APFB | Atlas Prévio dos Falares Baianos |
| AR | Atlas Regionais |
| BDI | Banco de Dados Interacionais |
| Carta M02 | Carta Linguística Morfossintática nº 2 |
| CIDS | Congresso Internacional de Dialetologia e Sociolinguística |
| CODEVASF | Companhia de Desenvolvimento dos Vales do São Francisco e do Parnaíba |
| COFECUB | Comitê Francês de Avaliação da Cooperação Universitária com o Brasil |
| D2 | Diálogo entre dois informantes |

| | |
|-------------------|---|
| DID | Diálogo entre Informante e Documentador |
| EALMG | Esboço de um Atlas Linguístico de Minas Gerais |
| EDUEL | Editora da Universidade Estadual de Londrina |
| EF | Elocução Formal |
| GNs | Gramáticas Normativas |
| HARAS | Homem, Adulto, Rurícola, Analfabeto e Sedentário |
| IBGE | Instituto Brasileiro de Geografia e Estatística |
| IDEMA | Instituto de Desenvolvimento Sustentável e Meio Ambiente |
| INF. | Informante |
| INQ. | Inquiridor |
| IPHAN | Instituto do Patrimônio Histórico e Artístico Nacional |
| LE | Texto para Leitura |
| MicroAFERJ | Micro Atlas Fonético do Estado do Rio de Janeiro |
| MS-DOS | Microsoft Disk Operating System |
| NORPOFOR | Norma Oral do Português Popular de Fortaleza |
| NURC | Norma Linguística Urbana Culta |
| NURC/SSA | Projeto Norma Linguística Urbana Culta de Salvador |
| PEPP | Português Popular em Salvador |
| PEUL | Projeto de Estudos sobre os Usos da Língua |
| PIB | Produto Interno Bruto |
| PB | Português brasileiro |
| PCNs | Parâmetros Curriculares Nacionais |
| PM | Perguntas Metalingüísticas |
| PTSF | Ponto de Transposição do rio São Francisco |
| QFF | Questionário fonético-fonológico |
| QMS | Questionário morfossintático |
| QP | Questões de prosódia |
| QSL | Questionário semântico-lexical |
| RIDE | Região Integrada de Desenvolvimento |
| SIG | Sistemas de Informação Geográfica |
| SPSS | Pacote Estatístico para as Ciências Sociais (Statistical Package for the Social Sciences) |
| TDS | Temas para Discurso Semidirigido |

| | |
|-----------------|--|
| UEL | Universidade Estadual de Londrina |
| UEPA | Universidade do Estado do Pará |
| UFBA | Universidade Federal da Bahia |
| UFERSA | Universidade Federal Rural do Semiárido |
| UFES | Universidade Federal do Espírito Santo |
| UFPA | Universidade Federal do Pará |
| UFRGS | Universidade Federal do Rio Grande do Sul |
| UFRJ | Universidade Federal do Rio de Janeiro |
| UFSC | Universidade Federal de Santa Catarina |
| UNB | Universidade de Brasília |
| UNESCO | Organização das Nações Unidas para a Educação, a Ciência e a Cultura (United Nations Educational, Scientific and Cultural Organization) |
| UNIFAP | Universidade Federal do Amapá |
| VALEXTRA | Variação lexical: teorias, recursos e aplicações |
| VARBRUL | Regras Variáveis (Variable Rules) |

LISTA DE FIGURAS E CARTAS

| | | |
|------------|--|-----|
| Figura 1: | Carta do uso de <i>tu</i> e <i>você</i> nas capitais brasileiras | 36 |
| Figura 2: | Ocorrência dos pronomes <i>tu/você</i> em Imperatriz (MA) e Uberlândia (MG) | 46 |
| Figura 3: | Totais de referência à segunda pessoa nos <i>corpora</i> analisados | 49 |
| Figura 4: | Distribuição dos pontos da rede do Nordeste | 134 |
| Figura 5: | Localidades do Rio Grande do Norte que fazem parte da rede de pontos do ALiB | 136 |
| Figura 6: | Localidades da Paraíba que fazem parte da rede de pontos do ALiB | 141 |
| Figura 7: | Localidades de Pernambuco que fazem parte da rede de pontos do ALiB | 147 |
| Figura 8: | Localidades de Alagoas que fazem parte da rede de pontos do ALiB | 157 |
| Figura 9: | Localidades de Sergipe que fazem parte da rede de pontos do ALiB | 161 |
| Figura 10: | Perfil dos informantes do Projeto ALiB | 167 |
| Figura 11: | Divisão dialetal proposta por Nascentes | 181 |
| Figura 12: | Distribuição geral do <i>tu</i> e <i>você</i> nos Estados do Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e Sergipe | 194 |
| Figura 13: | Distribuição da variante <i>tu</i> por faixa etária em Pernambuco | 211 |
| Figura 14: | Carta linguística 1: Distribuição diatópica de ocorrência de <i>tu</i> em Pernambuco | 213 |
| Figura 15: | Distribuição da variante <i>tu</i> por faixa etária na Paraíba | 217 |
| Figura 16: | Carta linguística 2: Distribuição diatópica de ocorrência de <i>tu</i> na Paraíba | 219 |
| Figura 17: | Carta M02: Tratamento do interlocutor (com <i>tu</i> ou <i>você</i>) nas capitais | 226 |

LISTA DE QUADROS

| | | |
|------------|--|-----|
| Quadro 1: | Pronomes-sujeito em Cunha e Cintra (2007) | 33 |
| Quadro 2: | Formas de tratamento segundo alguns pesquisadores | 34 |
| Quadro 3: | Estudos sobre a 2 ^a pessoa: distribuição por região | 40 |
| Quadro 4: | Diferenças semântico-pragmáticas entre <i>tu</i> e <i>você</i> | 67 |
| Quadro 5: | Síntese da frequência de uso das variantes <i>tu</i> e <i>você</i> nas regiões do Brasil | 70 |
| Quadro 6: | Atlas que já foram concluídos/publicados e Atlas que estão em andamento | 115 |
| Quadro 7: | Regionais do Projeto ALiB | 128 |
| Quadro 8: | Localidades da Amostra | 135 |
| Quadro 9: | Áreas semânticas do QSL | 165 |
| Quadro 10: | Quantidade de informantes por Estado | 168 |
| Quadro 11: | Comissão de informatização e cartografia | 190 |
| Quadro 12: | Total de ocorrências das variantes: <i>tu</i> e <i>você</i> por Estado | 231 |

LISTA DE TABELAS

| | | |
|------------|---|-----|
| Tabela 1: | Distribuição dos falantes de acordo com o uso dos pronomes | 41 |
| Tabela 2: | Totais de referência à segunda pessoa | 41 |
| Tabela 3: | Emprego das formas de tratamento <i>tu</i> e <i>você</i> em Sette (1980) | 43 |
| Tabela 4: | Efeito do gênero sobre o uso do <i>tu</i> em Bezerra (1994) | 45 |
| Tabela 5: | Índice geral de uso de <i>tu</i> e <i>você</i> em cartas particulares escritas em Feira de Santana (décadas de 1940, 1950 e 1960) | 48 |
| Tabela 6: | Variação entre <i>tu</i> e <i>você</i> de acordo com a década e o sexo | 48 |
| Tabela 7: | Atuação do tipo de fala sobre o pronome <i>tu</i> (<i>tu</i> x <i>você</i>) | 52 |
| Tabela 8: | Tipo de relação entre os interlocutores (dados da Vila) | 56 |
| Tabela 9: | Variação <i>você/cê/tu</i> em relação à função sintática | 57 |
| Tabela 10: | Distribuição dos pronomes de 2 ^a pessoa nos <i>corpora</i> investigados | 62 |
| Tabela 11: | Função sintática das formas <i>tu/você</i> em Santos-SP | 63 |
| Tabela 12: | Distribuição de <i>tu/você</i> por sexo e faixa etária nas capitais do Sul do Brasil | 68 |
| Tabela 13: | Atuação da faixa etária sobre a variante <i>tu</i> | 69 |
| Tabela 14: | Distribuição diatópica dos pronomes <i>tu</i> e <i>você</i> em capitais brasileiras .. | 195 |
| Tabela 15: | Frequência absoluta das formas de tratamento | 195 |
| Tabela 16: | A variante <i>tu</i> em função da Referenciação nos cinco Estados nordestinos analisados | 197 |
| Tabela 17: | Uso do <i>tu</i> pelo tipo de referência | 198 |
| Tabela 18: | A variante <i>tu</i> em função do Tipo do questionário nos cinco Estados nordestinos analisados | 199 |
| Tabela 19: | Resultado da variante <i>tu</i> em função da Localidade nos cinco Estados analisados | 203 |
| Tabela 20: | A variante <i>tu</i> em função da Escolaridade do informante nos cinco Estados nordestinos analisados | 205 |
| Tabela 21: | A variante <i>tu</i> em função da Referenciação no Estado de Pernambuco ... | 207 |
| Tabela 22: | A variante <i>tu</i> em função do Tipo do questionário no Estado de Pernambuco | 209 |
| Tabela 23: | A variante <i>tu</i> em função da Faixa etária do informante em Pernambuco | 210 |
| Tabela 24: | A variante <i>tu</i> em função da Localidade em Pernambuco | 212 |

| | | |
|------------|--|-----|
| Tabela 25: | A variante <i>tu</i> em função do Tipo do questionário no Estado da Paraíba | 215 |
| Tabela 26: | A variante <i>tu</i> em função da Faixa etária do informante na Paraíba | 216 |
| Tabela 27: | A variante <i>tu</i> em função da Localidade na Paraíba | 218 |
| Tabela 28: | A variante <i>tu</i> em função do Sexo do informante em Alagoas | 222 |
| Tabela 29: | A variante <i>tu</i> em função do Tipo do questionário no Estado de Alagoas | 223 |
| Tabela 30: | Distribuição dos pronomes <i>você</i> e <i>tu</i> no Rio Grande do Norte | 225 |
| Tabela 31: | Distribuição dos pronomes <i>você</i> e <i>tu</i> por Referenciação no Rio Grande do Norte | 227 |
| Tabela 32: | Distribuição dos pronomes <i>tu</i> e <i>você</i> em Sergipe | 228 |
| Tabela 33: | Distribuição dos pronomes <i>tu</i> e <i>você</i> com relação à Referenciação em Sergipe | 229 |
| Tabela 34: | Distribuição dos pronomes <i>tu</i> e <i>você</i> por Sexo do informante em Sergipe | 230 |

SUMÁRIO

| | |
|---|----|
| INTRODUÇÃO | 22 |
| 1 FORMAS DE TRATAMENTO | 28 |
| 1.1 FORMAS DE TRATAMENTO: UM POUCO DE HISTÓRIA | 28 |
| 1.2 FORMAS DE TRATAMENTO NA VISÃO DE ALGUNS AUTORES | 31 |
| 1.3 PRONOMES DE TRATAMENTO NO BRASIL | 35 |
| 1.3.1 De Vossa Mercê a você: uma abordagem diacrônica | 37 |
| 1.4 TRABALHOS REALIZADOS SOBRE O <i>TU</i> E O <i>VOCÊ</i> DE NATUREZA <i>DIALETOLÓGICA E SOCIOLINGUÍSTICA</i> | 39 |
| 1.4.1 Região Norte | 40 |
| 1.4.1.1 <i>Martins (2010)</i> | 40 |
| 1.4.2 Região Nordeste | 42 |
| 1.4.2.1 <i>Sette (1980)</i> | 42 |
| 1.4.2.2 <i>Soares (1980)</i> | 44 |
| 1.4.2.3 <i>Bezerra (1994)</i> | 44 |
| 1.4.2.4 <i>Herênio (2006)</i> | 45 |
| 1.4.2.5 <i>Almeida (2012)</i> | 46 |
| 1.4.2.6 <i>Nogueira (2013)</i> | 49 |
| 1.4.2.7 <i>Guimarães (2014)</i> | 51 |
| 1.4.3 Região Centro-Oeste | 53 |
| 1.4.3.1 <i>Lucca (2005)</i> | 53 |
| 1.4.3.2 <i>Dias (2007)</i> | 54 |
| 1.4.3.3 <i>Andrade (2010)</i> | 56 |
| 1.4.3.4 <i>Scherre et al. (2015)</i> | 57 |
| 1.4.4 Região Sudeste | 60 |
| 1.4.4.1 <i>Salles (2001)</i> | 60 |
| 1.4.4.2 <i>Paredes Silva (2003)</i> | 61 |
| 1.4.4.3 <i>Modesto (2006)</i> | 62 |
| 1.4.4.4 <i>Mota (2008)</i> | 64 |
| 1.4.4.5 <i>Silva (2017)</i> | 65 |
| 1.4.5 Região Sul | 66 |
| 1.4.5.1 <i>Ramos (1989)</i> | 66 |
| 1.4.5.2 <i>Loregian-Penkal (2004)</i> | 67 |
| 1.4.5.3 <i>Franceschini (2010)</i> | 69 |
| 1.5 SÍNTESE DO CAPÍTULO | 69 |
| 2 FUNDAMENTAÇÃO TEÓRICA | 72 |
| 2.1 SOCIOLINGUÍSTICA | 72 |
| 2.1.1 Etapas da pesquisa sociolinguística | 72 |
| 2.1.2 Variação linguística | 76 |
| 2.1.2.1 <i>Tipos de variação</i> | 78 |
| 2.1.2.1.1 Variação diatópica | 79 |
| 2.1.2.1.2 Variação diastrática ou social | 79 |

| | |
|---|-----|
| 2.1.2.1.3 Variação diafásica ou estilística | 81 |
| 2.1.2.1.3.1 Estilo em Labov | 82 |
| 2.1.2.1.3.2 Estilo em Allan Bell | 84 |
| 2.1.2.1.3.3 Um ponto de congruência | 86 |
| 2.1.2.1.3.4 Monitoramento | 87 |
| 2.1.2.1.4 Variação entre fala e escrita ou diamésica | 88 |
| 2.1.2.2 <i>Variável linguística e variantes</i> | 89 |
| 2.1.2.3 <i>Mudança linguística</i> | 90 |
| 2.2 DIALETOLOGIA | 94 |
| 2.2.1 Dialetos | 97 |
| 2.2.2 Sobre a dialetação do português brasileiro: as propostas de divisão dialetal | 98 |
| 2.2.3 Geolinguística: geografia linguística no Brasil | 102 |
| 2.2.4 Os atlas brasileiros | 103 |
| 2.2.4.1 <i>Atlas em andamento (incluindo teses)</i> | 110 |
| 2.3 DIALETOLOGIA E SOCIOLINGUÍSTICA: UM DIÁLOGO POSSÍVEL | 116 |
| 2.4 DA DIALETOLOGIA À SOCIOLINGUÍSTICA: CONTRIBUIÇÕES PARA O ENSINO DE LÍNGUA PORTUGUESA | 120 |
| 2.5 SÍNTESE DO CAPÍTULO | 122 |
| 3 ASPECTOS METODOLÓGICOS | 123 |
| 3.1 PROJETO ALiB: ANTECEDENTES HISTÓRICOS | 124 |
| 3.2 CONSTRUÇÃO DO PROJETO ATLAS LINGUÍSTICO DO BRASIL | 126 |
| 3.3 O <i>CORPUS</i> | 131 |
| 3.3.1 Divisão territorial do Nordeste | 131 |
| 3.3.2 Localidades que perfazem a amostra | 135 |
| 3.3.2.1 <i>Rio Grande do Norte</i> | 135 |
| 3.3.2.1.1 Mossoró (Ponto 51 do ALiB) | 136 |
| 3.3.2.1.2 Angicos (Ponto 52 do ALiB) | 137 |
| 3.3.2.1.3 Natal (Ponto 53 do ALiB) | 138 |
| 3.3.2.1.4 Pau dos Ferros (Ponto 54 do ALiB) | 139 |
| 3.3.2.1.5 Caicó (Ponto 55 do ALiB) | 140 |
| 3.3.2.2 <i>Paraíba</i> | 141 |
| 3.3.2.2.1 Cuité (Ponto 56 do ALiB) | 142 |
| 3.3.2.2.2 Cajazeiras (Ponto 57 do ALiB) | 143 |
| 3.3.2.2.3 Itaporanga (Ponto 58 do ALiB) | 143 |
| 3.3.2.2.4 Patos (Ponto 59 do ALiB) | 144 |
| 3.3.2.2.5 Campina Grande (Ponto 60 do ALiB) | 145 |
| 3.3.2.2.6 João Pessoa (Ponto 61 do ALiB) | 146 |
| 3.3.2.3 <i>Pernambuco</i> | 147 |
| 3.3.2.3.1 Exu (Ponto 62 do ALiB) | 147 |
| 3.3.2.3.2 Salgueiro (Ponto 63 do ALiB) | 148 |
| 3.3.2.3.3 Limoeiro (Ponto 64 do ALiB) | 149 |
| 3.3.2.3.4 Olinda (Ponto 65 do ALiB) | 149 |
| 3.3.2.3.5 Afrânio (Ponto 66 do ALiB) | 150 |

| | |
|---|------------|
| 3.3.2.3.6 Cabrobó (Ponto 67 do ALiB) | 151 |
| 3.3.2.3.7 Arcoverde (Ponto 68 do ALiB) | 152 |
| 3.3.2.3.8 Caruaru (Ponto 69 do ALiB) | 153 |
| 3.3.2.3.9 Recife (Ponto 70 do ALiB) | 153 |
| 3.3.2.3.10 Floresta (Ponto 71 do ALiB) | 154 |
| 3.3.2.3.11 Garanhuns (Ponto 72 do ALiB) | 155 |
| 3.3.2.3.12 Petrolina (Ponto 73 do ALiB) | 156 |
| 3.3.2.4 Alagoas | 157 |
| 3.3.2.4.1 União dos Palmares (Ponto 74 do ALiB) | 157 |
| 3.3.2.4.2 Santana do Ipanema (Ponto 75 do ALiB) | 158 |
| 3.3.2.4.3 Arapiraca (Ponto 76 do ALiB) | 159 |
| 3.3.2.4.4 Maceió (Ponto 77 do ALiB) | 159 |
| 3.3.2.5 Sergipe | 160 |
| 3.3.2.5.1 Propriá (Ponto 78 do ALiB) | 161 |
| 3.3.2.5.2 Aracaju (Ponto 79 do ALiB) | 162 |
| 3.3.2.5.3 Estância (Ponto 80 do ALiB) | 162 |
| 3.4 QUESTIONÁRIOS | 163 |
| 3.5 INFORMANTES | 166 |
| 3.6 RECOLHA E LEVANTAMENTO DE DADOS | 168 |
| 3.6.1 Pesquisa de campo do ALiB | 168 |
| 3.6.2 Constituição do <i>corpus</i> | 170 |
| 3.7 AS VARIÁVEIS CONSIDERADAS | 171 |
| 3.7.1 Variável dependente | 171 |
| 3.7.2 Variáveis independentes | 172 |
| 3.7.2.1 <i>Fatores linguístico-estruturais</i> | 173 |
| 3.7.2.1.1 Referenciação | 173 |
| 3.7.2.1.2 Parte do inquérito | 174 |
| 3.7.2.1.3 Função sintática do pronome | 176 |
| 3.7.2.1.4 Tempo verbal | 177 |
| 3.7.2.1.5 Paralelismo | 178 |
| 3.7.2.2 <i>Fatores extralingüísticos</i> | 179 |
| 3.7.2.2.1 Localidade | 180 |
| 3.7.2.2.2 Sexo | 181 |
| 3.7.2.2.3 Faixa etária | 182 |
| 3.7.2.2.4 Escolaridade | 183 |
| 3.8 PROCEDIMENTOS PARA A ANÁLISE DOS DADOS | 183 |
| 3.8.1 Codificação dos dados | 183 |
| 3.8.2 Aspectos da análise estatística | 184 |
| 3.8.3 Identificação de <i>nocautes</i> | 186 |
| 3.8.4 Peso relativo | 187 |
| 3.9 CARTOGRAFIA DOS DADOS | 189 |
| 3.10 SÍNTESE DO CAPÍTULO | 191 |
| 4 ANÁLISE DOS DADOS | 193 |

| | |
|--|------------|
| 4.1 DISTRIBUIÇÃO GERAL DAS VARIANTES NA AMOSTRA | 193 |
| 4.1.1 Variáveis selecionadas | 196 |
| 4.1.1.1 <i>Referenciação</i> | 197 |
| 4.1.1.2 <i>Parte do inquérito</i> | 199 |
| 4.1.1.3 <i>Localidade</i> | 202 |
| 4.1.1.4 <i>Escolaridade</i> | 205 |
| 4.2 DESCRIÇÃO DAS RODADAS POR ESTADO | 206 |
| 4.2.1 Pernambuco | 207 |
| 4.2.1.1 <i>Referenciação</i> | 207 |
| 4.2.1.2 <i>Parte do inquérito</i> | 208 |
| 4.2.1.3 <i>Faixa etária</i> | 210 |
| 4.2.1.4 <i>Localidade</i> | 211 |
| 4.2.2 Paraíba | 214 |
| 4.2.2.1 <i>Parte do inquérito</i> | 214 |
| 4.2.2.2 <i>Faixa etária</i> | 216 |
| 4.2.2.3 <i>Localidade</i> | 217 |
| 4.2.3 Alagoas | 221 |
| 4.2.3.1 <i>Sexo</i> | 222 |
| 4.2.3.2 <i>Parte do inquérito</i> | 223 |
| 4.2.4 Rio Grande do Norte | 224 |
| 4.2.4.1 <i>Referenciação</i> | 226 |
| 4.2.5 Sergipe | 228 |
| 4.2.5.1 <i>Referenciação</i> | 229 |
| 4.2.5.2 <i>Sexo</i> | 230 |
| 4.3 SÍNTESE DO CAPÍTULO | 231 |
| CONSIDERAÇÕES FINAIS | 233 |
| REFERÊNCIAS | 239 |
| APÊNDICE – CHAVE DE CODIFICAÇÃO..... | 249 |
| ANEXO A – PANORAMA DA COMPOSIÇÃO DOS ATLAS LINGUÍSTICOS NO BRASIL | 251 |
| ANEXO B – DOCUMENTOS PROJETO ALiB | 252 |
| ANEXO C – CDS <i>DESCREVENDO A LÍNGUA, FORMANDO JOVENS PESQUISADORES</i> | 253 |
| ANEXO D – AUTORIZAÇÃO DO COMITÊ NACIONAL PARA USO DO CORPUS DO PROJETO ALiB | 254 |

INTRODUÇÃO

“No Brasil, ...é preciso, antes de mais nada, criar mentalidade dialetológica, preparando um ambiente favorável às pesquisas de campo”. (SILVA NETO, 1957, p. 9)

Em debate sobre o português do Brasil, o professor Silva Neto (1986, p. 65) afirma que a unidade não quer dizer igualdade e que, no tecido linguístico brasileiro, há gradações de cores e que um estudo minucioso determinaria, com segurança, várias áreas. O pesquisador salienta ainda que o conjunto dos falares brasileiros se coaduna com o princípio da unidade na diversidade e da diversidade na unidade. A diversidade a qual Silva Neto se refere está relacionada ao caráter variável da língua e conduz, a cada dia, à investigação de novos aspectos linguísticos passíveis de descrição e análise.

No imenso território brasileiro da língua portuguesa, têm sido desenvolvidos importantes trabalhos sobre as diferenças de natureza fonética, morfossintática, lexical, dentre outros, que particularizam as variedades regionais nele existentes. Esses estudos baseados nos aspectos socioculturais da língua são cada dia mais numerosos. Como sabemos, as pesquisas nesta área são também muito produtivas porque abrangem todos os aspectos da estrutura e do uso da linguagem que diz respeito às suas funções sociais e culturais.

Pensamos que a língua é um mecanismo fundamental para a comunicação entre os falantes e para que estes possam se posicionar criticamente em relação ao mundo que os cerca. Nesse sentido, para atuar criticamente, o falante utiliza estratégias linguísticas, dentre as quais destacamos a forma como ele se refere à segunda pessoa em um evento de fala. A forma de tratamento que um indivíduo utiliza para dirigir-se ao seu interlocutor depende, assim, do tipo de relação estabelecida entre os falantes, do gênero destes, do contexto, da situação conversacional, dentre outros fatores.

A motivação para esta pesquisa teve início na interação com os estudos sobre variação linguística durante a Pós-Graduação realizada na Universidade Estadual de Feira de Santana. Nesse ínterim, atentamos para a variação das formas *tu* e *você*, o que gerou uma inquietação sobre o motivo pelo qual alguns falantes da cidade de Santo Antônio de Jesus - Ba utilizavam, em determinadas situações, o pronome *tu* e, em outras, o pronome *você* (DIVINO, 2008). Por conseguinte, a dissertação de Mestrado sobre o tema foi orientada pela professora Dra. Suzana Alice Marcelino Cardoso e, a partir daí, os laços com a Dialetologia foram sendo estreitados. Algum tempo após o início das disciplinas no Programa de Pós-Graduação em Língua e Cultura, da Universidade Federal da Bahia, pensamos em um estudo de reanálise da variação

dos pronomes *tu* e *você*, envolvendo dados do Atlas Linguístico do Brasil (ALiB). Assim, a vontade de reacender pesquisas anteriores foi o motivo intenso para chegar ao tema desta tese.

Pretendemos, com esta pesquisa, compor um panorama acerca dos usos dos pronomes de 2^a pessoa, bem como aprofundar as análises precedentes e, através das evidências empíricas, avançar sobre a caracterização e o comportamento do fenômeno.

Sabemos que as formas de tratamento *tu* e *você* coocorrem no português do Brasil. Porém, as gramáticas tradicionais insistem em incluir apenas o *tu* para a 2^a pessoa do singular no quadro pronominal. Muitos gramáticos, todavia, já observam que esse pronome tem uma aplicação limitada e que no Brasil vem sendo substituído pela forma *você*.

O gramático Evanildo Bechara (2004), de perfil descritivista, não apresenta a variação dos pronomes, enquanto Azeredo (2008) considera *tu* e *você* como pronomes pessoais do caso reto. O pronome *você*, segundo Azeredo (2008), é a forma pronominal característica da interlocução coloquial na maior parte do território brasileiro. Em alguns Estados do Nordeste e do Sul, prevalece o *tu*. Em determinadas regiões, *tu* ocorre em par com *você*, utilizando o verbo na terceira pessoa, na maioria das vezes.

Poucas gramáticas tradicionais debatem sobre o uso das formas de tratamento. Cunha e Cintra (2007) fazem um paralelo entre o português falado no Brasil e o falado em Portugal e sinalizam que as mudanças no nosso português fizeram com que o *você* fosse usado como uma forma mais polida, indicando respeito ou um tratamento de igual para igual, como também forma de intimidade.

Estudos como o de Monteiro (1994) e o de Ilari (1996), ambos baseados em inquéritos do Projeto NURC¹, consideram *você* a verdadeira forma pronominal de segunda pessoa no português do Brasil. Para estes últimos, o pronome *tu* só sobrevive no Sul do País. Silva Antônio (2003) chega a declarar que:

Há que ressaltar que no Brasil há dois pronomes que têm a mesma função: *tu* e *você*. Atualmente, predomina este último em quase todo o território brasileiro, pois o *tu* se restringe ao Sul do País e a algumas regiões do Norte e Nordeste. (SILVA, 2003, p. 179)

Analizar o uso das formas de tratamento específicas *tu* e *você*, para a referência ao interlocutor no português, considerando contextos linguísticos e sociais que costumam favorecer uma das variantes no Nordeste do país, justifica-se por: (i) descrever a realidade do

¹ Projeto de Estudo da Norma Urbana Culta das Principais Cidades Brasileiras.

português brasileiro, com base em dados empíricos; (ii) fornecer resultados para a realização de estudos preliminares para o ALiB; e (iii) fornecer aos docentes de língua portuguesa documentação, a partir desta tese, para o aprimoramento do ensino-aprendizagem da língua vernácula.

A hipótese central do trabalho pauta-se no pressuposto de que os falantes da região Nordeste que residem no interior utilizam preferencialmente o pronome *tu* e os moradores das capitais dão preferência ao emprego de *você*, nas suas interlocuções.

Formulamos as hipóteses seguintes para o desenvolvimento desta tese: a) o pronome *você* é mais usado por falantes da faixa etária mais jovem; b) o pronome *tu* ocorre com maior frequência na faixa etária mais velha; c) as mulheres tendem a utilizar mais a forma *você* do que os homens; d) o falante tende a usar o pronome *você* com referência indeterminada nos discursos semidirigidos.

A motivação para as nossas hipóteses foram surgindo quando observávamos que o pronome *tu* era utilizado por pessoas íntimas e em momentos informais, enquanto o *você* era restrito a contextos regidos por relações assimétricas, entre pessoas mais distantes e/ou para marcar a situação formal. Outra motivação para a hipótese de que falantes das capitais utilizam o pronome *você*, enquanto falantes dos interiores utilizam o pronome *tu*, teve origem na observação feita, em 2008, quando foi finalizada a dissertação *Como trato o meu receptor?* (a propósito do uso do *tu* e *você* em Santo Antônio de Jesus). Enfim, por ter bastante contato com moradores da capital do Estado, Salvador, percebemos que os mesmos utilizam, majoritariamente, o pronome *você*. Quanto à hipótese de os jovens utilizarem mais a variante *você*, acreditamos que se justifique pelo fato de o pronome *você* ser inovador.

Para o desenvolvimento desta tese, temos como objetivo geral investigar o uso das formas de tratamento *tu* e *você* em cinco Estados do Nordeste, a partir da análise de um extrato do *corpus* do ALiB, enfocando todos os questionários (questões de prosódia – QP, questionário semântico-lexical – QSL, questionário morfossintático – QMS, temas para discurso semidirigidos – TDS e questões metalingüísticas – PM), exceto o Questionário Fonético-Fonológico (QFF). A amostra analisada conta com dados de fala de 140 informantes, oriundos de cinco Estados, a saber: Paraíba, Pernambuco, Rio Grande do Norte, Sergipe e Alagoas, compreendendo as suas referidas capitais e 25 municípios do interior, totalizando 30 localidades. Os informantes foram distribuídos de forma equânime entre ambos os sexos e por duas faixas etárias (18 a 30 anos e 50 a 65 anos) e com distribuição regular entre dois graus de escolaridade nas capitais.

Para verificar se, de fato, as hipóteses aqui levantadas nos ajudam a esclarecer a variação observada no que se refere à escolha dos tratamentos usados para a segunda pessoa nos Estados escolhidos, esta pesquisa tem como objetivos específicos: (i) analisar a relevância das variáveis sociais – localidade, faixa etária, sexo, escolaridade – no comportamento dos falantes, com relação ao uso das formas *tu* e *você*; (ii) analisar a relevância das variáveis linguísticas – parte do inquérito, tipo de referência, função sintática da variante, tempo verbal e paralelismo – na ocorrência das formas *tu* e *você*; (iii) representar, em cartas linguísticas, o uso das formas em estudo, com vistas a delinear a realidade dialetal nos Estados nordestino analisados; (iv) observar como as formas *tu* e *você* estão distribuídas, em cada localidade, entre os informantes; (v) verificar a percentagem de ocorrência e o peso relativo dos fatores que favorecem ou não esses pronomes.

A análise foi desenvolvida segundo os pressupostos teóricos da Dialetologia Pluridimensional Contemporânea e da Sociolinguística Variacionista. Enquanto a Dialetologia Pluridimensional Contemporânea tem por foco observar a distribuição espacial da língua e suas variantes, utilizando como método a Geolinguística Contemporânea, que leva em conta não apenas a procedência geográfica dos indivíduos mas também fatores tais como o sexo, a faixa etária e o nível escolaridade, a Sociolinguística Variacionista prioriza os princípios gerais para o estudo da mudança linguística descritos por Weinreich, Labov e Herzog (2006 [1968]), ou seja, a variação é inerente à língua; é sistemática e condicionada por fatores intra e extralingüísticos; apresenta funções sociais; e é fundamental para que a mudança linguística ocorra de maneira ordenada.

Nesta perspectiva, verificamos os efeitos dos fatores linguísticos e extralingüísticos que favorecem ou inibem o uso dos pronomes *tu* e *você* nos cinco Estados do Nordeste mencionados. Para tanto, estruturamos esta tese em quatro capítulos.

No Capítulo 1, intitulado “Formas de tratamento”, apresentamos alguns aspectos históricos das formas de tratamento. Em seguida, discorremos acerca dos pronomes de tratamento utilizados pelos falantes brasileiros e elencamos diferentes trabalhos que tratam da referência de 2^a pessoa no português do Brasil, com base em amostra tanto da língua oral quanto da língua escrita. As pesquisas analisadas neste capítulo ajudaram no mapeamento das formas *tu* e *você*, por região.

O Capítulo 2, intitulado “Fundamentação teórica”, foi dedicado aos aportes teóricos da Dialetologia e da Sociolinguística. De forma geral, discutimos aspectos da variação linguística e reiteramos sobre os tipos de variação. Mencionamos a Geografia Linguística, método

utilizado pela Dialetologia para a investigação de seus dados, bem como apresentamos aspectos da metodologia da Sociolinguística laboviana que abrangem a quantificação dos dados em variação, em relação aos fatores que podem estar contribuindo para a seleção de uma ou outra forma variante. Enfocamos as áreas dialetais do Brasil, especialmente aquelas apontadas por Nascentes (1952), acrescentando ainda as novas propostas de dialetação para o PB. Citamos os Atlas linguísticos que foram concluídos e os que estão em andamento, no território brasileiro. Destacamos a interface entre a Dialetologia e a Sociolinguística, bem como as contribuições dessas áreas de estudo para o ensino da língua materna.

Na sequência, o Capítulo 3, intitulado “Aspectos metodológicos”, apresenta, como sugere o título, o aporte metodológico utilizado na pesquisa. Inicialmente, trazemos informações sobre a metodologia empregada pelo Projeto ALiB, cuja base de dados possibilitou a constituição do *corpus* analisado nesta pesquisa. Posteriormente, destacamos o processo empreendido na recolha e seleção dos dados, procedimentos para o levantamento e codificação das ocorrências. Em seguida, mostramos o suporte estatístico utilizado para a análise quantitativa dos dados, o pacote GoldVarb X. Por fim, incorporamos elementos necessários para a confecção das Cartas linguísticas, ferramenta de grande valia para representação espacial dos dados desta pesquisa.

No Capítulo 4, intitulado “Análise dos dados”, apresentamos e discutimos os resultados obtidos na pesquisa. Optamos, inicialmente, em explicar sobre a variável dependente e as variáveis independentes (fatores linguístico-estruturais e fatores extralinguísticos) bem como sobre as hipóteses aventadas na pesquisa. Procedemos à análise dos dados e à interpretação dos resultados de cálculos estatísticos, discutindo os resultados dos grupos de fatores selecionados como estatisticamente relevantes pelo programa GoldVarb X na escolha dos falantes dos Estados estudados quanto aos pronomes *tu* e *você*. Foram feitas duas análises: uma geral, incluindo todos os Estados estudados e outra de cada Estado, separadamente. Os resultados apontaram que a variante escolhida com o valor de aplicação para análise estatística dos dados foi o pronome *tu*. A rodada binária do GoldVarb X apontou como variáveis relevantes no emprego de *tu*: referênciação, parte do inquérito, localidade e escolaridade. Na descrição da análise por Estado, as variáveis selecionadas como significativas foram: em Pernambuco – referênciação, parte do inquérito, faixa etária e localidade; na Paraíba – parte do inquérito, faixa etária e localidade; em Alagoas – sexo e tipo de questionário; no Rio Grande do Norte e em Sergipe nenhuma variável foi selecionada. A discussão desses resultados incluiu uma análise

comparativa com outras pesquisas, buscando situar as nossas hipóteses e melhor visualizar o fenômeno em outros estudos.

Por fim, nas Considerações finais, reunimos os resultados obtidos, confrontando-os com as hipóteses e objetivos aventados. Em seguida, apresentamos as “Referências” utilizadas neste trabalho e, logo após, os “Anexos”. De modo geral, almejamos contribuir para o estudo das formas de tratamento no português brasileiro bem como fornecer resultados empíricos que sirvam de base para o ensino da língua materna, que deve ser pautado no respeito e na reflexão sobre a diversidade linguística, visando à formação de cidadãos conscientes da variação e da riqueza da língua que fala e desprovidos de preconceitos.

1 FORMAS DE TRATAMENTO

A língua portuguesa oferece diferentes opções de formas de tratamento: *senhor(a)*, *você*, *tu*. Fazer uma escolha, nessa variedade de formas, é manifestar claramente o tipo de relação que há entre o falante e o seu interlocutor, e essa escolha é o ponto em que os falantes da língua mais explícita e conscientemente desempenham um papel.

Tratamos, neste capítulo, de forma breve, de alguns aspectos históricos das formas de tratamento e como estas foram conceituadas por alguns linguistas.

Vários trabalhos, entre os quais a pesquisa pioneira de Brown e Gilman (1960), mostraram grande interesse pela escolha das formas de tratamento como objeto de estudo em inúmeras línguas. Esses autores sugerem recorrer-se ao percurso histórico das formas de tratamento para explicar as relações que sempre existiram entre elas e as estruturas sociais. Dentre esses estudos, destacamos na seção 1.4 deste capítulo algumas pesquisas feitas no Brasil sobre a variação das formas de tratamento.

1.1 FORMAS DE TRATAMENTO: UM POUCO DE HISTÓRIA

A abordagem histórica é importante, no presente estudo, uma vez que a língua é um produto social e uma atividade do espírito humano, moldada de acordo às necessidades de quem dela se serve.

A história das formas de tratamento liga-se à história do próprio ser. Nos séculos passados, quando as sociedades eram altamente hierarquizadas, estáticas, com suas classes sociais bem delimitadas e definidas, era normal que as relações de poder fossem marcadas por formas de tratamento próprias a cada classe social (MODESTO, 2006, p. 3).

Considerando o percurso histórico das formas de tratamento para elucidar as relações que sempre existiram entre elas e as estruturas sociais, tentamos reconstituir aspectos do ambiente sociocultural que desencadeou a introdução de novas formas de tratamento do interlocutor em português. Desse modo, a história das formas de tratamento em português proporciona exemplos interessantes de como fatos socioculturais podem acarretar mudanças linguísticas, além de fornecer dados da língua que são mal interpretados pelos gramáticos normativos, os quais “insistem em apresentar algumas realidades próprias do português arcaico como modelos a serem ainda seguidos” (FARACO, 1996, p. 53).

É sabido, portanto, que as diversas línguas passaram por diferentes mudanças que as afastaram menos ou mais do sistema inicialmente herdado. O latim, por exemplo, sofreu alterações por conta do declínio do Império Romano e subdividiu-se em: *latin clássico* e *latin vulgar*. O primeiro, reservado à língua literária, também chamado de *sermo urbanus*, tornou-se estável por ser a língua dos escritores latinos e da gente culta, utilizado apenas por uma minoria (SETTE, 1980, p. 10). Já o *latin vulgar*, denominado de *sermo vulgaris*, era a língua popular e, por isso, apresentava muitas alterações em relação ao latim dos escritores. Esse latim refletia ainda o dia a dia de todo o Império Romano, era dinâmico e deu origem à língua portuguesa, assim como às demais línguas românicas.

Na România medieval, quando se constituíram os padrões literários e normativos das línguas românicas modernas, a estrutura social dividia-se em: nobreza, clero e povo. As relações entre esses estratos, onde o poder era indicado pela forma de tratamento, eram assimétricas.

Curioso é o fato de que no começo das monarquias lusitana e castelhana não se utilizava o título de Dom, a não ser quando estivesse se referindo ao rei ou aos seus filhos legítimos. Dom, derivado de *dominus*, que significava senhor, não se aplicava nem aos governantes nem aos grandes príncipes. No final do século XII, essa forma de tratamento passou a ser usada pelos cronistas referindo-se aos senhores feudais e ricos-homens (BIDERMAN, 1972, p. 343). Nesse período, aumentou a quantidade dos títulos de Conde, Marquês, Duque e Dom. Visto assim, à medida que a nobreza adquiriu muitos privilégios, consequentemente, começou o desgaste. Por exemplo, o Dom já não se caracterizava como uma forma respeitosa e elegante para uma pessoa dirigir-se a um fidalgo, sendo aplicado, então, indiscriminadamente (BIDERMAN, 1972, p. 343).

Na Itália, deu-se a ampliação da sociedade do topo devido à Renascença, o que acarretou numa relação de servilismo entre os senhores e os príncipes italianos. Ao lado dessa relação desenvolveu-se um tratamento complexo entre a senhoria local e os seus cortesãos: Maestro, Reverendo, Signor, Signor Mio, Illustríssimo, Illustríssimo Signore, Vostra Illustríssima Signoria, Reverendíssimo e Magnífico, Loro Signorie Reverendíssime, Madama, Vostra Maesta, Vostra Santita, Sereníssima Regina. Dentre essas formas de tratamento, mencionam-se algumas de caráter respeitoso e até servis como: Maestra, Santita, Eccelenza, Signoria, derivadas de abstratos e que foram usadas no latim medieval clássico. (BIDERMAN, 1972, p. 344).

Na França, as formas de tratamento foram mais simples: *tu* (para inferiores e também para íntimos, ou para marcar a emoção) e *vous* para as pessoas bem-nascidas (entre iguais).

Assim, a língua francesa diferenciava as posições hierárquicas pelo uso assimétrico de *tu* e *vous*. “Se considerarmos o sistema dual dos pronomes de tratamento em francês ao longo da sua história, o que mais impressiona é a permanência inalterável desse sistema” (BIDERMAN, 1972, p. 349).

Na Espanha do século XVII, *vos* passa a denotar dois valores: tratamento dado a um inferior e tratamento dado à pessoa íntima com quem se tem familiaridade. Nisso, o tratamento de cerimônia fazia-se ausente no sistema, o que resultou na utilização de uma forma mais longa, *vuestra merced*, no espanhol, que passou por uma série de alterações fonológicas até se tornar o pronome *usted* (*você*), o que acarretou várias alterações no sistema verbal da língua em questão.

Na língua portuguesa, o pronome *vós*, utilizado em um contexto de plural, foi substituído pelo plural das novas formas (*vocês*, *senhoras*, *senhores* etc.), tornou-se um traço arcaizante das formas verbais de 2^a pessoa, limitando-se o seu uso à linguagem bíblica ou religiosa oficial, ao gênero oratório ou literário.

Segundo Biderman (1972, p. 375), no português brasileiro percebe-se uma grande mistura entre os hábitos do passado e os novos conceitos do presente. As mais recentes publicações da Bíblia no Brasil misturavam a 2^a pessoa do plural (antiga) com a 2^a pessoa do singular. Na década de 1960, muitos livros litúrgicos foram publicados e, em alguns deles, encontrava-se a mistura de tratamento de 2^a pessoa (*tu*, *vós*) com a 3^a pessoa (*você*). A Virgem Maria às vezes era tratada por *você*, ora por *vós*, ora por *tu*. “O *vós* e o *tu* representam versões canônicas das preces e leituras bíblicas; o *você* figura nas livres criações paralitúrgicas” (BIDERMAN, 1972, p. 376). Enfim, a Igreja Católica Romana vive um conflito buscando equilibrar uma vivência dentro e fora do tempo. De um lado, quer ter um pé na terra tentando dialogar com os homens usando a sua linguagem e, do outro, tenta manter uma linguagem sagrada própria. Hoje, defendem-se relações mais humanas e diretas com Jesus Cristo e muda-se a maneira como se trata um membro da família, abandonando a antiga hierarquia defendida pela Igreja Romana.

Cintra (1972) analisa as formas de tratamento em português a partir do contexto histórico-social que as licencia na sociedade portuguesa. Ao iniciar a reconstrução das origens das formas de tratamento, admite que o português herdou do latim tardio um sistema bifurcado de referência à 2^a pessoa do discurso – *tu* e *vós*. Segundo o autor, entre o final do século XIII e a primeira metade do século XI, dava-se preferência aos pronomes *tu*, no plano da intimidade, e *vós*, no plano da cortesia. O sistema de formas de tratamento da língua portuguesa

assemelhava-se ao do francês, que continua, até hoje, assumindo as formas pronominais *tu* e *vous* na expressão do tratamento informal e formal, respectivamente. Até princípios do século XV, o pronome de tratamento *vós* representou a forma de tratamento cerimonioso preferida pela realeza portuguesa.

1.2 FORMAS DE TRATAMENTO NA VISÃO DE ALGUNS AUTORES

Apoiando-se na concepção de Bechara, Soares (1980, p.19) faz distinção entre as formas de tratamento do português, classificando-as em formas diretas e formas indiretas. As formas diretas são os pronomes pessoais (*eu*, *nós*, *tu* e *vós*); as formas indiretas são os pronomes de tratamento (*você*, *o senhor*, *vossa senhoria* etc.), que designam a pessoa com quem falamos, isto é, a 2^a pessoa, e levam o verbo para 3^a pessoa.

Admite, também, outras formas indiretas – *o senhor professor*, *o amigo*, *o doutor* etc.–, que podem ser empregadas como o sujeito da oração: “*o senhor diretor fez o pedido*”; “*o doutor deseja sair*?”.

Sette (1980, p. 15) salienta que nas formas de tratamento da língua portuguesa incluem-se também os chamados pronomes de reverência (*Vossa Excelência*, *Vossa Majestade*, *Vossa Santidade* etc.), os quais se referem a uma virtude, um título, uma qualidade ou uma função que tenha a pessoa com quem falamos.

Cunha e Cintra (2007) assim se manifestam:

[...] denominam-se pronomes de tratamento certas palavras e locuções que valem por verdadeiros pronomes pessoais, como: *você*, *o senhor*, *Vossa Excelência*. Embora designem a pessoa a quem se fala (isto é, a 2^a), esses pronomes levam o verbo para a 3^a pessoa. (CUNHA; CINTRA, 2007, p. 303-304)

Visto que esses autores excluem o pronome pessoal reto de 2^a pessoa (*tu* e *vós*) das formas de tratamento por acreditarem que são “formas mortas” (SOARES, 1980, p. 18) no Brasil. Eles acreditam que a forma *tu* só existe no Sul do Brasil e com uma pouca incidência no Norte do país.

Silva (2003, p. 170) aponta que o sistema de tratamento no português pode ser representado por três formas:

a) as formas pronominais, que são os pronomes pessoais – *tu*, *vós*;

- b) as formas pronominalizadas, que são os pronomes de tratamento – *Você, o Senhor, Vossa Excelência, Vossa Senhoria* e suas variações;
- c) as formas nominais – formas constituídas por nomes próprios, prenomes, nomes de parentesco ou equivalente, antecedidos de artigo, com uso praticamente restrito ao português de Portugal, ou, ainda, por uma grande variedade de nomes empregados como vocativos ou formas de chamamento.

Silva (2003, p. 171) conceitua as formas de tratamento como “palavras ou sintagmas que o usuário da língua emprega para se dirigir e/ou referir à outra pessoa”. Ele estabelece quatro níveis para essas formas de tratamento:

- a) formas pronominalizadas;
- b) formas nominais;
- c) formas vocativas;
- d) outras formas referenciais.

Salientamos que a classificação de Preti (2000) está dividida em formas pronominalizadas, formas nominais e formas vocativas. Assim, a diferença entre a classificação de Preti (2000) reside no fato de que Silva (2003) trata de outras formas referenciais, isto é, palavras usadas com referência à pessoa de quem se fala. Por exemplo, duas amigas conversando e uma delas faz a seguinte pergunta sobre uma outra amiga: “**A Nara** vai à festa?” (forma usada no âmbito subjetivo). Além disso, o autor une as formas nominais e pronominalizadas em um só grupo, que ele chama de “formas pronominalizadas”.

Segundo Soares (1980) o sistema do português atual apresenta uma variedade muito grande dos tratamentos do tipo nominal e pode estar estruturado em três planos: formas de tratamento próprias de intimidade; formas usadas no tratamento de igual para igual (ou superior para inferior) e que não acarretam intimidade; e formas chamadas “de reverência” ou “cortesia”.

A maioria das gramáticas normativas (GNs) apresenta um quadro pronominal da língua portuguesa há muito discutido por pesquisas linguísticas. Uma das controvérsias gira em torno da composição da 2^a e da 3^a pessoas do singular. Como exemplo da classificação adotada pelas GNs, o Quadro 1 reproduz o paradigma adotado por Cunha e Cintra (2007):

Quadro 1: Pronomes-sujeito em Cunha e Cintra (2007)

| PRONOMES PESSOAIS RETOS | | |
|-------------------------|-----------------------------|------------|
| SINGULAR | 1 ^a pessoa | Eu |
| | 2^a pessoa | Tu |
| | 3 ^a pessoa | Ele, ela |
| PLURAL | 1 ^a pessoa | Nós |
| | 2 ^a pessoa | Vós |
| | 3 ^a pessoa | Eles, elas |

Fonte: Adaptado de Cunha e Cintra (2007, p. 291).

Um dos motivos de discórdia entre os linguistas e as GNs tem sido a não inclusão de *você* ou *cê*, resultado de processo de gramaticalização de *Vossa Mercê*, no quadro pronominal de segunda pessoa. O pronome *você* “vale por verdadeiro” pronome pessoal, mas não é tratado como tal. Esta forma leva o verbo para a terceira pessoa. Menon (1993) contra-argumenta a questão nos seguintes termos:

[...] a forma **você** (s) origina-se de uma locução nominal [...] e, nessa categoria, passa a requerer o verbo na terceira pessoa. No entanto, durante o processo de modificação fonética e de valor social, a forma se pronominalizou, isto é, passou por um processo de gramaticalização, mudando de categoria: de nome [...] para pronome. Este novo pronome é de segunda pessoa, logo a forma verbal que o acompanha também passa a ser uma forma de segunda pessoa. Então não faz sentido algum continuar a dizer que o verbo está na terceira pessoa com um pronome de segunda pessoa. [...] a língua portuguesa passa a ter [...] uma reestruturação no seu paradigma verbal, em que a segunda pessoa do singular passa a ter duas formas... (MENON, 1993, p. 3-4)

O fato de nessas gramáticas normativas não ser incluído o *você* na classe dos pronomes de segunda pessoa, segundo Martins (2010), tem sido contestado por vários linguistas. Ainda segundo esse autor, a forma *você* faz com que a concordância verbal fique na terceira pessoa, com o que Menon (1993) não concorda, uma vez que, segundo ela, *você*, que se originou de uma forma nominal – *Vossa Mercê* –, ao passar por um “processo de modificação fonética e de valor social” mudou da categoria de nome para a de pronome de segunda pessoa, razão pela qual, ainda segundo a autora, a forma verbal que deve acompanhar esse novo pronome também deve ser de segunda pessoa, não fazendo sentido, portanto, dizer que o verbo concorda na terceira pessoa com um pronome de segunda.

Reiteramos que o pronome de 2^a pessoa (*você*) é utilizado com o verbo na conjugação de 3^a pessoa (*você vai*), ou seja, o pronome *você* transfere a concordância verbal para a terceira pessoa.

Por fim, o Quadro 2 mostra, de maneira simplificada, o conceito de formas de tratamento defendido por alguns autores.

Quadro 2: Formas de tratamento segundo alguns pesquisadores

| PESQUISADORES | FORMAS DE TRATAMENTO E CONCEITO |
|-----------------------|--|
| Soares (1980) | Formas diretas: pronomes pessoais (<i>eu, nós, tu e vós</i>) Formas indiretas: pronomes de tratamento (<i>você, o senhor, vossa senhoria</i> etc.) Aceitam-se outras formas indiretas: o senhor professor, o amigo, o doutor etc. |
| Sette (1980) | Incluem-se os pronomes de reverência (<i>Vossa Excelência, Vossa Majestade, Vossa Santidade</i> etc.) |
| Cunha e Cintra (1985) | <i>Você, o senhor, Vossa Excelência</i> |
| Silva (2003) | a) formas pronominalizadas (<i>tu e vós</i>); b) formas nominais (<i>você, o senhor, Vossa Excelência, Vossa Senhoria e suas variações</i>); c) formas vocativas; d) outras formas referenciais. |
| Martins (2010) | <i>Você</i> (pronome de 2 ^a pessoa / forma verbal de 3 ^a pessoa) |
| Menon (1993) | <i>Você</i> (pronome de 2 ^a pessoa / forma verbal de 2 ^a pessoa) |

Fonte: Elaborado pela autora.

Através do quadro apresentado, constatamos uma realidade incompatível com o paradigma pronominal vigente no vernáculo brasileiro atual. Nota-se, dessa maneira, uma distância significativa entre o conceito de pronome de tratamento, prescrito em algumas gramáticas e pesquisas diferenciadas, do uso dos pronomes no Brasil. Essa dessemelhança, indubitavelmente, acarreta prejuízos no ensino-aprendizagem da língua materna.

De modo geral, as gramáticas normativas brasileiras registram o *tu* como pronome pessoal do caso reto de segunda pessoa do singular e o pronome *você* como forma de tratamento.

1.3 PRONOMES DE TRATAMENTO NO BRASIL

Os pronomes de tratamento em algumas línguas europeias traçaram um percurso diferente do percurso do português do Brasil, o qual, seguindo um processo mais simplificado do que o verificado nessas línguas, apresenta um sistema dual: tratamento de intimidade (*você/tu*) e tratamento formal (*senhor ou senhora*).

Biderman (1972) admite que no Brasil só há dois pronomes de tratamento: *você* (familiar) e *senhor* (formal), que correspondem ao par T/V, conforme a caracterização de Brown e Gilman (1960).

O português do Brasil (doravante PB) tem recursos que permitem ao falante evitar a expressão da relação social (a opção do pronome zero, o nome pronominalizado, e formas ambíguas como *lhe*). Utiliza títulos e termos de parentesco como pronomes sintáticos de tratamento, tem alternativas formais e informais disponíveis em vários pontos relativos ao tratamento, como assinala Ramos (1989, p. 27).

Menon (1997) salienta sobre a variedade da língua portuguesa falada no Brasil, mostrando que houve várias mudanças, entre as quais uma reorganização do sistema pronominal. Um processo contínuo de transformação no sistema pronominal do português brasileiro vem se arrastando ao longo dos anos e tem afetado as formas pronominais de sujeito, de clíticos e dos possessivos, resultando em um novo sistema pronominal. Com o objetivo de mostrar a relação de posse na segunda pessoa, Menon (1997) faz um estudo da variável *seu* ~ *de vocês* e conclui que ocorrem mais casos da forma inovadora *de vocês* do que da forma canônica *seu*.

Ainda falando das alterações no PB, Menon (1997, p. 102) afirma que temos dois pronomes possessivos para referência à 2^a pessoa: *teu* e *seu*. Em alguns dialetos, parece que a forma *seu* é empregada com maior frequência; outros, como parece ser o caso do dialeto curitibano, na situação de interlocução, o falante varia o pronome a ser empregado.

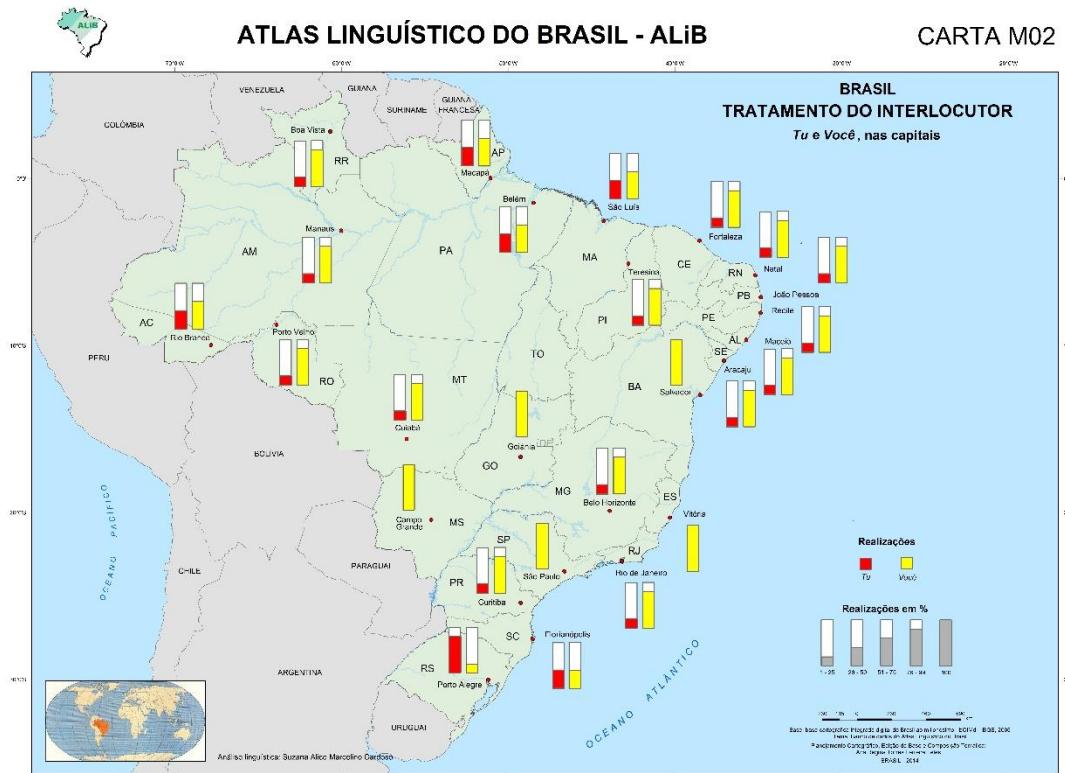
Nas gramáticas normativas, como se vê em Said Ali (1921, p. 93), os pronomes pessoais apresentam-se do seguinte modo: *eu*, pronome de 1^a pessoa, representa o indivíduo que fala e seu plural é *nós*. *Tu* e *vós* são pronomes de 2^a pessoa; *ele*, *ela*, *eles* e *elas* são os da 3^a pessoa.

Os gramáticos normativos, ao descreverem os pronomes de tratamento no português brasileiro, ainda recorrem às formas arcaicas, sem muita preocupação com o uso real da língua. Para exemplificar essa desarmonia, podemos comparar o que dizem as gramáticas com o que é encontrado nas análises da língua falada, como na descrição dos pronomes pessoais no

português brasileiro, assim apresentada por Ilari (2002, p. 79): primeira pessoa – *eu, nós e a gente*; segunda pessoa – *tu, você, o senhor, a senhora*; terceira pessoa – *ele, eles, ela, elas*.

A variação entre os pronomes *tu* e *você* no tratamento do interlocutor foi documentada pelo Projeto ALiB e está registrada na Carta MO2², que apresenta os dados das capitais brasileiras, reproduzida a seguir:

Figura 1: Carta do uso de *tu* e *você* nas capitais brasileiras



Fonte: Cardoso et al. (2014, p. 349).

Fazemos aqui uma análise intuitiva, pois não dispomos ainda dos números que serviram de base para a geração da carta. A Carta MO2 apresenta em vermelho os dados de *tu* e em amarelo os dados de *você* e o preenchimento das barras indica a faixa percentual dos dados exposta na legenda na base direita da carta.

Uma leitura visual nos permite dizer que, mesmo com um percentual baixo, ainda encontramos o pronome *tu* em algumas capitais do Brasil. A cidade de Porto Alegre apresentou a maior utilização da variante *tu* e, logo em seguida, Florianópolis e Belém. Percebemos, também, que nas cidades de Macapá, Rio Branco e São Luiz, coincidentemente, localizadas no

² O Projeto ALiB publicou a carta no volume 2 do Atlas, mas só no volume 3 (no prelo) sairá o texto com a análise dos dados da Carta e os respectivos números percentuais que lhe serviram de base. Por esse motivo, optamos por reproduzi-la aqui e não na seção 1.4, que revisita os trabalhos que já apresentaram resultados estatísticos.

Norte do país, há falantes empregando o *tu* de maneira significativa, atestando um percentual entre 26% a 50%.

Por outro lado, a Figura 1 demonstra que nas cidades de Campo Grande, Goiânia, Salvador e Vitória não foi encontrado nenhum falante utilizando o pronome *tu*.

Para obter a resposta *tu/você* sujeito, elaborou-se a questão 24 do Questionário Morfossintático que vem assim formulada: Quando se vê um amigo com uma mala e se quer saber para onde ele vai, como é que se pergunta?

Aludimos, ainda, que a Carta MO2 registra não somente os resultados atingidos como resposta à pergunta 24, bem como todas as ocorrências obtidas e documentadas na aplicação dos questionários: QFF, QSL, QMS, QP, TDS e PM (Cardoso et al., 2014, p. 348).

1.3.1 De *Vossa Mercê* a *você*: uma abordagem diacrônica

Almejando estudar a evolução da forma de tratamento *Vossa Mercê* e entender melhor o ponto de chegada, o pronome *você*, propomos tecer algumas considerações sobre a sua história entrecruzando a história social e a história linguística do pronome. Essa análise diacrônica objetiva situar a mudança social que acarretou na mudança linguística ocorrida no pronome *você*, que antes era uma forma nominal (SOUZA, 2008).

Faraco (1996, p. 52) já destacava que “as mudanças na forma de tratamento estão correlacionadas com as mudanças nas relações sociais e valores culturais” e o pronome *você* faz parte dessa regra.

Essas mudanças justificaram-se pelas transformações sociais que ocorreram a partir do século XIV na Europa ocidental. Era o começo de uma “nova organização econômica que estaria centrada nas cidades e resultaria na formação de uma nova classe social, a burguesia, que iria competir com a nobreza em termos de poder econômico e político (FARACO, 1996, p. 55). A alta burguesia torna-se a nova aristocracia e cria-se uma verdadeira nobreza burocrática com o crescimento de cargos devido à ampliação da estrutura administrativa do novo governo na expansão colonial. Muitas pessoas passam a depender da corte, criando, desse modo, novos costumes. Assim, novas formas de vestuário, de alimentação e de tratamento do interlocutor foram introduzidas para atender às necessidades da nova realidade. O rei passa a ter uma nova posição de autoridade, o que exigirá formas diferenciadas de tratamento para este, uma vez que a forma tradicional *vós* passa a ser insuficiente para marcar o novo *status* do rei.

Dessa maneira, considerado como um chefe militar, a quem cabiam as demandas da corte, o rei passou a ser o detentor do poder absoluto e, sendo assim, a forma pronominal de tratamento *vós* é substituída por não ser mais adequada para a referência real. Até aqui, essa forma era eleita pelos reis, rainhas, nobres para o tratamento com os vassalos e, concomitantemente, também era a forma utilizada pelos vassalos para se dirigirem aos seus superiores. Porém, com as desigualdades sociais, a linguagem sofre modificações. Em suma, a forma *vós*, antes usada para fazer referência à figura singular do rei, é substituída por *Vossa Mercê*, forma que, nesse período, demonstrava ter mais expressividade.

O português herdou seu sistema pronominal do latim e dentre as línguas românicas somente o francês conservou esse sistema tardio. Esse sistema se organizava em dois eixos: o paradigma do pronome *tu* para tratamento singular e não formal e o paradigma *vous* para tratamento singular formal, ou para tratamento plural (formal ou não formal). Não somente a língua portuguesa como as demais línguas sofreram diferentes mudanças no seu sistema de tratamento, uma verdadeira “revolução da terceira pessoa” (FARACO, 1996, p. 54).

Faraco (1996) detalha o processo sócio-histórico que acarretou a inserção na língua de formas de tratamento com estrutura *Vossa + N* (*Vossa Mercê*, *Vossa Senhoria*, *Vossa Alteza*, *Vossa Excelência*, *Vossa Majestade*) para substituir o pronome *vós*, tratamento herdado do latim, como já fora mencionado anteriormente.

Em relação à evolução de *Vossa Mercê* > *Você* tem-se um processo de gramaticalização, o que resulta numa mudança de categoria de locução nominal para pronome. No português arcaico, *mercê* era um substantivo comum que significava “favor, graça, benesse”. Esse substantivo, antecedido do pronome possessivo *vossa* cristalizou-se, adquirindo um novo significado e tornando-se uma forma de tratamento específica para o rei. Mais tarde, esse pronome estendeu-se a fidalgos fidalgotes que queriam o mesmo tratamento dos seus criados e subalternos e continuou alargando-se a todas as classes sociais, como afirma Nascentes (1956):

Vossa mercê agradava a todo mundo. A classe humilde não tardou a apoderar-se da fórmula nova para uso próprio, mas, sendo expressão um tanto longa e tendo de ser repetida a cada instante, a gente do povo abreviou-a em vossancê, vossemecê, vossecê e finalmente você. [...] *Vossa mercê* se transformou em *vossemecê*. De *vossemecê* passou a *vosmecê* e desta forma por intermédio das formas hipotéticas *vosm'cê* e *voscê'se* fez você, que ainda se alterou para *ocê* e finalmente para *cê*. (NASCENTES, 1956, p. 117)

Vitral (1996) defende a hipótese de que o processo de gramaticalização estaria ainda mais avançado em relação às três formas pronominais de 2^a pessoa: *você*, *ocê* e *cê* e, analisando

sintaticamente essas formas, argumenta ter havido uma cliticização, isto é, a forma *cê* significaria a etapa mais avançada da gramaticalização, caracterizando-se como um *pronomé clítico*. O uso da variante *cê* com interpretação indefinida é mais frequente tanto na área urbana quanto na área rural.

No tocante ao português, nota-se que, nos primeiros textos, a forma de tratamento herdada era visível e, com o decorrer do uso, as modificações foram sendo feitas. A forma de referência ao interlocutor não se aliava mais à segunda pessoa verbal, mas com a terceira, como assinala Faraco (1996):

Esse fato introduziu na língua uma duplicidade de formas (as herdadas se combinando com a segunda pessoa verbal e as novas se combinando com a terceira pessoa verbal) que acabou por gerar grande instabilidade nos paradigmas verbais e pronominais redesenhando-os, por consequência; e definindo vários dos traços que caracterizam o português atual. (FARACO, 1996, p. 54)

Vimos que novas formas de tratamento foram criadas para substituir a forma tradicional, *vós*, e dentre essas formas, a mais antiga parece ser *Vossa Mercê*, que deu origem a *você*, tratamento utilizado no português moderno. A forma cortês de tratamento *Vossa Mercê* passou por um processo de gramaticalização, o que promoveu uma mudança de categoria de locução nominal para pronomé, acarretando, portanto, em perda de significado (valor) e desgaste fonético.

1.4 TRABALHOS REALIZADOS SOBRE O *TU* E O *VOCÊ* DE NATUREZA DIALETOLÓGICA E SOCIOLINGUÍSTICA

Muitos estudos já foram feitos com o objetivo de analisar o quadro pronominal do português brasileiro. Apresentamos, a seguir, alguns estudos significativos sobre a variação dos pronomes de segunda pessoa no singular, selecionados por região, conforme se pode observar no Quadro 3. Esses trabalhos têm em comum, além do objeto de estudo, a investigação de fatores externos à língua, como sexo, idade e escolaridade.

Quadro 3: Estudos sobre a 2^a pessoa: distribuição por região

| REGIÃO DO BRASIL | DADOS DE LÍNGUA ORAL | DADOS DE LÍNGUA ESCRITA |
|---------------------|--|--------------------------------|
| NORTE | Martins (2010) | - |
| NORDESTE | Soares (1980) Bezerra (1994) Herênia (2006) Nogueira (2013) Guimarães (2014) | Sette (1980) Almeida (2012) |
| CENTRO-OESTE | Lucca (2005) Dias (2007) Andrade (2010) Scherre et al. (2015) | - |
| SUDESTE | Paredes Silva (2003) Modesto (2006) Mota (2008) Silva (2017) | Salles (2001) |
| SUL | Ramos (1989) Loregian-Penkal (2004) Franceschini (2010) | - |

Fonte: Elaborado pela autora.

A seguir, comentam-se os trabalhos elencados por região no Quadro 3.

1.4.1 Região Norte

Sobre a Região Norte, apresentamos o estudo de Martins (2010), o único que conhecemos sobre os pronomes de tratamento nessa parte do País.

1.4.1.1 Martins (2010)

Para efetivação da pesquisa, Martins (2010) analisa a alternância *tu/você/senhor* tomando por base a Teoria da Variação e Mudança Linguística proposta por Weinreich, Labov e Herzog (2006 [1968]). Os dados do *corpus* foram coletados entre 2008 e 2009. Foram 19 entrevistas realizadas com 30 informantes: quatro entrevistas sem o conhecimento prévio dos informantes e 15 entrevistas nos moldes labovianos. Os informantes eram pessoas nascidas em Tefé, no Amazonas. A amostra inclui três faixas etárias (7 a 10, 20 a 35 e mais de 50 anos de idade), dois níveis de escolaridade (fundamental e universitário) e os gêneros masculino e feminino.

O objetivo inicial era avaliar apenas as ocorrências de *tu*, *você* e *senhor*, mas, como houve um número considerável de pronome zero, adicionou-se esse aspecto na investigação. Somou-se um total de 1.213 dados, dos quais 347 (28,6%) são de pronome zero. Na Tabela 1 apresenta-se a distribuição dos pronomes *tu/você/senhor* pelos falantes de Tefé.

Tabela 1: Distribuição dos falantes de acordo com o uso dos pronomes

| Pronome usado | Nº de falantes |
|---|----------------|
| só <i>tu</i> | 4 |
| só <i>você</i> | - |
| só <i>senhor</i> | - |
| <i>tu</i> e <i>você</i> | 10 |
| <i>tu</i> e <i>senhor</i> | 2 |
| <i>tu</i>, <i>você</i> e <i>senhor</i> | 14 |
| Total | 30 |

Fonte: Martins (2010, p. 40).

Observando a Tabela 1, percebemos que quase metade dos informantes (46,6%) usou as três formas do fenômeno sob análise. O pronome *tu* é o único usado por todos os informantes (100%), quatro dos quais usaram apenas esta forma. Martins (2010) interpretou que esses falantes com uso categórico de *tu* nos dados não são usuários categóricos de fato desse pronome; por esse motivo seus dados foram analisados juntamente com os dos demais.

As variáveis trabalhadas pelo autor foram: tipo de referência (genérica e específica); tipo de discurso (direto e relatado); paralelismo (isolado não precedido de forma pronominal, isolado precedido de *tu*, isolado precedido de *você*, isolado precedido de *senhor*, primeiro da série não precedido de forma pronominal). Os fatores sociais foram: escolaridade: fundamental (até cinco anos de estudo) e superior (cursando um curso universitário ou graduado) e gênero (masculino e feminino). Considerou também a variável pragmática tipo de relação entre interlocutores: simétrica e assimétrica. As gravações ocorreram de forma oculta e consciente.

Tabela 2: Totais de referência à segunda pessoa

| | <i>Tu</i> | Zero | <i>Você</i> | <i>Cê</i> | <i>Senhor</i> | Total |
|----------|-----------|------|-------------|-----------|---------------|-------|
| N | 520 | 347 | 278 | 8 | 60 | 1213 |
| % | 42,9 | 28,6 | 22,9 | 0,7 | 4,9 | 100% |

Fonte: Martins (2010, p. 52).

Enfim, os dados, após submetidos ao Varbrul, revelaram que o pronome *tu* (com 42,9% ocorrências) é o mais utilizado pelos informantes e que a concordância com a forma verbal canônica é muito baixa (apenas 3,7%). Além disso, é informada a frequência de uso do *senhor*, 4,9%, e do *você*, 22,9%. Dos 1213 dados, foram desconsiderados os dos pronomes zero, cuja frequência foi de 28,6%. Sobre o pronome *senhor*, usada nas relações íntimas entre pais e filhos, o autor afirma que o seu peso relativo não foi calculado, porque essa forma, nas situações em que é utilizada, é categórica.

Salientamos ainda que, dentre todas as variáveis selecionadas para *tu*, observa-se que: o paralelismo (precedido por outro *tu*, com 64,5% e peso relativo de 0,753) é o grupo mais relevante dentre todos; o grau de intimidade aparece como a variável social de maior relevância e mostra que o emprego do *tu* cresce com o aumento da intimidade, com 80,9% e peso relativo de 0,59; quanto ao tipo de gravação, as gravações ocultas se destacam com 90,7% e peso relativo de 0,78.

1.4.2 Região Nordeste

Sobre os pronomes de tratamento na Região Nordeste, apresentamos os trabalhos de Sette (1980), Soares (1980), Bezerra (1994), Herênia (2006), Almeida (2012), Nogueira (2013) e Guimarães (2014).

1.4.2.1 Sette (1980)

Sette (1980) procurou fazer o levantamento de algumas formas de tratamento na linguagem coloquial do recifense e descobrir quais os fatores extralingüísticos responsáveis pela escolha de uma determinada forma de tratamento numa situação de comunicação específica. Para a constituição da amostra, a pesquisa foi dividida em duas fases: na primeira, os informantes responderam a um questionário direcionado; na segunda, foram gravadas algumas conversas espontâneas.

Os informantes foram distribuídos em dois grupos: os fixos e os não-fixos. Os informantes fixos são os informantes que trabalham no local em que a pesquisa foi realizada e os não fixos fornecem os dados para a segunda parte da coleta de dados. Por exemplo, o informante fixo é um garçom e os fregueses são os não-fixos. Portanto os informantes fixos

participam das duas etapas da pesquisa, enquanto que os não-fixos participam apenas da segunda etapa.

Alguns fatores sociais foram analisados: sexo, idade, profissão, nível social e o grau de instrução. Em relação à variável idade, verifica-se que parte dos informantes é relativamente jovem.

Analizando as respostas dadas ao questionário, Sette (1980) percebeu que, em relação às formas *tu* e *você*, a maioria dos informantes diz usar mais a forma *você* (65%), em segundo lugar vêm as pessoas que dizem usar somente *você* (20%), logo depois temos aquelas pessoas que dizem usar as duas formas da mesma maneira e com a mesma frequência (12,5%), e em último lugar vêm as que usam mais a forma *tu* (2,5%), conforme a Tabela 3, a seguir.

Tabela 3: Emprego das formas de tratamento *tu* e *você* em Sette (1980)

| SEXO | + <i>TU</i> | + <i>VOCÊ</i> | <i>SÓ TU</i> | <i>SÓ VOCÊ</i> | DO MESMO JEITO |
|--------------|-------------|---------------|--------------|----------------|-----------------------|
| MASC. | - | 12 | - | 6 | 2 |
| FEM. | 1 | 14 | - | 2 | 3 |
| TOTAL | 1 | 26 | - | 8 | 5 |
| % | 2,5% | 65% | - | 20% | 12,5% |

Fonte: Sette (1980, p. 105) (adaptada).

De acordo com a interpretação feita por Sette, o *você* realmente aparece com uma certa frequência na fala recifense, porém a forma *tu* parece ser considerada mais íntima e familiar do que a forma *você*. A maioria dos informantes conceituou o *tu* como forma muito íntima, reservada apenas para familiares ou pessoas com quem se tem grande aproximação (SETTE, 1980, p. 106).

Analizando uma parte do questionário direcionada ao âmbito familiar, no tratamento usado com os pais, foram encontrados tanto os pronomes *o senhor* (87,5%) e *a senhora* (82,5%) como *você* (pai: 12,5% e mãe: 17,5%). Os informantes se dividiram em três grupos: alguns afirmavam usar exclusivamente *tu* (2,5%); outros, exclusivamente *você* (52,5%); e outros, *tu* e *você* (45%) com os irmãos.

No âmbito de trabalho, os informantes afirmaram usar apenas *você* e *o senhor*. Nesse ambiente, o fator idade só é significativo na escolha de tratamento de superior para inferior. Quando é de inferior para superior, a preferência foi quase sempre *o senhor*. A idade tem menos importância na escolha de tratamento para os homens do que para as mulheres.

1.4.2.2 Soares (1980)

Soares (1980) estudou as formas de tratamento correntes na fala de Fortaleza. A pesquisa foi feita em duas etapas: a primeira consistiu numa entrevista com 20 pessoas com base em um questionário de 30 perguntas; e a segunda teve por objetivo a aplicação de um questionário composto de 27 gravuras. Para completar seu *corpus*, utilizou conversas informais sem o conhecimento dos informantes. Enfim, a amostra foi resultado de dados provenientes de 72 informantes.

O questionário foi aplicado com um número significativo de diádes para motivar os informantes a usarem tanto o tratamento simétrico, quanto o assimétrico. Os 46 informantes foram agrupados em sete classes e cada uma dessas classes abrangeu duas variáveis: sexo e escolaridade. Em geral, nas relações assimétricas ocorre a variação *você/ o senhor* e *tu* aparece quando o tratamento é dado de superior para inferior; já nas relações simétricas são usados *tu*, *você* e *o senhor*.

Soares (1980) encontrou em Fortaleza um sistema ternário com *tu*, *você* e *senhor* com usos variáveis. Os fatores que condicionam a variação são: situação discursiva, papel social dos interlocutores, idade e grau de intimidade. É comum a omissão de tratamento na posição de sujeito, compensada pelo uso de vocativos.

A forma *você* é a preferida em situações mais formais e o uso do *tu* é generalizado, mas sua concordância é variável, motivada por fatores como escolaridade, formalidade e atenção prestada à fala.

1.4.2.3 Bezerra (1994)

Bezerra (1994) analisa a fala espontânea de 19 crianças da Paraíba com idades entre 6 e 12 anos, pertencentes à classe média, com acesso aos bens culturais valorizados pela classe social a que pertencem. A coleta dos dados foi feita em 1992, no condomínio onde moram as crianças. Seu objetivo foi analisar a co-ocorrência de *tu* e *você*. Foram considerados os pronomes *tu* implícito e explícito e o *você* implícito e explícito na flexão verbal.

Na sua explicação sobre a importância do princípio de polidez, Leech (1986 apud BEZERRA, 1994, p. 100) apresenta diferentes funções ilocucionárias necessárias para sua análise:

- a) os atos competitivos – cujo objetivo ilocucionário compete com o objetivo social (ex: ordem);
- b) os atos de convivência – cujo objetivo ilocucionário coincide com o social (ex: cumprimento);
- c) os atos colaborativos – cujo objetivo ilocucionário se opõe ao social (ex: afirmação);
- d) os atos conflitivos – cujo objetivo ilocucionário é indiferente ao social (ex: ameaça).

O trabalho de Bezerra (1994) foi baseado em análise sociolinguística interacional e não quantitativa, tratando da questão da polidez e da preservação de faces, não dando prioridade aos dados estatísticos. A autora enfoca a predominância do *tu* sobre o *você*, pois 69% das ocorrências se referem ao *tu*, contra 31% de ocorrências de *você*. O uso de uma ou outra variante parece estar relacionado ao tipo de ato comunicativo envolvido na interação – pedido, ordem, pergunta, desculpa, insulto, consentimento, entre outros.

Os resultados de sua pesquisa apontam para o fato de que a variável sexo não influencia o fenômeno em análise, pois não há diferença no uso de *tu/você* na fala de meninos e meninas. Os resultados mostram que 71% das ocorrências entre as meninas são de *tu* e 65% de *você* entre os meninos. Observemos os resultados quantitativos na Tabela 4.

Tabela 4: Efeito do gênero sobre o uso do *tu* em Bezerra (1994)

| SEXO | FREQUÊNCIA / % |
|------------------|------------------|
| MASCULINO | 66 / 65% |
| FEMININO | 188 / 70% |
| TOTAL | 254 / 69% |

Fonte: Bezerra (1994, p. 115) (adaptada).

Por conseguinte, a pesquisadora aponta para o fato de que o *tu* prevalece em situações de intimidade e o *você* tende a ocorrer em situações de não-solidariedade, quando o falante ameaça a face do ouvinte.

1.4.2.4 Herênio (2006)

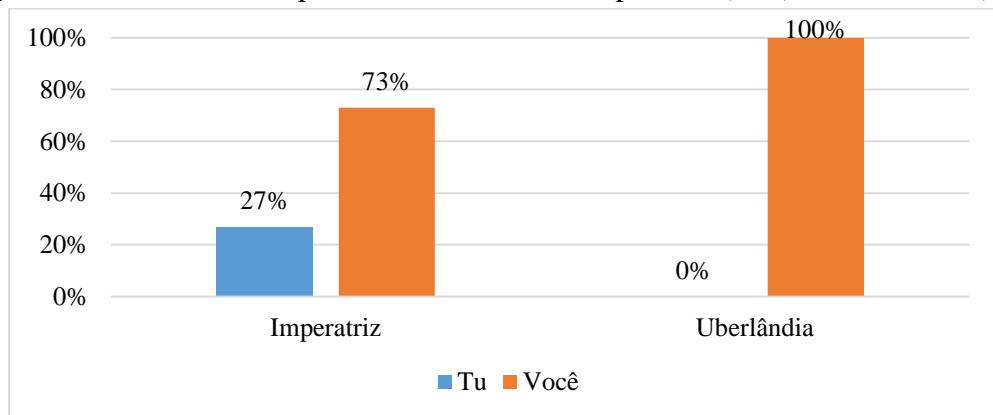
Herênio (2006) estudou a variação *tu/você* na fala de adultos das cidades de Uberlândia (MG) e Imperatriz (MA), analisando fatores linguísticos e extralinguísticos presentes nos discursos desses entrevistados. Para tanto, usou dois *corpora*, constituídos de 43 entrevistas dadas por informantes de Uberlândia e 43 de Imperatriz, totalizando 1.059 dados. Cada entrevista teve a duração de uma hora.

A pesquisadora defendeu como hipóste inicial que em Imperatriz o pronome *tu* apresentava alta frequência, alternando com o pronome *você* e que em Uberlândia o *tu* não ocorria, mas apresentava “resquícios da segunda pessoa caracterizados por correferentes como *te*, *ti*, *teu* etc.” (HERÊNIO, 2006, p. 58).

Em relação à faixa etária, a autora adotou a seguinte distribuição dos informantes, sendo cinco para cada faixa: 20 a 30 anos, 31 a 45 anos e acima de 45 anos, dos sexos masculino e feminino.

Mostramos os resultados de frequência do *tu/você* obtidos por Herênio (2006) para as cidades de Imperatriz e Uberlândia na Figura 2:

Figura 2: Ocorrência dos pronomes *tu/você* em Imperatriz (MA) e Uberlândia (MG)



Fonte: Herênio (2006, p. 69) (adaptada).

Conforme a Figura 2, notamos que Imperatriz apresenta 27% de uso do pronome *tu* e 73% de *você*, enquanto Uberlândia apresenta 0% de ocorrência do pronome *tu* e 100% de *você*. O resultado para Imperatriz contrariou a hipótese inicial da pesquisadora, que esperava uma alta frequência desse pronome no município. Já o resultado para Uberlândia confirmou a hipótese inicial da pesquisadora.

1.4.2.5 Almeida (2012)

Na pesquisa intitulada *Urbanização, escolarização e variação linguística em Feira de Santana-Bahia (século XX)*, o *corpus* utilizado para a análise de Almeida (2012) foi constituído por 45 cartas pessoais escritas em Feira de Santana-BA durante as décadas de 1940, 1950 e 1960 do século XX. A amostra foi constituída por pessoas de ambos os sexos, com escolaridade

diferenciada: alfabetizada, ensino fundamental completo e apenas duas delas eram estudantes universitárias. A amostra é:

- i) uma carta de M. P. A. de M., da década de 1940 do século XX – informante nascida em 08/05/1882; alfabetizada – primeiras letras – nascida na zona rural de Feira de Santana;
- ii) uma carta de A. A. de M., da década de 1940 do século XX – informante nascida em 15/11/1890 – alfabetizada – primeiras letras – nascida na zona rural de Feira de Santana;
- iii) uma carta de P. A. de M., da década de 1940 do século XX – informante nascida em 03/02/1894 – primeiras letras – nascida na zona rural de Feira de Santana;
- iv) 20 cartas de M. M., nascida em 28/08/1924 – informante com ensino fundamental completo (antigo ginásio) – nascida na zona rural de Feira de Santana – moradora da zona urbana;
- v) 20 cartas de D. P., nascida em abril de 1922 na Baixa Grande Bahia – moradora de Feira de Santana desde adolescente, informante com ensino fundamental completo;
- vi) duas cartas de M. S., nascida na zona urbana de Feira de Santana – estudante universitária.

A autora mostrou processos de variação e mudança ocorridos durante os períodos analisados e apresentou alguns elementos sobre o processo de constituição da língua portuguesa falada na Bahia, notadamente na microrregião de Feira de Santana, mais especificamente sobre aspectos que possam ter contribuído para o processo de formação da variedade linguística feirense ao longo do século XX.

Almeida (2012) analisa a coexistência das formas nominativas *você* e *tu* no espaço linguístico feirense, bem como as formas acusativas *te/lhe* e a variação entre pronomes possessivos *teu/tua* e *seu/sua*, com o objetivo de mostrar a articulação entre esses fenômenos dentro do quadro de variação da 2^a pessoa do século XX.

Um dos objetivos perseguidos na pesquisa de Almeida (2012, p. 14) e que deu origem ao trabalho foi tentar verificar se o uso de *tu/você* em Feira de Santana significa a manutenção, na oralidade, dessa antiga variação ou se, ao contrário, seria uma recuperação ocorrida nas décadas de 1970 e 1980 por conta da grande migração em condições diferentes da vivida em meados do século. Essa migração, ocorrida na região, teria levado a uma maior interação entre falantes cultos e populares, hipótese que ainda não pode ser confirmada só com esse estudo.

A Tabela 5 apresenta os resultados encontrados na análise das cartas pessoais escritas nas décadas de 1940, 1950 e 1960.

Tabela 5: Índice geral de uso de *tu* e *você* em cartas particulares escritas em Feira de Santana (décadas de 1940, 1950 e 1960)

| USO DO <i>TU</i> | USO DO <i>VOCÊ</i> |
|------------------|--------------------|
| 43/199 | 156/199 |
| 22% | 78% |

Fonte: Almeida (2012, p. 14).

Verificando os resultados apresentados, percebe-se que já havia, nas décadas em pauta, uma inclinação pelo uso do *você*. Nessa direção, a autora tenta relacionar o desaparecimento do pronome pessoal *tu* nas cartas pessoais produzidas na década de 1960 e o reaparecimento e/ou continuidade do uso, até os dias atuais, na oralidade, por influência dos dialetos nortistas e, em menor escala, dos dialetos rurais baianos, isto é, do processo de urbanização da cidade, uma vez que esse município reúne falares de diferentes localidades do Brasil.

Tabela 6: Variação entre *tu* e *você* de acordo com a década e o sexo

| PRONOME SUJEITO | MULHERES | | HOMENS | |
|---------------------------|-----------|-------------|-----------|-------------|
| | <i>Tu</i> | <i>Você</i> | <i>Tu</i> | <i>Você</i> |
| DÉCADA | | | | |
| DÉCADA DE 1940 | 6/6 | 0/0 | 0/0 | 0/0 |
| | 100% | 0% | 0% | 0% |
| DÉCADA DE 1950 | 25/44 | 19/44 | 12/70 | 58/70 |
| | 57% | 43% | 17% | 83% |
| DÉCADA DE 1960 | 0/0 | 67/67 | 0/0 | 12/12 |
| | 0% | 100% | 0% | 100% |

Fonte: Almeida (2012, p. 15).

Anunciando o que está em evidência na Tabela 6, o *tu* é majoritário na década de 1940, quando se trata do sexo feminino. Já na década de 1950, há equilíbrio no uso das formas entre as mulheres, enquanto na fala dos homens há um uso muito maior da forma *você*, o que pode indicar uma possível mudança.

Verificamos, assim, com a análise da variação entre os pronomes *tu* e *você*, que, a partir da década de 1960, o *tu* não mais aparece nas cartas. “O mais interessante nessa questão é que os autores das cartas são os mesmos, o que pode nos levar a inferir que houve alguma influência externa mais forte que os levou a não usarem mais a forma *tu*” (ALMEIDA, 2012, p. 16).

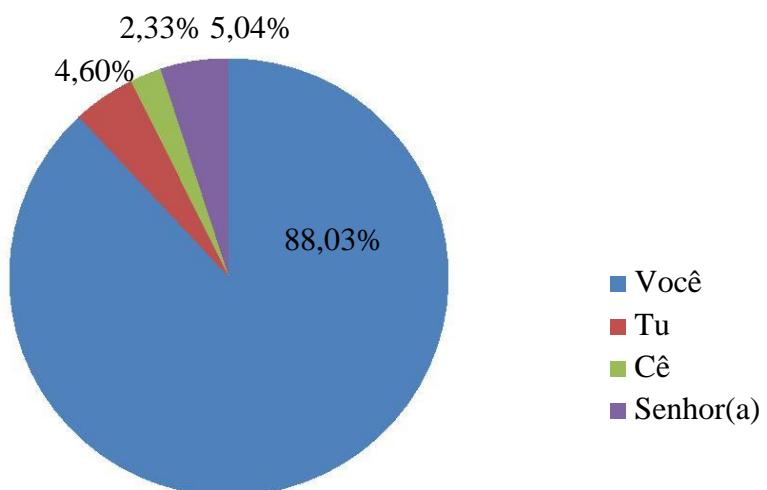
Por outro lado, a autora salienta que, com base tão somente nas cartas pessoais, não se pode afirmar que essa forma pronominal desapareceu completamente do espaço linguístico feirense nessa fase.

1.4.2.6 Nogueira (2013)

A pesquisa de Nogueira (2013) trata da variação entre as formas de tratamento *tu* e *você* no português culto e popular das cidades de Feira de Santana e Salvador. A amostra foi constituída por 48 gravações de entrevistas de aproximadamente 40 minutos cada, do tipo DID (Diálogo entre Informante e Documentador), sendo que 12 pertenciam ao Projeto Norma Linguística Urbana Culta de Salvador (NURC/SSA); 12 eram oriundas do Programa de Estudos sobre o Português Popular em Salvador (PEPP); 24 pertencentes ao Projeto A Língua Portuguesa no Semiárido Baiano; e sete conversações espontâneas entre informantes de Feira de Santana, com o objetivo de verificar a utilização das formas de tratamento nesse gênero do discurso. Os informantes foram estratificados de acordo com sexo, faixa etária (25-35; 36-55; mais de 56 anos) e escolaridade (ensino fundamental e médio), em uma amostra balanceada.

Nas conversações espontâneas, foram gravadas duas conversas, com cinco homens (três de 25 a 35 anos; dois de 45 a 55 anos) e sete mulheres (quatro entre 25 e 35; três entre 45 e 55 anos).

Figura 3: Totais de referência à segunda pessoa nos *corpora* analisados



Fonte: Nogueira (2013, p. 86).

A totalidade de ocorrências dos pronomes *você*, *tu*, *senhor(a)* e *cê* foi de 1804 dados. Houve 1588 ocorrências de *você* (88,03%), 83 ocorrências de *tu* (4,6%), 91 ocorrências de *senhor(a)* e 42 ocorrências da forma reduzida de *cê*.

Segundo a autora, um ponto importante a ser destacado é que a forma considerada mais formal *senhor(a)* chega a superar os dados de *tu*, com 5,04% dos dados, contrariando as expectativas da mesma.

Em relação às gravações espontâneas, Nogueira (2013) escolheu contextos nos quais pudesse captar conversas descontraídas, tais como salão de beleza, reunião entre amigas, visita a um colega, bate-papo entre colegas. A autora teve o cuidado de gravar apenas conversas cujos participantes eram naturais de Feira de Santana, ou que nesta cidade chegaram até os cinco anos de idade, seguindo, assim, os critérios do Projeto NURC. Os dados complementares para a pesquisa em pauta justificam-se para ilustrar as ocorrências do pronome *tu* em Feira de Santana em contextos mais descontraídos. O resultado dessa amostra mostrou-se diferente dos colhidos nos DIDs. Dos 147 dados, 57,8% são do *você* e 42,2% do *tu*. Quanto ao sexo, os homens usam mais o *você* (66,7%), enquanto as mulheres preferem o *tu* (50,5%).

Quanto à função sintática do sintagma nominal (SN), a autora adotou a categorização proposta por Ramos (1997). De um lado, analisou as ocorrências das variantes na função sujeito e, do outro, as ocorrências de objeto de verbo e objeto de preposição. Por seu turno, os resultados mostraram que *você* aparece mais como sujeito (97,8%), enquanto o *tu* aparece mais como não sujeito (19,6%). Outro fator analisado foi o tipo de frase: as declarativas favorecem o uso do *você* (95,4%) e as não declarativas favorecem o *tu* (9,2%), negando, portanto, a hipótese da pesquisadora de que as frases interrogativas favoreceriam o uso do *tu*.

Em relação à variável tempo verbal, o que se pode depreender dos dados é que *você* é favorecido pelo tempo passado (98,6%) e o *tu*, pelo não passado (5,15%). No discurso direto, há predomínio do *você* (96,1%) e no discurso relatado, do *tu* (8,5%). Já quanto ao tipo de referência, a específica favorece o *tu* (14%), enquanto a genérica beneficia o *você* (100%). Não houve concordância canônica com o pronome *tu*. A autora conclui que o gênero do discurso é uma das variáveis mais relevantes para a análise da alternância entre as formas de tratamento *tu/você*.

1.4.2.7 Guimarães (2014)

Guimarães (2014) analisa o comportamento variável das formas de tratamento pronominais *tu*, *você*, *cê* e *o(a) senhor(a)* e das formas nominais *macho*, *mulher*, *rapaz*, *minha filha*, *cara* e outras no português oral popular de Fortaleza. Para a constituição do *corpus*, a autora utilizou os inquéritos do projeto Norma Oral do Português Popular de Fortaleza (NORPOFOR). A distribuição dos informantes no *corpus* foi feita em função das variáveis sociais controladas no projeto, a saber: sexo (homens e mulheres), faixa etária (I: 15 a 25 anos, II: 26 a 49 anos e III: a partir dos 50 anos), escolaridade (A: 0 a 4 anos, B: 5 a 8 anos, e C: 9 a 11 anos) e tipo de registro (Diálogo entre Informante e Documentador: DID, Diálogo entre Dois Informantes: D2 e Elocução Formal: EF).

Na pesquisa em questão, as variáveis controladas foram: sociais (faixa etária, sexo e escolaridade); interacionais (grau de simetria entre os interlocutores e grau de intimidade); e linguísticas (tipo de referente, estrutura do verbo, posição em relação ao verbo, paralelismo, efeito gatilho, tipo de verbo, tipo de fala, tempo verbal, função sintática do pronome, tipo de entonação, tipo de relato, concordância com o verbo e polaridade da sentença).

Observando os dados na sua totalidade, com o auxílio do GoldVarb X, Guimarães (2014) obteve os seguintes resultados para as formas pronominais: *tu* com 47,2%, *você* com 46,5%, *o senhor(a)* com 4,3%, *cê* com 2% e *ocê* com 0,1%. E para as formas nominais: *macho* com 39%, seguido de *mulher* com 26,9% e *rapaz* com 17%, *cara* com 5,2%, *minha filha* com 4,3%, *amiga* com 3,3%, *meu amigo* com 1,5%, *menina* com 1,3%, *meu filho* com 1,1%, *meu irmão* com 0,2% e *menino* somente com 0,1%. Percebe-se que os pronomes mais frequentes são o *tu* e o *você*; os mesmos aparecem praticamente empatados. O *cê* dá indícios de aparecer no falar fortalezense, mas ainda de forma muito tímida, enquanto o *ocê*, que só apareceu uma vez, foi excluído, por esse motivo, da análise.

A pesquisadora salienta que as formas de tratamento nominais, assim como o *tu* e o *você*, são mais produzidas nas relações de solidariedade, mas o mesmo não acontece com *o(a) senhor(a)*. As formas *rapaz* e *você* seriam consideradas neutras. O uso do *cê* ainda é muito escasso, já o *ocê* praticamente não ocorre. Observou-se também que o uso de *cara*, *macho* e *rapaz* não se refere apenas a homens e que todas as formas em estudo não são estigmatizadas pela comunidade em questão.

De acordo com a autora, dos 17 grupos de fatores controlados na pesquisa, 12 foram selecionados pelo programa. Foram eles, por ordem decrescente de relevância: entonação, tipo

de fala, tipo de referente, faixa etária, sexo/gênero, grau de interação no discurso, tipo de relato, tipo de verbo, função sintática, paralelismo formal, escolaridade e tempo verbal. Os grupos excluídos foram somente cinco: posição em relação ao verbo, estrutura do verbo, polaridade, grau de simetria entre os informantes e efeito gatilho. Enfim, a variável selecionada em primeiro lugar foi a entonação, com as frases interrogativas (76%) favorecendo o uso do *tu*, ao passo que as frases declarativas e/ou exclamativas (43,5%) inibem a utilização do pronome *tu*.

Outra variável selecionada bastante interessante foi o tipo de fala, que se refere aos assuntos abordados pelos interlocutores na amostra em questão.

Tabela 7: Atuação do tipo de fala sobre o pronome *tu* (*tu x você*)

| OBSERVAÇÕES | APLIC./TOTAL | % | P.R. |
|--|--------------|------|-------|
| IRÔNICAS/BRINCADEIRAS | 48/53 | 90,6 | 0,910 |
| FOFOCA | 25/40 | 62,5 | 0,655 |
| CONVERSA SOBRE RELACIONAMENTO AMOROSO | 59/93 | 63,4 | 0,587 |
| RECORDAÇÕES | 40/99 | 41,5 | 0,579 |
| CONVERSA CASUAL | 541/1057 | 51,2 | 0,472 |
| CONVERSA SOBRE TRABALHO | 68/161 | 42,2 | 0,457 |
| RELIGIÃO | 10/48 | 20,8 | 0,218 |
| REPREENSÃO | 1/4 | 25 | 0,167 |

Fonte: Guimarães (2014, p. 14).

Considerando a Tabela 7, chega-se à conclusão de que as observações irônicas;brincadeiras (0,910) surgem como o fator que mais influencia positivamente o uso do *tu*. Embora com valores menos expressivos, a fofoca (0,655) também favorece esse pronome, assim como as conversas sobre relacionamento amoroso (0,587) e as recordações (0,579).

Guimarães (2014) conclui que o *tu* é mais usado em situações mais espontâneas, quando o assunto é íntimo ou só pode ser tratado apenas entre amigos. É usado na maioria das vezes com referente específico. Alude também que, apesar de não ser considerada a forma de prestígio, o *tu* também não é estigmatizado pelo fortalezense. Prova disso é que as mulheres, os mais jovens e os mais escolarizados são os que mais favorecem o seu uso.

1.4.3 Região Centro-Oeste

Sobre a Região Centro-Oeste, apresentamos os trabalhos de Lucca (2005), Dias (2007), Andrade (2010) e Scherre et al. (2015).

1.4.3.1 Lucca (2005)

Lucca (2005) analisa a variação *tu* ~ *você* na fala dos jovens brasilienses. Os dados foram colhidos especialmente para esse estudo em três regiões administrativas do Distrito Federal: Taguatinga, Ceilândia e Brasília. A autora acredita que o *tu* tende a aparecer em relações solidárias em momentos de intimidade. Seguindo os passos de Paredes Silva (2003), optou por fazer a coleta de dados recorrendo às gravações ocultas, em que informantes conversam entre si e um deles porta o gravador. Para esse estudo, foram selecionados somente falantes do sexo masculino na idade de 15 a 19 anos que cursam o Ensino Médio em escolas públicas. A autora trabalhou com sete grupos de fatores linguísticos e cinco grupos de fatores extralinguísticos.

Um dos fatores analisados foi o paralelismo linguístico. A hipótese criada para essa variável era de que a ocorrência dos pronomes *tu* ~ *você* favorecia o aparecimento subsequente do mesmo pronome. Essa hipótese foi confirmada e percebeu-se que a posição que mais favorece o uso do *tu* é aquela em que o pronome aparece como primeiro item da série.

Os resultados apontaram que 72% das referências de segunda pessoa ocorrem com o *tu*, enquanto 28% ocorrem com o *você*. Outra variável que foi analisada foi o tipo de estrutura considerando-se a entoação do falante. A hipótese era a de que as estruturas interrogativas e exclamativas favoreciam o uso do *tu*. Os resultados conseguidos corroboraram a hipótese, pois as estruturas que mais favoreceram o uso do *tu* foram as exclamativas, com peso relativo de 0,87, e as estruturas interrogativas tiveram um peso de 0,54. Já as estruturas declarativas tiveram o peso relativo de 0,33, atestando que desfavorecem o uso do *tu*.

A terceira variável selecionada foi o tipo de relação entre os interlocutores. A hipótese criada por Lucca (2005) era de que o pronome *tu* tenderia a ser usado em um estilo amplamente informal, em relações entre pares solidários³. A hipótese foi comprovada por meio dos resultados: os pares solidários apresentaram um efeito mais forte sobre o uso do pronome *tu*,

³ A solidariedade pode acontecer tanto em relações simétricas quanto em relações assimétricas, no sentido de que os interlocutores, pares ou não, podem estabelecer um vínculo de cooperação entre si.

com peso relativo de 0,57; já os pares em relação não-solidária apresentaram peso relativo de 0,22.

A conclusão a que chegou a autora nesse estudo foi a de que o *tu* é empregado nas três regiões administrativas de Brasília, em maior grau naquela em que os moradores são de procedência nordestina.

1.4.3.2 *Dias (2007)*

Outro estudo sobre a variação das formas *tu* ~ *você* no Centro-Oeste é a pesquisa feita por Dias (2007), cujo objetivo foi analisar a variação *tu* ~ *você* entre falantes brasilienses, descrevendo quais fatores linguísticos e sociais condicionam a variação em diferentes faixas etárias.

As hipóteses levantadas pela pesquisadora foram: a) o uso do pronome *tu* é tão mais frequente quanto mais jovem for o falante; b) na faixa etária com mais de 30 anos de idade, especialmente no caso das mulheres, o emprego do pronome *tu* dá indício de desrespeito; e c) a frequência do uso do *tu* está relacionada ao estilo de vida do falante.

O *corpus* foi constituído por meio de conversas espontâneas em que os falantes, pessoas residentes no Plano Piloto de Brasília, sabiam que estavam sendo gravados. A amostra foi composta por 18 informantes divididos em três faixas etárias. Foram escolhidos três falantes de cada sexo para cada uma das faixas etárias.

Os fatores linguísticos considerados foram: a) tipo de fala (casual, acadêmica, profissionais, brincadeiras, observações irônicas); b) tipo de referência (genérica ou específica); c) função sintática (sujeito, objeto, predicativo); d) forma verbal (presente do indicativo, pretérito perfeito, pretérito imperfeito, futuro do subjuntivo); e e) tipo de relato (fala própria, fala relatada de outra pessoa).

Os fatores sociolinguísticos analisados foram: a) faixa etária; b) sexo; c) estilo de vida do falante; d) relacionamento com o interlocutor (amigo, amigo íntimo, desconhecido, familiar); e) faixa etária do interlocutor; f) sexo do interlocutor; e g) lugar em que ocorreu o diálogo.

Dias (2007), depois de agrupar os dados de *cê* e *você*, obteve a frequência de uso do pronome *tu* de 12,8%, enquanto *cê* ~ *você* atingiram 87,2% dos dados.

Em relação à variável faixa etária do informante, uma das hipóteses levantadas pela autora foi de que o uso do *tu* era mais frequente na faixa etária mais jovem. Essa hipótese foi

confirmada. O uso do *tu* foi favorecido entre os mais jovens, neutralizado na faixa intermediária e desfavorecido entre os mais velhos.

Considerando a variável tipo de fala, a expectativa era a de que ironia, deboches e brincadeiras favorecessem o uso do *tu*, enquanto as conversas profissionais o desfavorecessem, já que as formalidades da situação e do tema sobre o qual se fala são levados em consideração por parte do falante. Porém essa variável não foi selecionada pelo programa VARBRUL.

Em relação ao sexo, a amostra foi formada por nove homens e nove mulheres. A autora levantou a hipótese de que o *tu* seria mais usado por falantes do sexo masculino, já que, segundo alguns estudos sociolinguísticos, as mulheres tendem a usar as variantes de maior prestígio em suas comunidades. O resultado confirma a hipótese, pois a frequência do uso do *tu* pelo sexo masculino foi de 14,9% e pelo sexo feminino foi de 10,8%.

Sobre a variável estilo de vida, a hipótese era a de que as pessoas com o estilo de vida alternativo usariam mais o *tu* do que aquelas com estilo de vida conservador. Essa hipótese foi confirmada. O critério adotado por Dias (2007) para determinar o estilo de vida do falante foi a profissão que o mesmo exerce. Dessa forma, considerou conservadores todos aqueles falantes que ocupam cargos públicos e alternativos os que exercem profissão autônoma ou trabalham em empresa privada.

Outra variável analisada foi o relacionamento com o interlocutor. Foram controlados os seguintes tipos de relacionamento: amigo íntimo, amigo ou colega, conhecido, pessoa da família, desconhecido. O peso relativo de amigo íntimo/familiar foi de 0,6 favorecendo o uso do *tu* quando o falante se dirige a pessoas com quem tem esse tipo de relacionamento.

Analizando os trabalhos de Lucca (2005) e Dias (2007), percebemos que eles, embora da mesma região, se distanciam no que tange à frequência de uso do pronome *tu*. Encontramos em Lucca (2005) 78% para o uso do *tu* e 28% para o uso do *você*, enquanto Dias (2007) apresenta para o pronome *tu* 12,8% e para *cê/você*, 87,2%.

No estudo de Dias (2007), o pronome *tu* ocorre na fala de informantes dos dois sexos e de todas as faixas etárias, porém é mais frequente entre falantes do sexo masculino mais jovens, corroborando os resultados encontrados por Lucca (2005). Com relação à idade dos interlocutores, vemos em Dias (2007) que o uso de *tu* é mais comum entre falantes da mesma faixa etária.

Em relação ao fator intimidade e distanciamento, os dois estudos se aproximam: o relacionamento que mais favorece o uso do *tu* é o de amigo íntimo ou familiar no estudo de Lucca (2005) e na pesquisa de Dias (2007).

1.4.3.3 Andrade (2010)

Andrade (2010) investiga as formas *você*, *cê* e *tu* no falar de crianças e adolescentes de Brasília, com o objetivo de verificar se, já nos primeiros anos, as crianças usariam o *tu*, que, em Brasília, se configura como variante inovadora. A amostra é composta por alunos, com idade entre 7 e 15 anos, de uma escola pública situada em Vila Planalto. São 43 informantes, 25 do sexo feminino e 18 do sexo masculino. Dessa seleção, nove não pertencem à comunidade de Vila Planalto (três meninas e um menino do Lago Sul; dois meninos da Asa Norte; duas meninas e um menino do Sudoeste). A análise dos dados foi baseada nos seguintes grupos de fatores:

- a) Variáveis extralinguísticas (ou externas): faixa etária, tipo de relação (simétrica *versus* assimétrica), localidade (região administrativa), origem dos pais, tipo de fala (relatada *versus* original);
- b) Variáveis linguísticas (ou internas): função sintática, presença ou ausência de vocativo, entonação (frases interrogativas *versus* não interrogativas), polaridade da sentença (afirmativa *versus* negativa), referencialidade (genérica *versus* específica), paralelismo formal.

A autora controlou a faixa etária, uma vez que esse fator já influenciou o fenômeno em análise em outros estudos do mesmo tipo. Ela escolheu, em especial, a faixa etária 7 a 15 anos, com a intenção de testar em que idade a variante *tu* entra para o repertório linguístico do brasiliense.

Tabela 8: Tipo de relação entre os interlocutores (dados da Vila)

| TIPO DE RELAÇÃO | VOCÊ | CÊ | TU |
|--------------------|------|------|------|
| SIMÉTRICA | 0,27 | 0,26 | 0,47 |
| ASSIMÉTRICA | 0,38 | 0,40 | 0,22 |

Fonte: Andrade (2010, p. 85).

Submetidos os dados ao GoldVarb X, realizou-se uma rodada eneária, o que impossibilitou o programa de efetuar a seleção dos grupos de fatores. Das 835 ocorrências, o pronome mais utilizado é o *tu* (48%).

A rodada eneária, além de confirmar a hipótese sobre o favorecimento do *tu* em relações simétricas, também revelou o favorecimento da variante *cê* em relações assimétricas. Houve também um favorecimento da variante *você* nas relações assimétricas, semelhante ao *cê*.

Em relação à variável origem dos pais das crianças estudadas, os números relativos sugerem que a origem que mais favorece a ocorrência de *tu* é a origem nordestina de ambos os

pais, com 63%, portanto acima da média de 48%. A ocorrência de *tu* também ficou acima da média, de forma surpreendente, quando os pais eram originários de Minas Gerais e do Nordeste. A autora acredita que o *tu* na fala brasiliense tem forte indício de ser proveniente do Nordeste.

Como se pode depreender dos dados, de forma geral, a variante *tu* é favorecida pelo sexo masculino. O *você* se destacou na fala feminina (com 34%) e em referências genéricas (com 64,2%), enquanto o *tu* obteve seu maior uso na referência específica, com 37,5%. Nas construções que não possuem verbos, o *você* também teve predominância, com 58,9%, conforme Tabela 9, com todos os dados da amostra.

Tabela 9: Variação *você/cê/tu* em relação à função sintática

| | VOCÊ | CÊ | TU |
|------------------|----------------------|-----------------|-----------------|
| SUJEITO | 285/779 = 36,6% | 224/779 = 28,8% | 270/779 = 34,7% |
| SEM VERBO | 33/56 = 58,9% | 5/56 = 8,9% | 18/56 = 32,1% |
| TOTAIS | 318/835 = 38,1% | 229/835 = 27,4% | 288/835 = 34,5% |

Fonte: Andrade (2010, p. 102).

Os resultados gerais indicam que as orações interrogativas favorecem relativamente as variantes *cê* (35,4%) e *tu* (46,7%) de forma semelhante, ficando ambas com os percentuais acima de suas médias. Por sua vez, a variante *você* foi favorecida pelas orações não interrogativas (45,8%).

A faixa etária mostrou-se relevante em todas as variantes. Para o *tu*, os resultados para as faixas etárias são: de sete a 11 anos = 27% e de 12 a 15 anos = 58%. Ou seja, a faixa maior de idade utiliza mais o *tu*. Por outro lado, as crianças de sete a 11 utilizam mais o *cê* (40%).

Na sua conclusão, a autora salienta que o *tu* em Brasília é um *tu* ainda não nativo, como o é nas regiões Sul e Norte; há em Brasília um estágio de focalização dialetal, que torna o jeito de falar brasiliense distinto dos demais. Alerta também que o *tu* tem ganhado cada vez mais espaço no repertório linguístico do brasiliense.

1.4.3.4 Scherre et al. (2015)

No artigo intitulado “Variação dos pronomes *tu* e *você*”, Scherre et al. (2015) apresentam resultados globais de diversos trabalhos sociolinguísticos sobre o uso de pronomes de segunda pessoa: *tu* e *você*. Segundo as autoras, a primeira razão do interesse pelos pronomes decorreu da análise da provável relação entre a variação gramatical do imperativo e os

pronomes em questão. A segunda razão foi decorrente das orientações feitas às dissertações e teses sobre o tema.

O estudo de Scherre et al. (2015) foi resultado da coletânea de 60 amostras diversificadas: 41 amostras de entrevistas sociolinguísticas; 12 amostras de conversas naturais estimuladas e não estimuladas, ocultas e não ocultas; seis amostras de entrevistas geolinguísticas e uma amostra de conversa estimulada por gravuras.

A metodologia mais recorrente, utilizada para essas diferentes pesquisas, foi a coleta de dados em situação de entrevista sociolinguística, com apresentação da gravura e com a presença de um entrevistado e um ou dois entrevistadores, não necessariamente naturais da comunidade local.

Sobre as entrevistas sociolinguísticas, as autoras destacam que as pesquisas variacionistas brasileira têm como objetivo explícito captar a fala espontânea da pessoa entrevistada. Por outro lado, “essas entrevistas podem não refletir toda a realidade dos usos dos pronomes de segunda pessoa, por, muitas vezes, evitarem constantes trocas de turno e/ou por não terem naturalmente toda a gama das relações interacionais” (SCHERRE et al., 2015, p. 135).

É importante salientar que os trabalhos analisados por Scherre et al. (2015) são de natureza heterogênea, o que pode acarretar em problemas para o alcance dos seus objetivos, uma vez que eles serão comparados e, em seguida, sistematizados.

Segundo as pesquisadoras, mesmo com todos os problemas inerentes ao registro dos pronomes de segunda pessoa, tem-se a oportunidade de ver que o pronome *tu* é mais difícil de captar em áreas em que ele não parece ser a primeira forma da comunidade, como, por exemplo, no Rio de Janeiro e em Brasília. Já no Rio Grande do Sul, o *tu* surge com facilidade e naturalidade, mesmo em entrevistas sociolinguísticas.

Para a proposta de síntese do mapeamento, as formas utilizadas para comparação foram *tu*, *você*, *cê* e *ocê* nas suas formas singulares. As autoras trabalham com a hipótese de que os pronomes em questão não têm o mesmo *status* de naturalidade e/ou de indicação de simetria/assimetria, ou seja, qualquer uma das quatro formas arroladas pode ter emprego diferenciado, a depender da área geográfica, como se pode perceber nas pesquisas em questão.

As autoras apresentam seis subsistemas, a saber:

- 1- Subsistema **só você**
- 2- Subsistema **mais tu com concordância baixa**
- 3- Subsistema **mais tu com concordância alta**

- 4- Subsistema ***tu/você* com concordância baixa**
- 5- Subsistema ***tu/você* com concordância média**
- 6- Subsistema ***você/tu***

Os resultados para o subsistema **só você** apontam que esse uso é supra-regional, porém se concentra na área central do país com bastante uniformidade. Atendendo à utilização somente do pronome *você* encaixam-se os Estados de Goiás, Mato Grosso e Mato Grosso do Sul, na Região Centro-Oeste, com exceção de Brasília. A Região Sudeste é representada pelos Estados do Espírito Santo, Minas Gerais (exceto a cidade de São João da Ponte) e São Paulo (exceto Santos). A Região Nordeste é representada pela Bahia, basicamente pela sua capital, Salvador; a Região Norte, pelo Estado de Tocantins; e a Região Sul, pelo Estado do Paraná.

O subsistema **mais *tu* com concordância baixa** aparece nas regiões Norte (Amazonas) e Sul (Rio Grande do Sul).

O subsistema **mais *tu* com concordância alta** é documentado na Região Norte (Pará) e na Região Sul (Santa Catarina).

O subsistema ***tu/você* com concordância baixa** é encontrado nas regiões Nordeste e Sul. Os Estados onde o fenômeno ocorre no Nordeste são Maranhão e Tocantins; e na região Sul, Santa Catarina.

O subsistema ***tu/você* com concordância média** aparece nas regiões Nordeste, Norte e Sul. Na Região Nordeste, contamos com os seguintes Estados: Maranhão, Piauí, Ceará, Paraíba e Pernambuco. Na Região Norte, o Estado representativo é o Amazonas; e, na Região Sul, Santa Catarina.

O subsistema ***você/tu*** sem concordância somente não aparece na Região Sul. Já na Região Centro-Oeste aparece no Distrito Federal; na Região Sudeste é representado pelos Estados do Rio de Janeiro, São Paulo e Minas Gerais; na Região Nordeste, pelos Estados do Maranhão e da Bahia; e na Região Norte, pelos Estados de Roraima e Acre.

As autoras concluem afirmando que o pronome *tu* pode ser de uso extremamente natural, como índice de identidade de algumas comunidades, com possibilidades de concordância, em maior ou menor grau, dependendo da formalidade da interação ou do grau de escolaridade do falante. São exemplos disso as localidades de Santa Catarina, Rio Grande do Sul, Amazonas, Pará e Maranhão.

Em contrapartida, a utilização do pronome *tu* pode ser marca interacional, sem absoluta possibilidade de concordância, como foi encontrado no Rio de Janeiro, em Santos e em algumas localidades da Bahia, e, recentemente, em Brasília.

As autoras acrescentam ainda que as formas *cê* e *ocê*, mais características do subsistema **só você**, apresentam suas peculiaridades: a forma *ocê* é marca mineira e/ou de ruralidade; já a forma *cê*, por seu turno, foi facilmente encontrada em todas as localidades consideradas nesse estudo.

1.4.4 Região Sudeste

Sobre a Região Sudeste, apresentamos os trabalhos de Salles (2001), Paredes Silva (2003) e Modesto (2006) e Silva (2017).

1.4.4.1 Salles (2001)

Salles (2001) trabalhou com os pronomes de tratamento na tentativa de descrever a evolução de cada um, vendo o percurso diacrônico dos mesmos. A amostra foi composta por documentos informais produzidos na cidade de São Paulo no século XIX. Eram, sobretudo, cartas pessoais selecionadas com a finalidade de preservar a naturalidade e espontaneidade do emissor na escrita. Foram analisadas também cartas oficiais, comerciais, de propaganda e correspondências dirigidas a seções de jornais ou revistas.

O autor elegeu algumas variáveis sociais, a saber: gênero do emissor, papel social do emissor, simetria e assimetria na interlocução e outras. Considerou as variáveis internas: função referencial, tratamento em 2^a pessoa, tratamento em 3^a pessoa, fator repetição e papel social dos emissores, dentre outras.

O *corpus* em análise registra 425 ocorrências de pronomes de tratamento em segunda pessoa, incluindo o paradigma *tu* e *vós*, documentando 75% de utilização do *tu*.

Em relação à variável externa grau de aproximação\distanciamento entre os interlocutores, foi feita uma comparação entre cartas cujos emissores eram religiosos e cartas cujos emissores eram amigos\familiares. A situação interlocutiva dos religiosos sugere maior distanciamento entre os interlocutores e, por outro lado, a situação interlocutiva dos familiares\amigos leva a maior proximidade.

Examinando os tratamentos em 2^a pessoa, o favorecido é o pronome *tu*. Analisando o grau de aproximação – distanciamento, o tratamento em 2^a pessoa pela forma *tu* foi predominante na interlocução de maior grau de aproximação, caracterizando um estilo mais

informal. Destacam-se também as relações simétricas, entre iguais, onde o *tu* predomina com um percentual de 89%.

1.4.4.2 Paredes Silva (2003)

Paredes Silva (2003), estudando a fala carioca, pesquisou sobre a variação *tu/você*. Em um primeiro momento, analisou amostras de fala pertencentes ao Projeto de Estudos sobre os Usos da Língua – PEUL/UFRJ, um conjunto de 64 entrevistas gravadas no início da década de 1980 que fazem parte do acervo desse Projeto que, por sua vez, constituiram o Projeto Censo da Variação Linguística no Rio de Janeiro. Das entrevistas utilizadas, inicialmente, houve apenas 10 ocorrências do pronome *tu*, estando estas resumidas a apenas cinco falantes – 10% da amostra examinada. A autora acreditou que o uso do *tu* foi inibido devido ao fato de os falantes saberem da existência da gravação, ou seja, o gênero do discurso pode ter interferido, bem como ao fato de o material ser de uma época na qual o pronome *tu* ainda não era difundido, início dos anos oitenta.

O interesse da autora era mostrar a mudança que parecia estar em curso no português brasileiro, mais precisamente no falar carioca – o retorno do *tu*, disputando espaço com o *você*. O estudo da autora teve início em cartas pessoais com escrita informal. Nesse *corpus*, ela observou o uso do *você* (com 70%) e do sujeito nulo (com 30%). A faixa etária não se mostrou relevante para a escolha dos pronomes nesse trabalho. Enfim, na década de 1990, observa-se a mudança: o (re)aparecimento do *tu*.

A pesquisadora, em um segundo momento, constituiu outro *corpus* de cunho oral, buscando captar a fala em situações mais naturais e menos controladas. Essas gravações pertencem ao Banco de Dados Interacionais – BDI, também integrante do acervo PEUL. As pessoas que fizeram parte da gravação eram conhecidas e, mesmo assim, colheu-se um número insignificante de ocorrências; foram encontradas apenas cinco ocorrências da variante *tu*, em um total de seis entrevistas analisadas. A autora explica que parte desse material foi adquirida com a permissão prévia dos participantes, o que pode tê-los inibido no uso de formas não padrão, como o pronome em estudo.

Depois de ter suas expectativas frustradas, Paredes Silva (2003) constituiu um novo *corpus* (intitulado Paredes 96) e usou como estratégia um minigravador oculto (só depois é que as pessoas souberam das suas falas gravadas) a fim de captar momentos de interações sem monitoramento, bastante espontâneos. A amostra foi constituída por oito gravações de cerca de

20 minutos cada, mostrando-se bastante relevante para a comprovação do uso da forma *tu* em contextos mais informais.

Os resultados encontrados nos três *corpora* estão apresentados na Tabela 10.

Tabela 10: Distribuição dos pronomes de 2^a pessoa nos *corpora* investigados

| | CENSO – PEUL | BDI – PEUL | PAREDES 96 |
|--------------|--------------|------------|------------|
| VOCÊ | 644 = 94% | 168 = 97% | 133 = 35% |
| TU | 42 = 6% | 5 = 3% | 235 = 65% |
| TOTAL | 686 | 173 | 368 |

Fonte: Paredes Silva (2003, p. 164) (adaptada).

Na Tabela 10, percebe-se que o pronome *você* foi majoritário nos dados colhidos por Paredes Silva (2003) na amostra CENSO, em 1980, com 94%, bem como na amostra do BDI, com 97% de favorecimento. Por outro lado, os dados de 1996 mostram que o pronome *tu* foi mais utilizado do que o *você*. Foram 368 referências pronominais ao sujeito de 2^a pessoa, das quais 235 foram do uso do *tu* (65%) e 133 ocorrências do pronome *você* (35%).

Verificamos, portanto, uma grande diferença na distribuição das variantes nos *corpora*, o que se justifica pelos diferentes modelos de entrevistas utilizados para a composição das amostras.

1.4.4.3 Modesto (2006)

Modesto (2006) analisou as formas de tratamento no português falado da cidade de Santos. A amostra foi formada por 10 gravações secretas e 10 gravações não-secretas (consentidas previamente), buscando explicitar até que ponto as diferentes situações interacionais levam o falante a escolher uma ou outra forma pronominal. Ele procurou verificar quais os grupos de fatores que podem influenciar na alternância de uso das formas *tu* e *você*. Selecionou, portanto, três fatores sociais – gênero, faixa etária e escolaridade; um fator linguístico – função sintática da forma de tratamento; e dois fatores discursivo-pragmáticos – referenciamento e monitoramento. O autor adotou a perspectiva funcionalista de análise, valendo-se do suporte metodológico da Sociolinguística Laboviana.

No que tange à faixa etária, Modesto (2006) estratificou como primeira (de 15 a 20 anos) e segunda (21 anos ou mais). Em relação à escolaridade, selecionou falantes do ensino médio e ensino universitário.

Os dados gerais sugerem que *você* é a forma preferida dos santistas, com 67%, enquanto o uso do *tu* restringiu-se a 32%. Em relação à variável escolaridade, constatou que os informantes com maior nível de escolaridade tendem a usar menos a forma *tu*.

Analizando a variável gênero, o autor verificou que homens e mulheres preferem a forma *você* (67% e 65%, respectivamente) à forma *tu* (32% e 34%, respectivamente).

Quanto à referência ao interlocutor, se direta ou indeterminada, o autor constatou que, em todos os tipos, a preferência é o pronome *você*. Já na referência direta, ocorreu o aumento da frequência de uso do *tu*.

Os dados apontam para 46% de ocorrência do pronome *tu* em falas não monitoradas, com peso relativo de 0,61, contra 17% de pronome *tu* em falas monitoradas, com peso relativo de 0,37. Referindo-se à variável discursiva expressividade do ato comunicativo (força de persuadir, emoção e capacidade apelativa aplicadas na enunciação), os resultados apontaram que, em enunciados com maior expressividade, a ocorrência da forma *tu* foi favorecida com o peso relativo de 0,65, ao passo que, em enunciados com menor expressividade, a forma *tu* foi desfavorecida com o peso de 0,39. Sobre ambos os fatores, Modesto (2006, p. 13) afirma que “a forma *tu* é usada na maioria das vezes em contextos de maior expressividade e menor monitoramento”. Quanto aos resultados em relação à função sintática, vejamos a tabela a seguir.

Tabela 11: Função sintática das formas *tu/ você* em Santos-SP

| FUNÇÃO SINTÁTICA | VOCÊ | TU |
|----------------------------|---------------|---------------|
| SUBJETIVA (SUJEITO) | 440/617 = 71% | 177/617 = 28% |
| OBJETIVA (OBJETO) | 36/91 = 39% | 55/91 = 60% |
| TOTAL | 476/708 = 67% | 232/708 = 33% |

Fonte: Modesto (2006, p. 85) (adaptada).

A partir da Tabela 11, verifica-se que a função objetiva da forma *tu* prevalece sobre as demais formas. Modesto (2006, p. 98-100) percebe que esse resultado é favorecido, principalmente, pela alta ocorrência da forma *te* na cidade de Santos.

O autor afirma que, diferentemente do que ocorre em outras regiões do Brasil, em Santos, “a forma *tu* não é selecionada em função exclusivamente do falante, mas em função da configuração do contexto de fala, em que estão interligados os fatores discursivo-pragmáticos” (MODESTO, 2006, p. 119).

1.4.4.4 Mota (2008)

Mota (2008) fez um estudo sobre a variação *tu* e *você* na cidade de São João da Ponte, no Norte de Minas Gerais. A amostra foi constituída por 24 informantes de ambos os sexos, com nível de escolaridade até o Ensino Fundamental, distribuídos por três faixas etárias (de 7 a 14 anos; de 15 a 25 anos; de 26 a 49 anos; e a partir de 50 anos).

Para a constituição do *corpus* da pesquisa, foram usadas, além de entrevistas sociolinguísticas, estratégias que criassem situações naturais de fala e que possibilassem enriquecer o *corpus*, aumentando o número de ocorrências relevantes. Foram feitas entrevistas sobre temas diversos, familiares e não-familiares ao falante. Alguns informantes narraram fatos de experiência pessoal e/ou fizeram retextualizações, narrativas recontadas, num ambiente de informalidade.

No cômputo da amostra, foram levantadas 509 ocorrências de formas de tratamento, das quais 49 foram do pronome *tu* (10%) e 460 foram do pronome *você* (90%). A pesquisadora faz um estudo, de forma isolada, das ocorrências da variante *tu*, embora o percentual de ocorrência seja somente 10% de realização desse pronome. Esse detalhamento justifica-se por a autora considerar inexistente no dialeto mineiro a variante *tu*. Nesse sentido, Mota (2008) buscou verificar quais fatores históricos, sociais e linguísticos poderiam explicar a presença do tratamento por *tu* no dialeto da comunidade em destaque.

Segunda a autora, das 49 ocorrências de *tu*, 19 realizaram-se como pronome sujeito, 27 como pronome objeto de verbo, uma foi como objeto de preposição e duas como pronome possessivo *tua*. Não foi encontrado no *corpus* o pronome possessivo *teu*. Dessa maneira, o número maior de ocorrências se situa nas posições de sujeito e de objeto de verbo.

Para a análise quantitativa, foram selecionados alguns grupos de fatores: três fatores linguísticos e seis extralinguísticos: função sintática da variante, tempo verbal, tipo de frase, grau de intimidade entre os falantes, tipo de discurso (atual ou relatado), área geográfica de atuação profissional do falante, sexo e faixa etária. Foram considerados significativos os grupos de fatores: função sintática da variante, tipo de discurso e área geográfica de atuação profissional do falante. Das variáveis sociais, foram selecionadas pelo Programa como estatisticamente relevantes os grupos de fatores: grau de intimidade entre os falantes, estatuto do interlocutor na interação e faixa etária.

A pesquisa deixou claro que a forma *tu* é uma marca do grupo de faixa etária de 15 a 25 anos e é também um fenômeno da zona rural. A hipótese inicial de que a variação *tu/você* no

Norte de Minas, especificamente em São João da Ponte, depende da alternância entre os estilos (formal ou informal) e do tipo de relação entre os interlocutores (maior intimidade/menor intimidade) foi confirmada.

1.4.4.5 Silva (2017)

Silva (2017) descreveu e analisou a variação dos pronomes de segunda pessoa do singular, *tu* e *você*, no falar ressaquinhense, identificando os fatores linguísticos e extralingüísticos que influenciam o uso de uma ou outra forma pronominal. A amostra foi composta de 24 entrevistas, estratificada em função da área geográfica (zona urbana e zona rural), da faixa etária (20 a 35 anos, 36 a 50 anos e 50 anos ou mais) e do sexo (feminino e masculino).

A cidade de Ressaquinha, *locus* da pesquisa, é muito importante para a história de Minas Gerais devido à extração de ouro que acontecera nesse local durante os primórdios da independência do Brasil.

Para a coleta dos dados, além do questionário, que tem como temática gêneros diferenciados, como por exemplo, receitas, instruções, aconselhamentos, exemplificações, também foi solicitado a alguns sujeitos que narrassem fatos recontados ou experiências pessoais. Em relação aos fatores escolhidos para análise, consideraram-se os seguintes grupos: contexto sintático da forma, contexto frasal, área geográfica, sexo, faixa etária e escolaridade.

É importante ressaltar que, além dos pronomes *tu* e *você*, os falantes ressaquinhenses também fazem uso das variantes *ocê* e *cê*. Assim, os dados referentes a essas formas foram contabilizados e, para a análise dos dados no Programa GoldVarb X, foram agrupadas todas as ocorrências dos pronomes *você*, *ocê* e *cê* como constituintes de uma única variante linguística: *você*.

Os resultados apontaram que o pronome *você* foi o mais favorável em Ressaquinha, com 61,5% de uso, contra 38,5% de uso do pronome *tu*. Das 579 ocorrências de pronomes de segunda pessoa, 356 foram de *você* e 223 do pronome *tu*. A incidência do pronome *tu* é um fato curioso em Minas Gerais, por se tratar de um Estado em que o uso da forma *você* é quase categórico.

No tocante à variável ‘sexo’, os resultados foram equilibrados, as mulheres usaram 61,2% de *você* e os homens 61,9% do mesmo pronome. Em relação ao pronome *tu*, os homens utilizaram 38,8% e as mulheres 38,1%.

Silva (2017) salienta que, diante dos dados apresentados, não se pode afirmar que o *tu* seja estigmatizado na fala da comunidade em estudo; o que se observa é que tal forma parece ser bem aceita entre os falantes.

1.4.5 Região Sul

Sobre dados da Região Sul, apresentamos os trabalhos de Ramos (1989), Loregian-Penkal (2004) e Franceschini (2010).

1.4.5.1 Ramos (1989)

Ramos (1989) utiliza na sua amostra falantes nascidos e educados em Florianópolis, considerando três níveis de escolaridade e três faixas etárias, no total de 18 células. Em cada uma dessas células, foram entrevistadas duas pessoas, o que resulta no total de 36 informantes. A autora apresenta uma abordagem das formas de tratamento-sujeito relativas à segunda pessoa do singular usadas no Sul do Brasil, especialmente na cidade de Florianópolis, onde se pode encontrar, segundo a autora, um sistema quaternário de tratamento.

Ramos (1989) afirma que as formas *tu* e *você* co-ocorrem no falar de Florianópolis sem que os falantes percebam. Comenta ainda que a presença da forma *você* em Florianópolis é mais evidente do que há alguns anos e, por isso, acha que essa forma de tratamento poderá predominar na ilha algum dia.

Ao final de cada entrevista, o entrevistador pedia a opinião do informante a respeito do uso dos pronomes *tu* e *você* e, de maneira consciente quanto ao uso das duas formas, respondiam que era por influência do interlocutor. Usavam sempre o *tu* em casa, com familiares, no trabalho e com amigos mais íntimos. E, como muitos turistas de outros Estados visitavam a ilha e estes utilizavam a forma *você*, nessas situações empregavam também esta forma. A seguir, transcrevemos um depoimento do informante GMYB3, que integra a amostra dessa pesquisa de Ramos (1989):

(01) Geralmente a gente usa mais o pronome de tratamento *tu*. Acho que por cultura mesmo porque todo mundo fala assim. A maior parte do pessoal. É costume aqui da Ilha. Geralmente cidade mais serena usa um jeito de falar mais correto, eu acho: *você*. O pessoal do litoral é mais aberto. *Você* é um negócio muito formal. É mais educado também, eu diria. Depende de quanto tempo conhece a pessoa. Com pessoal de fora fica meio rude, informal demais usar o *tu*. Em casa a gente tá falando todo dia, a gente conhece melhor e não

precisaria utilizar um jeito mais refinado. É mais informal, bem íntimo. Eu até diria que usar o *tu* as pessoas ficam mais soltas pra conversar do que o *você*. (RAMOS, 1989, p. 66)

A conclusão a que a autora chegou foi de que há uma diferença significativa entre os dois pronomes e as opiniões convergem para o seguinte resultado:

Quadro 4: Diferenças semântico-pragmáticas entre *tu* e *você*

| <i>TU</i> | <i>VOCÊ</i> |
|-----------------------------|----------------------------------|
| ÍNTIMO | <i>DISTANTE</i> |
| FAMILIAR | <i>COM ESTRANHOS</i> |
| EM AMBIENTE FAMILIAR | <i>INFLUÊNCIA DE FORA</i> |
| + DOS ILHÉUS | <i>+ BONITO</i> |
| RUDE | <i>EDUCADO</i> |
| INFORMAL | <i>FORMAL</i> |
| COLOQUIAL | <i>CORRETO</i> |
| DESRESPEITOSO | <i>RESPEITOSO</i> |

Fonte: Ramos (1989, p. 46).

A pesquisadora concluiu o estudo confirmando que na fala florianopolitana encontra-se um sistema quaternário de tratamento: a presença do *tu* como forma mais usada por estes falantes; a incorporação do pronome *você*, com características aproximadas do terceiro pronome, o pronome *senhor(a)*; e o pronome *zero*, contrariando, portanto, sua hipótese de um sistema ternário.

1.4.5.2 *Loregian-Penkal (2004)*

Loregian-Penkal (2004) amplia o *corpus* de 1996, analisando o resultado de entrevistas realizadas tanto nas capitais dos Estados sulistas (Curitiba, Florianópolis e Porto Alegre) quanto em três cidades do interior de Santa Catarina (Blumenau, Chapecó e Lages) e em três cidades do interior do Rio Grande do Sul (Flores da Cunha, Panambi e São Borja). O objetivo da pesquisa foi mapear a concorrência de *tu* e *você* e verificar se ocorre mesmo a substituição de *tu* por *você*.

Os dados das capitais mostram os seguintes resultados: Porto Alegre e Florianópolis apresentam semelhanças no uso dos pronomes: dos 24 informantes analisados em cada cidade, foram encontrados 14 falantes categóricos de *tu* em Porto Alegre em 13 e Florianópolis; alternam *tu/você* nove falantes em Porto Alegre e 10 em Florianópolis; foi encontrado apenas

1 falante de *você* em cada uma das duas capitais. Já em Curitiba, é categórico o uso de *você*. Em Porto Alegre e em Florianópolis há diferenças em relação ao comportamento de homens e mulheres. Há 10 mulheres em Porto Alegre e sete em Florianópolis com uso categórico de *tu*, enquanto entre os homens há uma distribuição mais equilibrada de *tu/você*, com probabilidade de que eles sejam os responsáveis pela entrada de *você* nessas comunidades.

Tabela 12: Distribuição de *tu/você* por sexo e faixa etária nas capitais do Sul do Brasil

| INFORMANTE | PORTO ALEGRE | | | FLORIANÓPOLIS | | | CURITIBA | | |
|-----------------|--------------|-------------|--------------|---------------|-------------|--------------|-----------|-------------|--------------|
| | <i>tu</i> | <i>você</i> | <i>T + V</i> | <i>tu</i> | <i>você</i> | <i>T + V</i> | <i>tu</i> | <i>você</i> | <i>T + V</i> |
| FA | 5 | - | 1 | 5 | - | 1 | - | 6 | - |
| FB | 5 | - | 1 | 2 | - | 4 | - | 6 | - |
| SUBTOTAL | 10 | - | 2 | 7 | - | 5 | - | 2 | - |
| MA | 2 | 1 | 3 | 4 | - | 2 | - | 6 | - |
| MB | 2 | - | 4 | 2 | - | 3 | - | 6 | - |
| SUBTOTAL | 4 | 1 | 7 | 6 | 1 | 5 | - | 12 | - |
| TOTAL | 14 | 1 | 9 | 13 | 1 | 10 | - | 24 | - |

Fonte: Loregian-Penkal (2004, p. 121) (adaptada).

A faixa etária de 25 a 49 anos é a que mais utiliza a forma *tu* em todas as cidades. Sugere-se que o uso do pronome *tu* talvez esteja associado a uma menor formalidade, ou a uma maior intimidade. E a autora diz não haver indícios de mudança em progresso em direção ao uso de só *você*.

A escolaridade não apresentou resultado homogêneo nas comunidades estudadas. Aparentemente, a educação formal não influencia a variação *tu/você* em algumas delas.

A partir das variáveis linguísticas analisadas, a autora descreveu os contextos em que o uso de *você* é favorecido: indeterminação do referente; discurso relatado de terceira pessoa; e discurso predominantemente narrativo. Esses são contextos em que há certo distanciamento do falante em relação àquilo a que se refere e ao interlocutor. O falante marca seu distanciamento usando *você* ao invés de *tu*. A autora aponta a possibilidade de ser por meio desses contextos que esteja se dando a entrada de *você* no repertório linguístico dos falantes das cidades analisadas.

As variáveis linguísticas selecionadas em Loregian-Penkal (2004) na comparação entre Porto Alegre, Florianópolis e o Ribeirão da Ilha juntas foram: gênero discursivo, determinação do discurso e explicitação do pronome.

1.4.5.3 Franceschini (2010)

No artigo “A influência dos fatores sociais nos pronomes *tu/você* em Concórdia – SC”, Franceschini (2010) apresenta resultados parciais de um estudo que investiga a variação no uso dessas formas pronominais em Concórdia – SC. A pesquisadora utilizou como *corpus* 12 entrevistas, tendo como variáveis independentes sexo, faixa etária (26 a 45 anos e 50 anos ou mais) e escolaridade (nível Fundamental I e II e nível Médio).

O objetivo da pesquisa foi descrever o uso dos pronomes *tu/você* na posição sujeito e verificar os grupos de fatores sociais que condicionam o uso de um ou outropronome.

O total de ocorrências do pronome *tu* na amostra de Franceschini (2010) foi de 204 (52%), enquanto o pronome *você* ocorreu 188 vezes (48%), totalizando 392 ocorrências. Faixa etária e escolaridade foram as variáveis sociais selecionadas como significativas pelo programa Varbrul. Reproduzimos, a seguir, os resultados obtidos pela autora.

Tabela 13: Atuação da faixa etária sobre a variante *tu*

| FAIXA ETÁRIA | APLIC./TOTAL | % | PESO RELATIVO |
|------------------------|----------------|-----------|---------------|
| 26-45 ANOS | 72/200 | 36 | 0,32 |
| 50 ANOS OU MAIS | 132/192 | 69 | 0,68 |
| TOTAL | 204/392 | 52 | - |

Fonte: Franceschini (2010, p. 779) (adaptada).

Perecebemos, portanto, que na pesquisa de Franceschini (2010) os falantes mais velhos, da faixa etária 2 (50 anos ou mais), favoreceram o uso do *tu* (peso relativo de 0,68) e que, ao contrário, os falantes mais jovens apresentam um uso reduzido deste pronome (peso relativo de 0,32), ou seja, os falantes mais jovens dessa amostra utilizam principalmente o pronome *você*, o que pode ser indício de uma mudança em tempo aparente.

1.5 SÍNTESE DO CAPÍTULO

São muitos os trabalhos desenvolvidos a respeito da variação dos pronomes *tu* e *você* no Brasil, tanto com dados da língua oral, quanto com dados da língua escrita, a saber: Sette (1980); Soares (1980); Ramos (1989); Salles (2001); Paredes Silva (2003); Modesto (2006); Franceschini (2010); dentre outros. Sabe-se que as formas de interlocução no português brasileiro vêm despertando interesse há muito tempo, o que tem levantado hipóteses e

explicações diferentes sobre o fenômeno em questão. Vejamos o Quadro 5 para sintetizar os resultados dessas pesquisas, por região, embora nem todos os autores tenha trabalhado com a oposição binária *tu/você* nem todos apresentem resultados numéricos.

Quadro 5: Síntese da frequência de uso das variantes *tu* e *você* nas regiões do Brasil

| REGIÕES | ESTUDO | <i>TU</i> | <i>VOCÊ</i> |
|---------------------|------------------------|-----------|-------------|
| NORTE | Martins (2010) | 42,9% | 22,9% |
| NORDESTE | Sette (1980) | 2,5% | 52,5% |
| | Soares (1980) | + | - |
| | Bezerra (1994) | 69% | 31% |
| | Herênia (2006) | 27% | 73% |
| | Almeida (2012) | 22% | 78% |
| | Nogueira (2013) | 4,85% | 95,15% |
| | Guimarães (2014) | 47,2% | 46,5% |
| CENTRO OESTE | Lucca (2005) | 72% | 28% |
| | Dias (2007) | 12,8% | 87,2% |
| | Andrade (2010) | 48% | 26% |
| | Scherre et al. (2015) | - | - |
| SUDESTE | Salles (2001) | 75% | 25% |
| | Paredes Silva (2003) | 65% | 25% |
| | Modesto (2006) | 32% | 67% |
| | Mota (2008) | 38,5% | 61,5% |
| | Silva (2017) | 10% | 89% |
| SUL | Ramos (1989) | 20% | 31% |
| | Loregian-Penkal (2004) | 65,6% | 34,4% |
| | Franceschini (2010) | 55% | 45% |

Fonte: Elaborado pela autora.

Segundo os estudos aqui revisados, notamos, com base no Quadro 5, que na Região Norte, apenas com um trabalho encontrado, a ocorrência do *tu* é bastante significativa na quantificação dos dados, aproximando-se dos resultados da Região Sul, enquanto no Nordeste a aparição do pronome *tu* é bem menor.

Na Região Sudeste, diferentemente do atestado por Salles (2001) e Paredes Silva (2003), na cidade de Santos, o pronome mais utilizado foi o *você*, como confirmam Modesto (2006), Mota (2008) e Silva (2017).

As pesquisas referentes à Região Centro-Oeste, feitas por Lucca (2005) e Andrade (2010) exibem favorecimento do uso do pronome *tu*, enquanto Dias (2007) apresentou uma predominância do pronome *você* na mesma região.

Na Região Sul, os trabalhos de Loregian-Penkal (2004) e Franceschini (2010) evidenciam preferência pelo pronome *tu*, diferentemente dos resultados encontrados por Ramos (1989), em que a forma *você* é a mais frequente.

Até aqui foi feita uma revisão das pesquisas sobre a variação do *tu* e do *você* nas regiões, com o propósito de mapear o comportamento desses pronomes. Algumas contribuições poderão ser elencadas para a nossa tese, a saber: Paredes (2003), em um segundo momento da sua busca pela variante *tu* na fala carioca, constituiu outro *corpus* de cunho oral, buscando captar a fala em situações mais naturais e menos controladas. Essas gravações pertencem ao Banco de Dados Interacionais – BDI e o resultado, devido à monitoração ocorrida no momento da gravação, foi insatisfatório; Soares (1980) salientou que a forma *você* é a preferida em situações mais formais e que o uso do *tu* é generalizado, mas sua concordância é variável, motivada por fatores como escolaridade, formalidade e atenção prestada à fala. Assim, na nossa pesquisa, percebemos que, na maioria das vezes nas entrevistas, a variante *você* foi utilizada em momentos formais, de relações de distanciamento.

Com base na teoria sociolinguística laboviana e nos pressupostos da geolinguística, analisamos nesta tese a variação no uso das formas de tratamento no PB, a fim de aprofundar a análise já existente sobre a variável em questão, desta vez com base em dados do Projeto ALiB, coletados neste século.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo abordamos os pontos teóricos que fornecem subsídios à pesquisa. Primeiramente, expomos sobre a variação linguística, uma vez que o nosso objeto de estudo é a variação pronominal das formas *tu* e *você*. São apresentados os aportes teóricos da Sociolinguística e da Dialetologia, com destaque para a interface entre essas áreas e para a divisão dialetal brasileira, bem como evidenciamos a respeito dos Atlas que foram concluídos e os que estão em andamento, no Brasil. Apontamos também algumas contribuições da Sociolinguística e da Dialetologia para o ensino da língua materna.

2.1 SOCIOLINGUÍSTICA

Para o desenvolvimento desta pesquisa, como já foi mencionado anteriormente, seguimos os pressupostos da Sociolinguística Quantitativa. A metodologia da Sociolinguística Variacionista abrange a quantificação dos dados em variação, em relação aos fatores que podem estar contribuindo para a seleção de uma ou outra forma variante. A seguir, apresentamos as etapas da pesquisa sociolinguística.

2.1.1 Etapas da pesquisa sociolinguística

A Sociolinguística surgiu no final da década de 1960 como uma resposta aos modelos teóricos que consideravam a língua um sistema homogêneo e invariável e à noção de língua que faz abstração da variação. Assim, a Sociolinguística Variacionista firmou seu lugar ao provar que a variação é inerente ao sistema linguístico.

Foi no início do século XX que começaram a germinar as sementes que viriam posteriormente – depois de cerca de meio século de domínio de correntes estruturalistas – a florescer e dar frutos no terreno de estudos da linguagem conhecido como Sociolinguística. Assim é que a partir da década de 1960, como herança de Meillet, volta a ganhar força a noção de língua como fato social dinâmico, cuja variação é explicada por forças externas e, como herança de Bakhtin, renova-se a perspectiva de que a língua é um fenômeno social cuja natureza é ideológica (COELHO, 2010).

Logo a Sociolinguística ocupa-se de questões como variação e mudança linguística, bilinguismo, contato linguístico, línguas minoritárias, política e planejamento linguístico, entre outras.

Na década de 1960, Weinreich, Labov e Herzog (2006 [1968]) manifestam interesse em inserir o componente social nos estudos linguísticos. A partir de então, William Labov, principalmente, começa a desenvolver uma série de estudos sobre fala, almejando explicar e sistematizar a variação nas línguas. O estudo da língua, sob este ponto de vista, é feito a partir da língua em uso, de forma que as escolhas que o falante faz dependem não somente de fatores internos à estrutura linguística como também de fatores relacionados à situação de uso.

Conceber a língua desta forma contrariava correntes teóricas, como o estruturalismo e o gerativismo, que examinavam a língua por ela mesma, como uma estrutura abstrata, desvinculada de fatores históricos e sociais, e que idealizavam o sistema linguístico de maneira uniforme, homogêneo e estático. Nessa nova concepção, a heterogeneidade não ocorre de maneira aleatória, mas, sim, fundamentada por um conjunto de regras previsíveis (PEREIRA, 2016).

As pesquisas anteriores as de Labov (1963) concebiam as unidades linguísticas, tais como fonemas, morfemas e todas as demais, “invariantes, discretas e qualitativas”. Depois dos estudos variacionistas, essa concepção começa a mudar e essas unidades linguísticas passam a ser vistas como: 1) **estruturas variantes**, porque se realizam de maneira diferente em situações diversificadas; 2) **contínuas**, pois o valor social de determinadas variantes está ligado à diferenciação em relação à variante padrão; 3) **quantitativa**, uma vez que a significação da variável é definida não só pela presença ou ausência de suas variantes mas por suas frequências relativas, razão pela qual essa teoria passou a chamar-se Sociolinguística Quantitativa (PEREIRA, 2016).

Assim, podemos definir a Sociolinguística, em sentido mais amplo, como uma subárea da Linguística responsável por estudar a língua aliada a fatores sociais, isto é, examinando a correlação entre as categorias internas ou externas ao sistema linguístico. O foco de estudo da sociolinguística é a variação, que, como afirma Mollica (2003, p. 10), é entendida “como um princípio geral e universal, passível de ser descrita e analisada cientificamente”.

O trabalho de William Labov em Martha's Vineyard foi um marco fundamental para a caracterização da Sociolinguística enquanto ciência dotada de método, cujo objetivo foi analisar um fenômeno de mudança linguística – fônica – em processo na fala de seus habitantes. A partir

de então, Labov passou a desenvolver uma série de pesquisas empíricas baseadas na teoria que ficou conhecida como Sociolinguística Quantitativa ou Variacionista.

É importante mencionar dois autores que influenciaram decisivamente Labov: Uriel Weinreich e Marvin Herzog. O primeiro foi seu professor na Columbia University e o orientou na dissertação de mestrado sobre a centralização dos ditongos /ay/ e /aw/ em Martha's Vineyard e depois na tese de doutorado sobre a estratificação do /r/ em lojas de departamentos na cidade de Nova Iorque. O segundo foi seu colega na mesma Universidade. Labov e Herzog receberam de Weinreich, em 1966, a tarefa de escreverem o ensaio *Fundamentos empíricos para uma teoria da mudança linguística*, que deveriam apresentar numa conferência na Universidade do Texas. Nesse ensaio, foram incorporados resultados das pesquisas em Martha's Vineyard e em Nova Iorque, de Labov; da dialetologia do iídiche no Norte da Polônia, de Herzog; além da visão global de Weinreich, que foi o criador do *Atlas Linguístico e Cultural dos Judeus Asquenazes* (LABOV, 2008 [1972], p. 35).

A teoria laboviana facilitou compreender que a variação linguística é passível de sistematização e que não é caótica, antes apresenta regularidades que não são devidas ao acaso. Mostrou que existe uma relação intrínseca e inseparável entre a língua e a sociedade e que a variação pode ser explicada por fatores internos e externos ao sistema. A língua passa a ser vista como um instrumento social de comunicação, sendo os atos linguísticos eminentemente sociais e pragmáticos, instrumentos para se estabelecer e manter o relacionamento entre indivíduos em sociedade.

Segundo Tarallo (2006), todos os que concebem a língua como um veículo de comunicação, informação e de expressão entre as pessoas devem também ser chamados de sociolinguistas. Entretanto vale destacar que a teoria laboviana sobressaiu-se por colocar em evidência a possibilidade de sistematizar a variação existente na própria língua falada, o que, por conseguinte, coloca a ‘fala’ em evidência.

A Sociolinguística Variacionista caracteriza-se em ocupar-se com números e dar tratamento estatístico aos dados coletados, opera com duas ou mais formas linguísticas que configuram o mesmo valor de verdade, tratadas como variantes de uma mesma variável. Através da Sociolinguística, podem-se investigar fenômenos estigmatizados ou não e, consequentemente, prover mecanismos visando à elaboração didático-metodológica para os estudos da oralidade em diferentes locais (ZILLI, 2009). Graças a essa teoria, podemos fazer o registro de diversos tipos de falas em diferentes localidades na mesma região. Ao passo que,

apontaremos a necessidade de se delimitar a comunidade de fala na qual a variação será pesquisada para que, assim, os dados sejam os mais fieis possíveis.

A metodologia da sociolinguística contribui também para a realização da seleção de informantes, chama atenção para o paradoxo do observador, orienta para um método de entrevista mais dinâmico, o que vai favorecer uma melhor coleta de dados e minimizar o efeito negativo causado pela presença do pesquisador.

A Sociolinguística laboviana dispõe de uma metodologia de pesquisa de campo bastante criteriosa para conduzir os estudos sobre variação e mudança linguística, a saber:

- Seleção dos informantes;
- Metodologia de coleta de dados;
- Envelope de variação;
- Levantamento de questões e hipóteses;
- Codificação de dados e análise estatística.

Sobre a metodologia laboviana para a coleta de dados, Coelho (2010) ressalva que:

Nas entrevistas sociolinguísticas o entrevistador deve tentar (i) neutralizar a força inibidora de sua presença (já que ele é uma pessoa estranha à comunidade) e do gravador, mostrando-se interessado, de fato, nas histórias que os informantes vão contar; (ii) realizar o mínimo de interferências no momento em que o informante estiver discorrendo sobre os assuntos que despertem o seu interesse. (COELHO, 2010, p. 119)

É essa, em suma, a postura investigativa que se adota no trabalho com a Sociolinguística. Com ela, identifica-se uma variável no uso corrente da língua de uma comunidade; identificam-se, a seguir, as variantes dessa variável; a partir das hipóteses que elaboramos quanto aos fatores que possam estar em jogo no favorecimento ou desfavorecimento das variantes, prossegue-se à coleta e, posteriormente, à análise de dados, para a confirmação ou refutação de nossas hipóteses iniciais. Coelho (2010, p. 119) acrescenta: “Mas essa é somente a síntese, uma vez que uma pesquisa sociolinguística deve ir muito além disso se quiser efetivamente esclarecer um pouco mais sobre a complexa relação que há entre língua e sociedade”.

Foi com base na orientação da Sociolinguística Quantitativa que verificamos os efeitos dos fatores linguísticos e extralingüísticos que favorecem ou inibem o uso dos pronomes *tu* e *você* nos cinco Estados nordestinos, o que está evidenciado no Capítulo 4, de análise dos dados.

2.1.2 Variação linguística

É por meio da linguagem que os indivíduos se comunicam e retratam o conhecimento e o entendimento de si próprios e do mundo que os cerca. A linguagem está intrinsecamente ligada à sociedade. Essa interação constitui o ser humano, uma vez que é através do exercício da linguagem que o homem constrói a sua identidade e suas relações com outros homens. De acordo com Benveniste (1976, p. 27), “é dentro da, e pela língua, que indivíduo e sociedade se determinam mutualmente”. Ao acatarmos que a linguagem é um fenômeno social, logo compreendemos que a mesma não é estática nem homogênea, manifestando-se de inúmeras formas.

A variação linguística é um fenômeno que acontece com a língua e pode ser atestada nas variações diatópicas e diastráticas. Em um mesmo País, com um único idioma oficial, a língua pode sofrer diversas alterações feitas por seus falantes. Como não é um sistema fechado e imutável, a língua portuguesa ganha diferentes formas. O português que é falado no Nordeste do Brasil pode ser diferente do português falado no Sul do País. Sabe-se que um idioma nos une, mas as variações podem ser consideráveis e justificadas de acordo com a comunidade na qual se manifestam.

Essas variações linguísticas podem estar relacionadas a fatores diversos, tais como: classe social, sexo, origem geográfica, profissão, idade, grau de escolaridade, entre outros. É evidente que há diversas formas de manifestação da língua, variando de acordo com a situação de uso e seus interlocutores.

Callou e Leite (2002, p. 9) dizem que a explicação para a variação linguística deve-se ao próprio processo de colonização do País: “dialetação horizontal por influxo indígena e diferenciação vertical entre a fala do luso e a fala do nascido e criado na terra”. O pluralismo étnico e cultural que se processaram no cenário nacional são responsáveis pelas diferenças existentes entre os diferentes falares brasileiros.

Alguns teóricos verificaram que as variações acontecem porque o princípio fundamental da língua é a comunicação, logo é compreensível que seus falantes façam readaptações de acordo com suas necessidades comunicativas. Os diferentes falares devem ser considerados como variações, e não como erros. Quando tratamos as variações como erro, incorremos no preconceito linguístico, que associa, equivocadamente, a língua a *status*. Não é justo que façamos da língua um mecanismo de segregação cultural, atestando a estratificação social da

mesma, ao julgarmos determinada manifestação linguística superior a outra, sobretudo superior às manifestações linguísticas de classes sociais ou regiões menos favorecidas.

Podemos perceber que as variedades linguísticas mais propícias a preconceito linguístico são, normalmente, as que possuem características associadas a grupos de pessoas com menos prestígio na escala social ou a grupos de pessoas da área rural ou do interior do País. Este fato decorre do sentimento de superioridade – muito claro na mídia – dos grupos vistos como mais privilegiados econômica e socialmente. Assim, numa comunidade de fala, temos variantes que estão sempre competindo: as variantes padrão e não padrão, aquelas que são conservadoras contra as que são inovadoras e, finalmente, as variantes que recebem algum tipo de estigma em oposição àquelas de prestígio. Na maioria das vezes, a variante padrão é considerada conservadora e possui maior importância sociolinguística dentro da comunidade. De outro modo, uma variante inovadora tende a ser não-padrão e, por isso, estigmatizada pelos falantes da comunidade.

A escola não pode ignorar as diferenças sociolinguísticas. Os professores e, por meio deles, os alunos têm que estar bem conscientes de que existem duas ou mais maneiras de dizer a mesma coisa. E mais, que essas formas alternativas servem a propósitos comunicativos distintos e são recebidas de maneira diferenciada pela sociedade. Algumas conferem prestígio ao falante, aumentando-lhe a credibilidade e o poder de persuasão; outras contribuem para formar-lhe uma imagem negativa diminuindo-lhe as oportunidades. (BORTONI-RICARDO, 2005, p. 15)

A escola tem como papel fundamental ensinar a língua da cultura dominante, mas desconstruindo a ideia de que tudo o que se afasta desse código é impróprio, defeituoso e deve ser eliminado. Infelizmente, na escola brasileira, as diferenças linguísticas socialmente condicionadas não são seriamente levadas em conta, uma vez que o prestígio linguístico está associado ao português padrão. Enfim, a escola não pode ignorar as diferenças dialetais e sociais, mas deve, portanto, valorizar e respeitar as peculiaridades linguístico-culturais dos falantes.

Ainda sobre a estigmatização social de formas linguísticas, Labov (2008 [1972] p. 360) assim se manifesta:

Um estereótipo social é um fato social, parte do conhecimento geral dos membros adultos da sociedade. Isso é verdade mesmo quando o estereótipo não corresponde a nenhum conjunto de fatos objetivos. Os membros da comunidade de fala se referem aos estereótipos e falam sobre eles; podem ter

um rótulo geral e uma frase característica que serve igualmente bem para identificá-lo.

Vivemos em uma sociedade complexa, na qual estão inseridos diferentes grupos sociais e alguns desses grupos tiveram acesso à educação formal, enquanto outros não tiveram muito contato com a norma culta da língua⁴. É notório que a língua varia de acordo com suas situações de uso, pois um mesmo grupo social pode se comunicar de maneira diferente, de acordo com a necessidade de adequação linguística. Prova disso é que o indivíduo não vai se comportar em uma entrevista de emprego da mesma maneira como conversa com amigos em uma situação informal. Dessa maneira, podemos afirmar que a variação linguística se manifesta desde o nível mais elevado e coletivo – quando comparamos, por exemplo, o português falado em dois países diferentes (Brasil e Portugal) – até o nível mais baixo e individual, quando observamos o modo de falar de uma única pessoa variando a depender da situação social (BAGNO, 2002). Entre esses dois níveis extremos, a variação aparece em diversos outros níveis: sexo, regiões dentro dos Estados, classes sociais, grupos de idade, níveis de renda, graus de escolarização, profissões, modalidades escritas e modalidades faladas, grau de formalismo.

A percepção de que a língua é variável provém desde a Antiguidade, quando os primeiros pesquisadores da língua grega tentaram estruturá-la para o ensino e para a crítica literária. Estes, imediatamente, fizeram um diagnóstico negativo da variação, acreditando que era um entrave para a unificação territorial e para a difusão da língua. Foi nesse período (século III a.C.) que surgiu a gramática no mundo ocidental, a partir de uma disciplina do helenismo⁵, empenhada em criar um modelo de língua que alcançasse uma posição superior à variação e servisse de instrumento de controle social.

2.1.2.1 Tipos de variação

Mencionamos, a seguir, os tipos de variação que podem resultar da influência de fatores extralingüísticos.

⁴ A expressão ‘norma culta’ está associada à dimensão da cultura escrita. É a norma linguística praticada em determinadas situações que envolvem certo grau de formalidade, por aqueles grupos sociais mais diretamente relacionados com a cultura escrita, legitimada historicamente por grupos que controlam o poder social.

⁵ O período helenístico foi delimitado historicamente como a fase da história da Grécia e parte do Oriente Médio.

2.1.2.1.1 Variação diatópica

É a variação geográfica, também conhecida como regional ou dialetal, a responsável por podermos identificar, às vezes com bastante precisão, a origem de uma pessoa através do modo como ela fala. É possível saber quando um falante é gaúcho, mineiro ou de um dos Estados do Nordeste, por exemplo.

São variações que acontecem de acordo com o local onde vivem os falantes, sofrendo sua influência. Esse tipo de variação ocorre porque diferentes regiões têm culturas diversas, com hábitos, modos e tradições distintos, estabelecendo assim outras estruturas linguísticas.

Podem ser observadas por diferentes palavras para os mesmos conceitos, diferentes sotaques, dialetos e falares, e até mesmo com reduções de palavras ou perdas de fonemas. “Em geral, itens lexicais particulares, certos padrões entoacionais e, principalmente, certos traços fonológicos respondem pelo fato de que falantes de localidades diferentes apresentem dialetos diferentes de uma mesma língua” (COELHO, 2010, p. 76).

Como exemplo de variações diatópicas, podemos mencionar a *abóbora* que conhecemos no Sudeste, chamada de *jerimum* no Nordeste brasileiro pela influência indígena, assim como a *mandioca*, que pode ser encontrada como *aipim* ou *macaxeira*, dependendo da região do País. Outro exemplo de variação regional é a pronúncia das vogais /e/ e /o/ pré-tônicas, como nas palavras *peteca* e *moderno*, que no dialeto nordestino são pronunciadas abertas (*p[ɛ]teca* – *m[ɔ]derno*) e no dialeto do Sudeste e do Sul do Brasil são pronunciadas fechadas (*p[e]teca* – *m[o]derno*).

Assim, a variação regional pode ser estudada colocando-se em oposição diferentes tipos de unidades espaciais: podemos dizer que existe variação regional entre Brasil e Portugal (dois países), entre o Nordeste e o Sul do Brasil (duas regiões de um mesmo País), entre Paraná e Santa Catarina (dois Estados de uma mesma região), entre Feira de Santana e Irecê (duas cidades de um mesmo Estado), bem como o estudo diferencial entre zona rural e zona urbana.

2.1.2.1.2 Variação diastrática ou social

É aquela variação que ocorre em virtude da convivência entre os grupos sociais. Como exemplo, podemos citar a linguagem dos advogados, dos surfistas, da classe médica, entre outras. Assim, da mesma forma que a fala pode carregar marcas de diferentes regiões, também pode refletir diferentes características sociais dos falantes. Os principais fatores sociais que

condicionam a variação linguística são a faixa etária, o grau de escolaridade, o nível socioeconômico, o sexo/gênero e mesmo a profissão dos falantes, conforme exemplificamos a seguir:

- Faixa etária (variação diageracional). A relação entre variação linguística e idade do falante tem acarretado muitas reflexões entre os pesquisadores, uma vez que entra em questão a mudança linguística. Algumas expressões e palavras, como, por exemplo, “dia da mulher” (menstruação), “barrer” (varrer), “prosa” (conversa), são utilizadas por pessoas com idade mais avançada. Por outro lado, vocábulos como: *irado* e *sinistro* são utilizados por jovens como forma de elogio.
- Grau de escolaridade. Supõe-se que, em geral, falantes com nível de escolarização elevado dificilmente produzirão modelos como *nós vai* ou *a gente vamos*, que são típicos de indivíduos com pouca ou nenhuma escolarização.
- Nível socioeconômico (variação diastrática). Os filhos de uma elite, que frequentaram colégios mais conceituados e cuja família utiliza um vocabulário mais culto no ambiente familiar, tendem a utilizar uma linguagem mais organizada. Já os filhos de pais analfabetos, cuja preocupação maior, muitas vezes, maior é alimentá-los, não terão os mesmos recursos. Enfim, “a questão de letrar-se é muito mais complexa. A diferença linguística entre classes sociais não se restringe às classes alta, média e baixa” (PAIM, 2019, p.41). Os médicos, por exemplo, utilizam uma linguagem particular, assim como os adolescentes criam uma linguagem para se diferenciarem de outros grupos.
- Sexo/gênero (variação diagenérica). A afirmação de que o sexo feminino é propulsor de variantes conservadoras já foi comprovada em muitas pesquisas. Numa pesquisa realizada por Scherre (1996) com uma variável morfossintática, ficou evidente que o índice maior de utilização da concordância entre os elementos do sintagma nominal é atestado em falantes do sexo feminino. Por outro lado, podemos encontrar também falantes do sexo feminino que apresentam comportamentos linguísticos diferentes, a depender da localidade, se zona rural ou zona urbana.
- Profissão. Há termos técnicos que são utilizados no exercício de alguns profissionais, como policiais, professores, médicos, advogados, farmacêuticos etc. Por exemplo, para o jargão policial, podemos encontrar com frequência os seguintes itens lexicais: “elemento”, “viatura”, “meliante”, “auto”, “ocorrência”.

2.1.2.1.3 Variação diafásica ou estilística

É a variação que depende do contexto comunicativo, ou seja, a ocasião é que determina a maneira como nos dirigimos ao nosso interlocutor, se deve ser formal ou informal. A linguagem formal ou culta, por exemplo, tenta levar em consideração as normas prescritivas da língua. Ela pode ser usada quando não há familiaridade entre os ouvintes da comunicação ou em situações que necessitam de maior elegância, acarretando, portanto, uma linguagem mais monitorada. Por outro lado, a linguagem informal é usada quando há familiaridade entre os ouvintes da comunicação ou em situações descontraídas, quando o falante está mais preocupado com *o que* dizer e não com o *como* dizer, ou seja, em situações de menor monitoramento.

Um mesmo falante pode usar diferentes formas linguísticas, dependendo da situação em que se encontra. Basta pensarmos que a maneira como falamos em casa, com nossa família, não é a mesma como falamos em nosso emprego, com o chefe. O que está em jogo aí são os diferentes ‘papéis sociais’ que as pessoas desempenham nas interações que se estabelecem em diferentes ‘domínios sociais’: na escola, na igreja, no trabalho, em casa, com os amigos etc. Os papéis sociais que desempenhamos vão se alterando em conformidade com as situações comunicativas (entre professor e aluno, patrão e empregado, pais e filhos, irmãos etc.) (COELHO, 2010, p. 82).

Como exemplo de trabalho pioneiro na consideração dos fatores estilísticos, aqui no Brasil, podemos citar a pesquisa de Naro e Lemle (1977) sobre a variação na concordância verbal. Os autores chegaram ao resultado de que, nos discursos informais, nos contextos familiares e em situações menos formais, os falantes eram menos propensos a realizar a marca de concordância verbal no plural do que nos discursos formais, não-familiares e em contextos mais formais, em que era favorecida a marcação explícita da concordância.

Labov (2008 [1972], p. 91) chama a atenção para o grande desafio que os linguistas encontram em utilizar técnicas para resolver os problemas de variação estilística, pois as mesmas são consideradas insuficientes e inadequadas. Na investigação feita pelo autor nas lojas de departamento de Nova York ele notou na comunidade de fala, que havia uma variação regular em estilos e contextos diferentes. Desse modo, na busca de dar conta da variação estilística – não menos importante que a social, por exemplo - as pesquisas labovianas foram pioneiras em observar a influência do condicionamento estilístico sobre o comportamento linguístico.

O princípio básico da *variação estilística* é que o falante não utiliza a língua da mesma forma em todas as ocasiões, o que acarreta na escolha de diversas possibilidades de expressão. Labov (2008 [1972]) considera que há um *continuum* que vai da máxima informalidade até a máxima formalidade.

Essa caracterização foi de grande valia para o estabelecimento dos fatores, arrolados nesta pesquisa, que podem influenciar no momento em que o interlocutor precisa fazer a escolha de um tratamento adequada para determinado momento, dentro de um determinado contexto discursivo, levando em conta, também, os interlocutores e suas características (escolaridade, sexo, classe social etc.).

Na sequência, discorreremos sobre as noções de estilo e indivíduo segundo Labov (2008 [1972]) e, em seguida, apresentaremos a proposta de Bell (1984) que se caracteriza pela importância dada aos falantes e destinatários e suas relações uns com os outros nos eventos de fala.

Nesta pesquisa, sobretudo, aceitamos a argumentação de Labov (2008 [1972]), bem como a tese de Bell (1984) a respeito da influência do interlocutor sobre a escolha estilística do falante, pois acreditamos que as duas teses são complementares: um falante diante de um interlocutor desconhecido utiliza uma fala monitorada, um estilo mais cuidado.

2.1.2.1.3.1 Estilo em Labov

Os indivíduos utilizam estilos linguísticos variados para se posicionarem uns em relação aos outros e desfrutam de estratégias discursivas para atender a esse fim.

Argumentos sobre a alternância de estilo já haviam sido levantados por Labov (2008 [1972]) quando investigou o uso do inglês na cidade de Nova Iorque. Pesquisando sobre a variação na fala e na comunidade, notou que não havendo nenhum modo claro de caracterizar os comportamentos adotados pelo falante na alternância de um estilo a outro, ou mesmo o falante recorrendo a um único estilo várias vezes, ele designou essas alternâncias como *variantes estilísticas*.

Labov (2008 [1972]; 2001) com base na análise quantitativa de dados em entrevistas sociolinguísticas, analisa a correlação entre o comportamento de determinadas variáveis e as partes da entrevista, classificadas em uma escala de níveis de formalidade. O pesquisador afirma que a organização de estilos contextuais não se trata de uma descrição geral para a observação de como os estilos são produzidos e organizados no discurso de cada falante, mas

sim uma organização sistemática da variação intrafalante observada em entrevistas sociolinguística.

Desde 1966 Labov empenhou-se em observar as trocas estilísticas na entrevista, as quais foram caracterizadas como “estilos contextuais”, subdivididos em um eixo de atenção prestada à fala na direção menos formal e mais formal. Nessa perspectiva de variação estilística, Labov (2008 [1972], p. 102) define de fala monitorada (*careful speech*) quando o falante presta mais atenção à sua própria fala e, de fala casual (*casual speech*), quando o falante presta menos atenção à sua própria fala, entendida aqui como a fala cotidiana usada nas situações informais. O maior ou menor “grau de espontaneidade ou entusiasmo” do indivíduo na produção do discurso pode variar em relação à fala monitorada, contextos mais formais, onde a atenção dirigida à fala é maior. O estilo de fala espontânea é correlato da fala casual e pode ocorrer tanto em contextos informais como formais, mas, que decorre “não em resposta à situação formal, mas apesar dela” (LABOV, 2008 [1972], p. 111).

Reiteramos que a fala cuidada apresenta um estilo mais simples de ser identificado numa entrevista, até porque Labov (2008 [1972]) sugere alguns recursos: estilo de leitura, lista de palavras e pares mínimos. Essas estratégias levam o falante a preocupar-se com a forma de manipular a língua. Por outro lado, o autor estabelece situações contextuais que permitem ao entrevistador controlar quando os traços mais formais da entrevista são abandonados pelo falante, a saber: ‘fala fora do formato entrevista’, ‘fala com uma terceira pessoa’, ‘fala que não corresponde às questões estabelecidas’, ‘fala sobre brincadeiras de crianças’ e ‘fala sobre o perigo de morte’ (2008 [1972], p. 111-121). Lembremos que a situação de entrevista define, por si só, um contexto mais adequado para falas monitoradas, acarretando portanto, em algumas implicações para o pesquisador que necessita dessa metodologia para coletar seus dados.

Labov (2008 [1972]) defende que o estilo de cada falante varia em um *continuum* em que de um lado temos o prestígio, como resultado da fala mais formal, variedade localizada na extremidade mais alta da escala da hierarquia socioeconômica e, do outro, temos a fala mais coloquial, a estigmatizada, situada na extremidade oposta do contínuo, na parte mais baixa. Percebemos, então, que embora a noção de prestígio desempenhe um papel importante na obra laboviana, é a atenção prestada à fala que o autor põe no centro da sua teoria estilística. Por fim, a atividade estilística do falante e as estratégias operadas com respeito a esta mesma hierarquia, estariam diretamente vinculadas ao lugar que ele ocupa na hierarquia socioeconômica.

Tendo como critério a metodologia utilizada nas suas entrevistas, Labov (2001) propõe um outro modelo para análise da fala nas entrevistas, conhecido como a *árvore da decisão*. Esse

modelo arbóreo é constituído de oito critérios organizados em quatro contextos classificados como “fala casual” e quatro como “fala cuidada”, arrumados em ordem decrescente de objetividade.

Nessa direção, as entrevistas podem ser encaixadas no estilo mais casual que são as passagens que dizem respeito às narrativas de cunho pessoal, a participação de outras pessoas, as passagens que tratam da infância e respostas tangenciais (que desviam do tópico introduzido). Em contrapartida, para o estilo cuidado, são considerados: as respostas às perguntas feitas pelo entrevistador, questões que tratam de avaliação da língua, opiniões generalizadas (*soapbox*) e os resíduos que não se encaixam em nenhum desses casos (LABOV, 2001, p. 94).

Como qualquer outro pesquisador, a proposta de Labov enfrentou algumas críticas, destacamos:

- a- dificuldade de se quantificar a atenção prestada à fala;
- b- dificuldade operacional de separar a fala casual da fala cuidada a partir das situações contextuais;
- c- caráter unidimensional da proposta, o que implica um *continuum* formalidade vs. informalidade;
- d- o falante é visto como passivo, alternando sua fala em resposta às mudanças na situação externa, mais do que creditando-lhes a capacidade de agenciar seu uso dos recursos estilísticos.

Outra crítica direcionada a abordagem laboviana da *árvore da decisão* está relacionada à mistura de critérios de natureza diferenciada, como: a audiência (grupo), o tópico (infância, língua), o controle conversacional (resposta vs. tangente) e o gênero (narrativa vs. *soapbox*). Soapex é um fundamento caracterizado como a expressão extensa de opiniões generalizadas que, necessariamente, não precisam ser direcionada a um determinado receptor (interlocutor), mas pode ser organizada como se fosse dirigida a um receptor indeterminado, ou seja, uma audiência mais ampla.

2.1.2.1.3.2 Estilo em Allan Bell

Partindo da premissa de que, normalmente, a mudança de estilo acontece quando o falante tenta se acomodar ao estilo da pessoa com a qual está falando, reiteramos aqui o comportamento monitorado do falante quando o interlocutor é uma pessoa não íntima.

Allan Bell (1984), em sua pesquisa sobre variação estilística, buscou responder os questionamentos sobre esse tipo de variação. O autor apresenta o modelo da *Audience Design*, estudo baseado na Teoria da Acomodação. A abordagem voltada para a “audiência” assegura que os falantes mudam seu estilo de fala em resposta aos seus diferentes interlocutores (ouvintes), isto é, o falante molda o seu estilo, principalmente, em relação ao seu destinatário.

Bell aperfeiçoou seus estudos, em 1977, quando pesquisou para o doutorado a variação no estilo da fala de locutores de rádio da Nova Zelândia e suas influências geradas nos ouvintes, ou seja, o programa era feito pelos mesmos locutores, porém o público ouvinte era diferente. Enfim, ele constatou a ocorrência de alternância nas variações de estilo da fala que resultaram nas características fonéticas e morfológicas, de acordo com a audiência radiofônica.

Seguindo a proposta belliana, “o estilo é essencialmente uma resposta do falante a sua audiência”, a mudança de estilo decorre das diferenças linguísticas dos indivíduos que fazem parte de grupos sociais distintos; seu argumento revela que a variação estilística é indubitavelmente reflexo do social. Nesse sentido, salienta que nos estudos sociolinguísticos encontramos dois grupos de fatores que estão associados à variação linguística. Os fatores linguísticos: fonológicos, morfológicos e sintáticos que possibilitam ou inibem a aplicação de uma regra variável e os fatores extralingüísticos, subdivididos em duas esferas que, segundo ele, desde Labov com seu estudo na cidade de Nova York, foram intitulados como “eixos: *social* e *estilístico* da variação linguística, onde o social indica as diferenças entre a fala de diferentes falantes e o estilístico aponta as diferenças dentro da fala de um único falante” (BELL, 1984, p. 145-146).

De acordo com o pesquisador, a variação estilística e a social operam simultaneamente e pode ser difícil distinguir entre a fala casual de um indivíduo escolarizado e a fala cuidada de um falante menos escolarizado. “Se a variação estilística deriva da variação social, a variação social vem primeiro”. Assim, algumas variáveis linguísticas exibem a variação social bem como a estilística, porém, algumas variáveis têm somente variação social, mas, nenhuma variação somente estilística, porque a dimensão estilística pressupõe a social.

Ainda a respeito da posição de Bell (1984, p. 186), a mesma se ampara em razão do falante relacionar a mudança de tópico ao tipo de audiência, pois é possível redefinir seu estilo a partir de um ouvinte pré-existente na interação, os chamados “referees”: participantes que, mesmo ausentes, podem influenciar a mudança estilística pelo fato de estarem presentes na mente do falante.

Árbitros são terceiras pessoas não presentes fisicamente a uma interação, mas possuem tanta saliência para um falante que influenciam a fala até mesmo quando ausentes. (...) O efeito da configuração do árbitro é fazer um falante mudar de estilo como se de fato estivesse falando com o árbitro em lugar do destinatário. (BELL, 1984, p. 186)

O indivíduo ao reproduzir a sua fala leva em conta o ouvinte, ou seja, a sua “audiência”, que é constituída não só do destinatário, a segunda pessoa, mas de outras terceiras pessoas. O destinatário, a segunda pessoa, é conhecido e validado, os demais não precisam ser, necessariamente validados. Os interlocutores que são conhecidos e validados, são denominados de “auditor”. Aqueles que o falante sabe que estão lá, mas que não são participantes ratificados, são os “overhearers”. Outros cuja presença é desconhecida são os “eavesdroppers”.

Assim, esses papéis da audiência (*auditor, overhearer, eavesdropper e addressee*) estão ordenados de forma que sejam validados e conhecidos, seguindo uma hierarquia, cada um com seu papel, atribuído pelo falante, o que acarreta para a variação sociolinguística em interpretações qualitativa e quantitativa, segundo Bell (1984).

2.1.2.1.3.3 Um ponto de congruência

Compreendemos que a proposta de Bell (1984) não é incompatível com a de Labov (2008 [1972]), uma vez que na tentativa de configurar a sua fala, o indivíduo tem consciência de que pode fazer alterações de acordo com o ambiente e com as características de seu ouvinte, podendo prestar mais ou menos atenção a sua fala. Compactuamos, neste ponto, com a posição de Bortoni-Ricardo (2002, p. 335), segundo a qual as visões de Labov (2008 [1972]) e Bell (1984) são complementares porque um falante, diante de interlocutor desconhecido, de maior poder na hierarquia social ou a quem ele precisa ou deseja impressionar, sente-se na obrigação de usar um estilo mais cuidado, monitorado.

Outra contribuição de Bortoni-Ricardo (2002, p. 336) é apontar quatro fatores que estabelecem o nível de monitoramento do falante, interferindo, portanto, na variação estilística, são eles:

- i) a acomodação do falante ao seu interlocutor, fator que se refere ao tipo de relação existente entre os interlocutores;
- ii) o apoio contextual na produção dos enunciados;
- iii) a complexidade cognitiva envolvida na produção temática (que se refere ao quanto o falante entende sobre o tópico discursivo); e

- iv) a familiaridade do falante com a tarefa comunicativa que está sendo desenvolvida.

Mesmo com tamanha atenção direcionada à fala e projeção que o falante faz do seu receptor, devemos considerar também estes fatores no processo de escolha estilística que o locutor faz mediante determinada interação.

Na verificação dos dados da tese, parece-nos adequado lidar com o conceito de ‘grau de atenção prestada à fala’ que pode variar, no *corpus* em análise. Por fim, nas entrevistas selecionadas, teremos condições de verificar esses elementos acrescentados por Bortoni-Ricardo (op. cit.) às análises de Labov (2008 [1972]) e Bell (1984).

2.1.2.1.3.4 Monitoramento

A constituição do *corpus* com falantes da região Nordeste foi resultado de dados coletados exclusivamente para o Projeto ALiB, perfazendo 30 localidades. Em se tratando da aplicação do questionário pré-elaborado as gravações caracterizaram-se como não-secretas, pois o entrevistador conscientizou os informantes de que as respostas dadas ao questionário seriam gravadas. Talvez por estar diante de um gravador, o falante passe a medir suas palavras, se monitora linguisticamente a fim de produzir o efeito desejado. Acreditamos que como resultado desta monitoração, sua fala tende a se aproximar mais da forma padrão.

A pesquisadora Scherre et al. (2015) salientam a respeito das limitações encontradas ao realizar estudos sobre a variação *tu* e *você* a partir de entrevistas sociolinguísticas, enfatizando que apesar das entrevistas terem como objetivo captar a fala espontânea do entrevistado, em alguns casos isso não acontece, podendo assim não demonstrar toda a realidade dos usos da língua por muitas vezes o entrevistado evitar trocas de turno ou por não ter de forma natural toda a gama de relações interacionais. Porém, as autoras afirmam que “qualquer coleta de dados é sempre um pequeno recorte da realidade” (SCHERRE et al. 2015, p. 135).

Nessa perspectiva, queremos ressaltar que nas entrevistas, num momento de maior monitoramento da fala e com a consciência de que a conversa está sendo gravada, os falantes tendem a prestar mais atenção à fala, evitando a utilização do *tu*, uma vez que esse pronome deixa explícito sua marca de mais informalidade e mais envolvimento entre os participantes do evento comunicativo.

Consoante Bortoni-Ricardo (2004, p. 63), ao longo de uma conversa, o falante pode estar movido por fatores que o levam à monitoração do estilo, a saber:

- o ambiente,
- o interlocutor e
- o tópico da conversa.

Assim, um mesmo interlocutor pode optar por um estilo mais monitorado em detrimento de um menos monitorado, em função do alinhamento que assume em relação ao tópico da conversa e ao próprio interlocutor. Todavia, na mudança de um estilo para outro, damos pistas, ou “metamensagens”, nas palavras de Bortoni-Ricardo (2004, p. 63), sendo estas verbais ou não verbais que transmitem informações (alertas), do tipo: ‘isso é uma brincadeira’, ‘estou falando sério’, ‘estou ralhando com você’. “A variação ao longo do contínuo de monitoração estilística tem, portanto, uma função muito importante de situar a interação dentro de uma moldura ou enquadre”.

É cabível aceitar que a monitoração estilística permite-nos observar a avaliação que os falantes fazem das estruturas da língua, na proporção em que a escolha por uma determinada forma em detrimento de outra está condicionada aos “modos de ver” a língua pelo falante.” A monitoração estilística, nesse caso, está relacionada ao problema da avaliação linguística, como proposto pela teoria da variação e mudança linguística” (SOUZA; SANTOS, 2018, p. 14).

2.1.2.1.4 Variação entre fala e escrita ou diamésica

Esse eixo da dimensão externa da variação linguística refere-se às diferenças entre a fala e a escrita. Quando consideramos a condição de produção da fala, temos um texto espontâneo, improvisado e suscetível a variação nos mais diversos níveis, exceto se esse texto é uma palestra. Por outro lado, a escrita caracteriza-se como uma atividade artificial (não espontânea), planejada e monitorada.

“Devido a esse caráter naturalmente improvisado da produção do texto falado, as sequências linguísticas produzidas estão sujeitas a falsos inícios, hesitações, a reformulações e a correções que não podem ocorrer senão no momento em que o texto é produzido” (COELHO, 2010, p. 84). O texto escrito se diferencia do texto falado, uma vez que este dispõe de tempo de planejamento para sua produção bem como há um monitoramento linguístico voltado para a normatização.

É indubitável que o texto falado tem uma organização própria, com entonação particular, reformulações, marcadores discursivos, enfim, um caráter espontâneo. Em relação ao texto escrito, o mesmo se diferencia do texto falado, uma vez que este dispõe de tempo de

planejamento para sua produção bem como há um monitoramento linguístico voltado para a normatização.

Segundo (COELHO, 2010, p. 85), “na variação diamésica é necessário relativizar: assim como não existe uma oposição polarizada entre os registros formal e informal, também não há uma fronteira rígida entre as modalidades oral e escrita da língua”.

Como exemplo de variação diamésica, podemos mencionar a pesquisa de Oliveira (2006) sobre a expressão do futuro verbal em português. A autora constata que predomina na modalidade escrita o futuro simples (*farei*), ao passo que predomina na modalidade falada o futuro perifrástico (*vou fazer*).

2.1.2.2 Variável linguística e variantes

O marco inicial da Teoria da Variação e Mudança Linguística foi estabelecido no trabalho de Weinreich, Labov e Herzog, publicado em 1968, em que houve um rompimento teórico com a corrente de estudos linguísticos que instituía uma identidade entre estrutura e homogeneidade. Contrariando essa visão, foi proposto o estudo da heterogeneidade ordenada da língua, nas palavras dos autores: “a possibilidade de descrever a diferenciação ordenada dentro da língua” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 88). Labov (2008 [1972]), falando sobre essa heterogeneidade das línguas, afirma:

A existência de *variação e de estruturas heterogêneas* nas comunidades de fala investigadas está certamente bem fundamentada nos fatos. É a existência de qualquer outro tipo de comunidade de fala que deve ser posta em dúvida. (LABOV, 2008 [1972], p. 238)

A variabilidade linguística presente em todas as línguas naturais é objeto de estudo da Sociolinguística. Essa variabilidade é o que caracteriza a *variável linguística*. Entendemos por *variantes* as diversas formas alternativas que configuram *um fenômeno variável*, tecnicamente chamado de *variável dependente*. Segundo Labov (2008 [1972], p. 313), *variantes* são duas ou mais formas diferentes de se dizer a mesma coisa com o mesmo valor de verdade. A concordância entre o verbo e o sujeito, por exemplo, é uma variável linguística, uma vez que a mesma se realiza por meio de duas variantes, duas alternativas possíveis e semanticamente equivalentes: a marca de concordância no verbo ou a ausência da marca de concordância.

Assim, cada variável linguística é como se fosse uma forma abstrata que se materializa nas suas formas variantes. Nesta pesquisa, a variável linguística é a *referência de segunda*

pessoa do singular e suas formas variantes são *tu* e *você*.

Uma variável é concebida como dependente quando o uso das variantes não é aleatório, mas influenciado por grupo de fatores de natureza social ou estrutural. Por conseguinte, as variáveis independentes ou grupo de fatores podem ser de natureza interna ou externa à língua. No bloco de variáveis internas, encontram-se os fatores de natureza: fonológica, morfossintática, semântica, discursiva e lexical. Estes referem-se às características da língua em várias dimensões, considerando o nível do significante e do significado, bem com os diversos subsistemas de uma língua. Já no bloco das variáveis externas à língua, juntam-se os fatores inerentes ao indivíduo, a saber: sexo, etnia, escolarização, classe social, nível de renda, profissão etc.

2.1.2.3 Mudança linguística

Entendemos que as variantes de uma mesma variável podem competir em determinados contextos, o que ocasiona duas situações possíveis: (i) tal competição pode ocorrer indefinidamente; ou (ii) uma das formas pode se sobrepor à outra, o que acarreta um processo de mudança.

O fenômeno variável deve ter alta frequência de uso e deve, sobretudo, estar sujeito à influência de fatores sociais e de fatores linguísticos. Além disso, a proposta de regra variável de Labov (2008 [1972]) implica as seguintes questões: (i) a necessidade de um grande número de dados, para que a regra variável seja analisada quantitativamente; (ii) o objeto de descrição deve ser necessariamente um dado real; e (iii) a análise deve envolver diferentes fatores independentes, uma vez que esses são responsáveis por influenciar a alternância das formas.

Nem todas as mudanças são altamente estruturadas, e nenhuma mudança ocorre num vácuo social. Até mesmo a mudança em cadeia mais sistemática ocorre num tempo e num lugar específicos, o que exige uma explicação. (LABOV, 2008 [1972], p. 20)

A variação é a condição essencial para que a mudança linguística ocorra, embora nem todo processo de variação acarrete, necessariamente, uma mudança. Dessa maneira, pelo fato de a mudança não ocorrer de maneira repentina e aleatória, é possível identificarmos os diferentes momentos do processo de mudança. Conforme Weinreich, Labov e Herzog (2006 [1968]), para o estudo do processo de mudança, cinco problemas empíricos precisam ser considerados:

1. O problema das restrições (*the constraints problem*): define exatamente as possibilidades estruturais que restringem e definem as possíveis mudanças linguísticas;
2. O problema da transição (*the transition problem*): relaciona-se com os diferentes estágios do processo de mudança, isto é, o falante alterna o uso das formas variantes em alguns contextos, até que se torne principal em todos os contextos existentes;
3. O problema do encaixamento (*the embedding problem*): de acordo com os autores, toda variação/mudança acarreta e decorre de outras mudanças, seja em nível linguístico ou social;
4. O problema da avaliação (*the evaluation problem*): os falantes não teriam consciência nos estágios iniciais da mudança, entretanto, em estágios mais avançados, tal consciência começa a se tornar possível e perceptível ao falante. Nesse sentido, torna-se importante saber como os falantes avaliam a mudança e quais valores são passados através dela, além do fato de que uma possível avaliação negativa poderia interferir no processo de mudança;
5. O problema da implementação (*the actuation problem*): a compreensão dos mecanismos de implementação da mudança é movida pela seguinte indagação: por que razão um determinado processo de mudança linguística aconteceu em certo momento e lugar e não em outro?

Nesta tese, buscamos validar algumas das questões apresentadas por Weinreich, Labov e Herzog (2006 [1968]), considerando, principalmente, os problemas da restrição, da transição e do encaixamento. O primeiro é bastante elementar nos estudos variacionistas pelo controle que se faz dos fatores que condicionam o fenômeno variável, no caso, a variação entre as formas de segunda pessoa. O segundo refere-se ao mapeamento realizado para identificar estágios e/ou etapas para incremento do pronome *você* em referência à segunda pessoa, além da identificação dos valores que essa forma foi assumindo. Por último, temos o problema do encaixamento, pois uma das hipóteses de trabalho é a de que a forma *você* é predominante entre os falantes de algumas localidades do Nordeste.

O papel da mudança linguística é muito importante para os estudos da Sociolinguística. Sendo assim, o linguista deve compreender que

A diferença entre uma mudança em progresso e uma mudança avançada pode ser vista claramente às vezes no padrão da distribuição social. Uma mudança

pode começar primeiro num grupo social localizado em qualquer ponto da hierarquia social. Enquanto está se desenvolvendo e se expandindo, ainda se pode ver o padrão em pirâmide através de diversas faixas etárias, com os valores mais altos nos falantes mais jovens do grupo original. (LABOV, 2008 [1972], p. 336)

Percebemos, portanto, que o linguista deve definir se o caso é de variação estável ou de mudança em progresso.

No que concerne ao mecanismo da mudança linguística, Labov (2008 [1972], p. 192) observa que há cinco questões importantes que orientam a evolução linguística:

1. Existe uma direção geral para a evolução linguística?
2. Quais são os condicionantes universais da mudança linguística?
3. Quais as causas do surgimento contínuo de novas mudanças linguísticas?
4. Por meio de quais mecanismos as mudanças ocorrem?
5. Existe uma função adaptativa na evolução linguística?

Segundo Labov (2008 [1972], p. 192), algumas dessas questões sobre o mecanismo da mudança linguística podem ser melhor estudadas com base numa análise sincrônica detalhada da mudança linguística em “tempo aparente”, que se refere ao “padrão de distribuição do comportamento linguístico através de vários grupos etários num determinado momento do tempo”. Entende-se que, “se o uso da variante inovadora for mais frequente entre os jovens, decrescendo em relação à idade dos grupos mais idosos, tudo indica que se trata de uma situação de mudança em progresso” (MONTEIRO, 2000, p. 132).

Por seu turno, a mudança em “tempo real” refere-se ao aspecto diacrônico da língua ao analisar a evolução linguística num período arbitrário de tempo. É possível investigar a mudança em tempo real a partir de dois métodos básicos: o primeiro consiste em buscar textos que no passado registraram as variantes em estudo na intenção de compará-las com registros mais novos. Porém esta não é uma tarefa simples, haja vista que a Sociolinguística Quantitativa desenvolveu-se a partir da década de 1960, com Labov, e as diferenças metodológicas entre as pesquisas atuais e os estudos anteriores podem dificultar as comparações. O segundo método é mais difícil e elaborado que o primeiro, consiste em retornar à mesma comunidade depois de um certo período de tempo (dez ou vinte anos) e repetir o estudo, seguindo a mesma metodologia adotada no primeiro, isto é, realizar novas gravações com os mesmos informantes.

Depois de Weinreich, Labov e Herzog (2006 [1968]) atentarem para os problemas empíricos a serem investigados no processo de mudança, Coelho (2010, p. 106-107) evidencia

seus achados teóricos na formulação de sete princípios gerais para o estudo da mudança linguística, elencados, a seguir:

1. A mudança linguística se inicia quando a generalização de uma dada alternância em um certo subgrupo da comunidade de fala assume o caráter de diferenciação ordenada, isto é, não é uma deriva aleatória;
2. A associação entre estrutura e homogeneidade é uma ilusão. A estrutura linguística inclui a diferenciação ordenada de falantes e de estilos através de regras que regem a variação na comunidade de fala. O domínio de uma língua pressupõe o controle de tais estruturas heterogêneas;
3. Nem toda variabilidade e heterogeneidade na estrutura linguística envolve mudança, mas toda mudança implica variabilidade e heterogeneidade;
4. A generalização da mudança linguística através da estrutura linguística não é nem uniforme nem instantânea. A generalização envolve a correlação de mudanças ao longo do tempo e aparece refletida em diferentes áreas do espaço geográfico;
5. As gramáticas nas quais a mudança linguística ocorre são gramáticas da comunidade de fala. Devido ao fato de as estruturas variáveis contidas no sistema serem determinadas por funções sociais, não é possível falar em gramáticas individuais;
6. A mudança linguística é transmitida dentro da comunidade como um todo. Ela não está confinada a etapas discretas dentro da família. Toda e qualquer descontinuidade encontrada na mudança resulta de descontinuidades específicas observadas dentro da comunidade de fala. É muito mais do que o resultado de diferenças de geração (entre pai e filho);
7. Fatores linguísticos e sociais encontram-se intimamente relacionados no desenvolvimento da mudança linguística. Explicações apenas de um ou outro aspecto falharão ao descrever as regularidades que podem ser observadas nos estudos empíricos do comportamento linguístico.

O objeto central dos estudos de variação linguística é a língua do dia a dia, o que Labov (2008 [1972], p. 244) chamou de vernáculo: “o estilo em que se presta o mínimo de atenção ao monitoramento da fala”. Assim, o estudo da língua, segundo a Teoria da Variação, se processa no seu contexto social e, para Labov (2008 [1972], p. 244), “a melhor forma de se obter bons dados de fala em quantidade suficiente é mediante entrevista individual, gravada”. Essa exigência metodológica de descobrir como as pessoas falam quando não estão sendo sistematicamente observadas e a necessidade de gravá-las para obterem-se os dados faz emergir

o chamado *paradoxo do observador*. Um dos desafios dos pesquisadores nesta corrente de pesquisa é como resolver esse paradoxo. Uma solução apresentada por Labov (2008 [1972], p. 119) é tentar envolver o entrevistado com perguntas que recriem emoções fortes e, assim, ele pode esquecer que está sendo gravado e fazer uso de uma fala mais espontânea.

Mais um aspecto importante relacionado à variação é o fato de que esse fenômeno não está limitado a um dos níveis da gramática: encontramos variação no nível fonético bem como no morfológico, no sintático, no lexical e no discursivo.

No nível fonético, podemos realizar certos ditongos tanto de maneira plena quanto reduzida, como em *caixa/caxa* e em *outro/otro*; no morfológico, encontramos variação, por exemplo, na marcação do plural de alguns nomes (*degraus/degrais, chapéus/chapéis* etc.); na sintaxe, encontramos variação na realização das orações relativas (*Esse é o livro de que eu gosto/Esse é o livro que eu gosto/Esse é o livro que eu gosto dele*); no discursivo, um estudo com dados de Florianópolis (VALLE, 2001) mostra o uso alternado dos marcadores discursivos *sabe?/não tem?/entende?*; em relação ao nível lexical, temos *aipim/mandioca/macaxeira*, com distintos traços regionais. Enfim, a variação linguística não só é um fenômeno inerente às línguas naturais mas também se manifesta em qualquer nível de análise.

2.2 DIALETOLOGIA

A Dialetologia foi responsável pelos primeiros estudos dedicados à variação linguística, embora sem basear seus resultados num número maior de dados para que se pudesse ter um quadro mais claro das variações diatópica e social. Assim, a Dialetologia já interpretava os fatos linguísticos segundo diferenças sociais, profissionais, de nível de escolaridade, etárias, de sexo etc. Ela já faz uso de recursos interpretativos há muito tempo e, mais tarde, estes são aprofundados pela Sociolinguística.

Isto posto, quando se fala em Dialetologia e Sociolinguística, entendem-se ambas como o estudo da diversidade da língua dentro de uma perspectiva sincrônica e concretizada nos atos da fala. Ferreira e Cardoso (1994) salientam que a Sociolinguística e a Dialetologia são consideradas, até certo ponto, sinônimas, uma vez que ambas as disciplinas estudam a língua falada. No entanto, é indiscutível que a Dialetologia trouxe importantes contribuições à Sociolinguística e à Linguística Geral.

A Dialetologia surge no final do século XIX, antes mesmo de os estudos Sociolinguísticos se iniciarem, como resposta a uma afirmação teórica dos Neogramáticos. Ao

questionar alguns pressupostos tradicionais da prática histórico-comparativa, que fazia estudos entre as línguas clássicas, como o latim, o grego, o sânscrito, o persa, dentre outras, a fim de chegar à língua original indo-europeia, esses estudiosos estabeleceram uma orientação metodológica diferenciada e um conjunto de postulados teóricos para a interpretação da mudança linguística.

De acordo com esses pesquisadores (os neogramáticos), as mudanças fonéticas estavam submetidas a uma sistematicidade absoluta, isto é, todas as mudanças fonéticas eram governadas por uma regra, afetando sempre a mesma unidade fônica, no mesmo ambiente, em todas as palavras, não admitindo exceções. Caso houvesse alguma exceção seria ou porque o princípio regular efetivo ainda era desconhecido, ou porque a regularidade da mudança havia sido influenciada pelo processo da comparação.

Inferimos que, no interior do comparativismo do século XIX, a vertente que trabalhava com o subgrupo românico das línguas indo-europeias foi responsável pela motivação do estudo da variação geográfica de uma mesma língua e para o início das especulações sobre a variação social; isso devido às possibilidades de reconstrução de um “latim corrente”, tradicionalmente, chamado de “vulgar”, o qual fazia oposição ao latim clássico (MATTOS E SILVA, 1996).

As variações dialetais conviventes no âmbito das línguas românicas, nas suas fases antigas como modernas, e as fontes não clássicas para o conhecimento do diassistema latino fizeram com que, ainda no século passado, a Linguística Histórica no seu sentido estrito, ou seja, o que estuda a mudança das línguas no tempo, se conjugasse com a Dialetologia, também no seu sentido estrito, isto é, a que estuda a variação espacial das línguas. (MATTOS E SILVA, 1996, p. 24)

A ideia de expor em materiais cartográficos a extensão dos fenômenos linguísticos deve-se, segundo Chambers e Trudgill (1994), ao linguista alemão George Wenker, que, em 1876, empreende um estudo voltado para os fenômenos fonéticos, a partir de um questionário, enviado por correspondência a 40.736 localidades, com 40 frases pequenas que deveriam ser “traduzidas” pelos professores primários do Norte da Alemanha para os seus respectivos dialetos. Em 1881, os mapas elaborados com os dados obtidos por Wenker contrariam a concepção de que uma mudança fonética afeta todas as palavras do mesmo modo, como presumiam os neogramáticos.

O trabalho de Wenker (1881), apesar de não se ter estabelecido um controle sistemático no que se refere às variáveis sociais, representou um grande avanço para a Dialetologia, pois a coleta de dados em diferentes regiões possibilitou a intercomparação entre os mesmos.

A Dialetologia desenvolveu-se principalmente entre 1898 e 1912, com o *Atlas Linguistique de la France* (ALF), de Jules Gilliéron, em colaboração com Edmond, cuja preocupação recaía sobre o vocabulário dos camponeses, atendo-se a um grupo de palavras restritas a uma dada região. Diferentemente de Wenker (1881), Gilliéron (1912) privilegiou a coleta de dados *in loco*. Conforme Gilliéron (1912 apud BRANDÃO, 1991), esta deveria basear-se no fundamento de que o retrato fiel da realidade fonética só poderia ser fornecido por um leigo, visto que a fala desse não estaria “contaminada” pelo conhecimento, expectativa ou preconceitos linguísticos. Para ele, os fatos linguísticos deveriam ser estudados tomando-se como parâmetro a sua distribuição espacial, sendo os pontos de inquérito determinados conforme a equidistância geográfica. Com o trabalho de Gilliéron (1912), inaugurou-se o método da geografia linguística, na França.

Nesse novo método de investigação geográfica, a técnica de escolha dos informantes não é pré-determinada nem registrada nas cartas do ALF, embora seja depreensível a partir do exame do perfil dos informantes: as idades se estendem dos 15 aos 85 anos, o número de informantes do sexo feminino é menor do que os do sexo masculino e o nível de escolaridade é identificado por dedução das categorias de informantes (CARDOSO, 2001a, p. 28).

Seguindo Cardoso (2010, p. 15), entendemos por Dialetologia um ramo dos estudos linguísticos que assume a tarefa de identificar, descrever e situar os diferentes usos em que uma língua se diversifica, consoante a sua distribuição espacial, sociocultural e cronológica. As pesquisas nessa linha, como já dito anteriormente, procuram observar as relações entre espaço geográfico e fatos linguísticos no intento de compreender o fenômeno da variação linguística a partir dos estudos feitos em campo de pesquisa (*in loco*). Conforme Oliveira (2005), o mérito da pesquisa dialetológica volta-se para a preocupação com a diversidade linguística existente no País, em virtude de sua extensão territorial e das influências linguísticas recebidas.

Nos estudos dialetológicos, as cartas têm fundamental importância como instrumento metodológico de descrição. Por meio delas, o pesquisador tem uma ampla observação horizontal e plana de uma dada região a partir de determinados fatores linguísticos. Investigá-las, porém, exige um nível de sondagem maior do que a superfície de um mapa pode oferecer e leva o pesquisador a diferentes interpretações diante de questões que devem ser consideradas.

Nesse sentido, o resultado mais imediato do trabalho dialetológico são as cartas, ferramentas nas quais os resultados obtidos nas pesquisas de campo são expostos e oferecem, entre outras possibilidades, a de encontrar isoglossas, com as quais se torna possível a definição de áreas dialetais.

É importante lembrar que os limites dialetais não coincidem, necessariamente, com os limites políticos. O dialetólogo, ao definir uma rede de pontos para coleta de dados por meio de pesquisa de campo, não dá conta do emaranhado de questões sociais, históricas e culturais que caracteriza uma localidade isolada ou uma microrregião.

Para o desenvolvimento de pesquisas dialetológicas, especialmente no que concerne à produção de atlas linguísticos, tornou-se necessário o estabelecimento de normas gerais para orientar e coordenar a pesquisa *in loco*. Desse modo, as pesquisas dialetais seguem alguns passos metodológicos na intenção de bons resultados na investigação. Ferreira e Cardoso (1994) determinam quatro etapas principais que devem ser observadas na pesquisa dialetal:

- a preparação da pesquisa;
- a execução dos inquéritos;
- a explicitação e a análise dos materiais recolhidos;
- a divulgação dos resultados obtidos.

Por seu turno, as pesquisas dialetais confirmam a relevância dos trabalhos linguísticos para o conhecimento da língua e os aportes metodológicos servem para disciplinar a investigação, permitindo, por conseguinte, que os pesquisadores alcancem com mais critério e científicidade os seus objetivos.

2.2.1 Dialetos

A Dialetologia é, como a própria palavra sugere, o estudo dos dialetos. Salientamos que não há um consenso entre os estudiosos em relação ao conceito de dialeto, uma vez que alguns pesquisadores consideram o critério geográfico e outros, os critérios sociais, o que originaria os dialetos sociais ou os socioletos. Na linguagem coloquial, diz-se que um dialeto é uma língua não padrão, que geralmente se associa a grupos que não ostentam prestígio ou uma língua que não tem tradição escrita. Entretanto Chambers e Trudgill (1994, p. 19) não aprovam estes conceitos e dizem que “todos os falantes são falantes de, pelo menos, um dialeto” e que a forma padrão de uma língua constitui-se em si um dialeto. Esses autores afirmam que uma língua é um conjunto de dialetos mutuamente inteligíveis, embora evidenciem que tal definição não seja totalmente satisfatória porque esta tal inteligibilidade nos traz alguns problemas.

Leite de Vasconcelos (1970), considerado um dos maiores nomes da Dialetologia portuguesa, com a finalidade de sistematizar o estudo da língua portuguesa, divide-a em dialetos, reconhecendo que estes podem se subdividir em subdialetos e variedades. O

dialetólogo português guiou-se pelo critério geográfico para a sua classificação dos dialetos em: continentais, insulares e ultramarinos. O pesquisador considera a língua portuguesa do Brasil entre os dialetos de ultramar, evidenciando ainda que o português do Brasil é um dialeto do português europeu. Na nossa opinião, a língua portuguesa foi passada para nós de maneira irregular, em um contínuo de níveis diferenciados de socialização, uma segunda língua adquirida massivamente, de maneira quase “imperfeita”, em contextos sócio-históricos específicos (LUCCHESI, 2003, p. 274).

Ferreira e Cardoso (1994), para definir dialeto, mobilizam o conceito de isoglossa, pois, para as pesquisadoras, o dialeto define-se como um feixe de isoglossas, ou seja, “um conjunto de isoglossas que se somam e que, portanto, mostram uma relativa homogeneidade dentro de uma comunidade linguística em confronto com outras” (FERREIRA; CARDOSO, 1994, p. 16).

Chambers e Trudgill (1994) também traçam um paralelo entre o conceito de dialeto e sotaque. Segundo eles, sotaque caracteriza-se pela forma como um falante pronuncia a língua e, portanto, se aplica à variação fonética ou fonológica da língua; ao passo que dialeto refere-se às variedades que são diferentes do ponto de vista gramatical ou lexical, além do nível fônico. Contudo eles esclarecem que os sotaques e os dialetos frequentemente se fundem uns com os outros sem que tenham uma separação definida. Esses autores ressaltam que, embora essas observações acerca de diferenças dialetais sejam muito comuns, o estudo dos dialetos só começa de maneira sistemática na segunda metade do século XIX.

2.2.2 Sobre a dialetação do português brasileiro: as propostas de divisão dialetal

Em relação ao desenvolvimento dos estudos dialetais do Brasil, podem ser observadas diferentes fases que são classificadas segundo a predominância de produção de cada época. Não há, assim, uma data fixa para cada fase nem, tampouco, uma classificação única, uma vez que não são fases estanques, mas propostas que se complementam na medida em que avançam as pesquisas nessa área.

A primeira proposta de ordenação dos estudos dialetais em nosso País, e que até hoje é ponto de referência entre dialetólogos e geolinguistas, é de autoria de Antenor Nascentes (1952), na qual o autor sugere duas fases: a **primeira** inicia-se com a publicação do estudo feito pelo Visconde de Pedra Branca, em 1826, e vai até 1920. Esta fase caracteriza-se, principalmente, por obras de caráter lexicográfico. É o momento dos dicionários, dos glossários

regionais e dos vocabulários; a **segunda** tem como marco inicial a publicação de *O dialeto caipira*, de Amadeu Amaral, em 1920, e estende-se até nossos dias.

Desse modo, é dada a partida para a história dos estudos dialetais no Brasil. Cardoso e Ferreira (1994, p. 37-62) sugerem **três diferentes fases**, reformulando, portanto, a periodização proposta por Nascentes (1952; 1953). E a mais recente proposta de Cardoso e Mota (2005) considera **quatro fases**.

A primeira manifestação específica de natureza dialetal no Brasil foi o capítulo escrito pelo Visconde de Pedra Branca para o *Atlas Ethnographique du Globe*, de Adrien Balbi, em 1826, no qual apontava caracteres comparativos do português brasileiro com o português europeu em relação ao léxico. A partir desse trabalho, marcou-se o início da **primeira fase** e, assim, surgem outros trabalhos voltados para o léxico, para a produção de glossários e dicionários. Exemplificamos algumas obras publicadas nesse período, cujo interesse era o estudo do léxico e de suas especificidades no português do Brasil, a saber: *Glossário de vocábulos brasileiros, tanto dos derivados como daqueles cuja origem é ignorada*, publicado pelo visconde de Beaurepaire-Rohan, na *Gazeta Literária*, de 1883 a 1884, posteriormente transformado em *Dicionário de vocábulos brasileiros* (1976); *O tupi na geografia nacional* (1901), de Theodoro Sampaio; *Glossário paraense* (1905); *A criação de gado no Marajó* (1912) (CARDOSO, 2010).

A **segunda fase**, voltada para a produção monográfica, atendo-se à realidade de áreas específicas, com propostas de abordagens em níveis fonético, morfológico, sintático e semântico, teve como marco inicial o trabalho produzido por Amadeu Amaral em 1920, *O dialeto caipira*. Caracteriza-se por trabalhos voltados para os estudos gramaticais, embora os lexicográficos continuem numerosos. Amaral (1955[1920]) é o destaque desta fase, pois, sem contar com as técnicas atuais de pesquisa, procurou adotar uma metodologia de abordagem voltada para a nossa realidade linguística. Amaral (1955[1920], p. 42) faz um estudo sobre a linguagem popular e diz que pretende “[...] caracterizar esse dialeto ‘caipira’”, ou, se assim acham melhor, *esse aspecto da dialetação portuguesa em São Paulo*. Esclarece que os fonemas do dialeto são os mesmos do português, apresentando ligeiras variantes fisiológicas.

Cardoso (2010) salienta que nessa fase percebe-se uma preocupação com uma metodologia de abordagem dos fenômenos orientada para o exame da realidade *in loco*, buscando, portanto, uma nova metodologia de trabalho na tentativa de fornecer dados nos diferentes níveis da língua.

Dois anos depois de publicado *O dialeto caipira*, Antenor Nascentes publica, em 1922, *O linguajar caipira*, obra que, depois da segunda edição, passa a chamar-se *O linguajar carioca* (1953). O referido autor demonstra uma preocupação em definir o que entende por falar brasileiro e em situar o linguajar carioca no conjunto desses falares. Ele discute “o processo de dialetação do português do Brasil, apresenta uma divisão dos falares brasileiros a que declara ter chegado depois de haver realizado o ardente desejo de percorrer todo o Brasil, do Oiapoque ao Chuí, de Recife a Cuiabá” (CARDOSO, 2010, p. 135).

A obra de Nascentes (1922), além de ajudar conhecer o dialeto carioca, uma vez que traz aspectos fonéticos, morfológicos, sintáticos e lexicais, destaca-se também por apresentar uma proposta de divisão dialetal do Brasil, a qual é a primeira.

Já a **terceira fase** foi marcada pelo Decreto 30.643 de 20 de março de 1952, que define como principal finalidade da Comissão de Filologia da Casa de Rui Barbosa a elaboração de um Atlas Linguístico do Brasil. Dessa forma, os estudos geolinguísticos tomam um novo rumo. Nessa ocasião, divulga-se o *Atlas Prévio dos Falares Baianos* (APFB), sob a coordenação do professor Nelson Rossi, primeiro atlas linguístico produzido no Brasil. Nessa fase, os pesquisadores Antenor Nascentes, Serafim da Silva Neto, Celso Cunha e Nelson Rossi se destacaram, pois, graças aos seus esforços, deu-se início aos estudos de geografia linguística no Brasil.

Ainda nesse período, Antenor Nascentes publica as *Bases para elaboração do atlas linguístico do Brasil*, obra em dois volumes – o primeiro em 1958; o segundo em 1961 – na qual se firmam passos fundamentais para o início dos estudos dialetais. O autor

[...] discute, na pequena introdução, as vantagens de um atlas feito, ao mesmo tempo, para todo o país, mas reconhece a impossibilidade de sua concretização, nesses termos entre nós, argumentando com a vastidão do nosso território e a dificuldade de acesso aos diferentes pontos. (CARDOSO, 2010, p. 139-140)

Segundo a proposta de Cardoso e Mota (2005), a **quarta fase** tem como marco a retomada de um projeto do Atlas Linguístico do Brasil, em 1996, por ocasião do *Seminário Caminhos e Perspectivas da Geolinguística no Brasil*, em Salvador. A nova postura metodológica impressa ao Projeto Atlas Linguístico do Brasil (Projeto ALiB) pelo tratamento pluridimensional dado aos fatos linguísticos é um dos aspectos que, segundo Cardoso e Mota (2005), confirmam a escolha do Projeto ALiB como marco divisório entre a terceira e a quarta fases da história dos estudos dialetais no Brasil. A **quarta fase** se caracteriza pela

implementação de pesquisas na área da geolinguística, como se observa nos trabalhos de pós-graduação, em comunicações em congressos e nos vários projetos de atlas regionais; pela visão pluridimensional e incorporação de dados em níveis morfossintático, pragmático-discursivo e metalinguístico; e pela relação com outras ciências, principalmente, com a Etnolinguística.

Chega-se à conclusão de que, na segunda proposta de Cardoso e Ferreira (CARDOSO, 2003), a qual aponta para três diferentes tendências ou fases, nas duas primeiras fases, há uma coincidência com a proposta de Nascentes (1952), apresentando apenas uma pequena diferença no que diz respeito à subdivisão da segunda fase em quatro grupos, consoante características comuns.

A terceira e última proposta, de autoria de Cardoso e Mota (2005), é igual à segunda, isto é, possui três fases. A diferença, entre esta e aquela, reside no fato de que as autoras, ao revisarem a divisão da história dos estudos dialetais no Brasil, decidiram acrescentar a esta proposta uma quarta fase que inclui, além da construção do Atlas Linguístico do Brasil (Projeto ALiB), as inovações dos estudos dialetais brasileiros bem como os avanços da geolinguística em direção a outros atlas regionais.

Acrescentamos aqui uma **quinta fase**, proposta lançada por Ana Regina Teles (2018). Segundo a pesquisadora, a publicação dos dois primeiros volumes do ALiB, que ocorreu na Universidade Estadual de Londrina (UEL), em Londrina, durante o *III Congresso Internacional de Dialetologia e Sociolinguística* (CIDS), no ano de 2014, é um marco que dá início a essa quinta fase.

Ao longo de 18 anos, após a implantação do Projeto, a publicação desses dois volumes foi uma referência para a Geolinguística brasileira, uma vez que muitos trabalhos têm sido desenvolvidos a partir da publicação desses volumes. Além disso, é notório o efeito atingido pelo ALiB nas mídias sociais e em bibliotecas de diferentes Universidades.

Assim se manifesta Teles (2018):

[...] 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 as cartas publicadas passíveis de reedições e complementações a qualquer época. (TELES, 2018, p. 76)

É fato que a publicação desses dois primeiros volumes do *Atlas Linguístico do Brasil* veio corroborar as pesquisas dialetais. Percebe-se também uma maior expressividade no cenário da Dialetologia e da Sociolinguística; prova disso são os relatórios encontrados em *Documentos 7*, obra publicada em 2017, na qual se disponibiliza a história de 20 anos do ALiB.

2.2.3 Geolinguística: geografia linguística no Brasil

Segundo Cardoso (2016),

Para dar cumprimento ao objetivo de detectar e descrever a variação, a Dialetologia recorre ao seu método por excelência, a *Geografia Linguística*, ou, como atualmente vem sendo denominada, a *Geolinguística*, descrita de forma exemplar por Coseriu, em *La geografia linguística* (1955). Essa técnica metodológica consiste em apresentar os dados linguísticos sob a forma de mapa ou carta geográfica, distribuídos por pontos espacialmente identificados. (CARDOSO, 2016, p. 17)

A Geolinguística oportuniza, a partir do estado atual da língua, reviver as palavras antigas esquecidas no tempo. Segundo Busse (2010, p. 55), “é possível mesmo perceber nessas palavras, ecoadas no tempo, o objeto mesmo da Geolinguística: os dialetos, tomados como formas pelas quais é possível determinar a ‘estratigrafia’ de uma palavra”.

Desse modo, a Dialetologia tem como função descrever os diferentes dialetos, estudando a variação diatópica pelo método geolinguístico. Por meio desses estudos, verificamos as relações entre o espaço geográfico e os fenômenos linguísticos apresentados, conhecendo, assim, a norma diatópica da localidade investigada. Em síntese, a Dialetologia não deve ser confundida com geografia linguística ou geolinguística, pois esta é um método utilizado pela Dialetologia. Segundo Cristianini (2007),

considerando que a ideia de norma está definitivamente caracterizada pela presença da alta frequência e pela distribuição regular das variações em uma comunidade linguística, e que cabe à Geolinguística a descrição das variações diatópicas, poderemos especificar, por meio das respostas obtidas de entrevistas aplicadas a sujeitos de um certo grupo numa dada localização, a norma característica do local, além de precisar a repetição topográfica dos fenômenos apresentados em cartogramas. (CRISTIANINI, 2007, p. 50)

Conforme Cardoso (1996, p. 130-131), “é reconhecida a relevância da contribuição que esse ramo de estudos da linguagem pode trazer à reconstituição da história, ao entendimento da organização demográfica, às questões de natureza antropológica, em geral, e ao próprio ensino

de língua materna”. A pesquisadora, quando questiona o ensino da língua materna, alerta sobre a contribuição dos estudos dialetológicos, pois é na escola que se tem o contato com a mistura de vários dialetos, os quais devem ser respeitados, já que renegá-los é o mesmo que ignorar a história de nosso País.

2.2.4 Os atlas brasileiros

No Brasil, até o momento, foram concluídos dezessete atlas linguísticos, tendo onze deles sido publicados: *Atlas Prévio dos Falares Baianos* (1963), *Esboço de um Atlas Linguístico de Minas Gerais* (1977), *Atlas Linguístico da Paraíba* (1984), *Atlas Lingüístico de Sergipe* (1987), *Atlas Lingüístico do Paraná* (1994), *Atlas Lingüístico-Etnográfico da Região Sul* (2002), *Atlas Lingüístico de Sergipe II* (2005), *Atlas Lingüístico Sonoro do Estado do Pará* (2004), *Atlas Lingüístico do Mato Grosso do Sul* (2007), *Atlas Linguístico do Estado do Ceará*, *Atlas Linguístico do Amapá* (2017). Os atlas ainda não publicados, ou seja, que estão em andamento, são: *Atlas Linguístico do Espírito Santo*, *Atlas Linguístico do Mato Grosso*, *Atlas Linguístico do Maranhão*, *Atlas Linguístico do Rio Grande do Norte*, *Atlas Linguístico do Acre*, *Atlas Linguístico de Pernambuco*, *Atlas Linguístico de Rondônia*, *Atlas Linguístico do Pará*, *Atlas Linguístico de Goiás* e *Atlas Linguístico do Piauí*.

Além desses atlas estaduais e regional, foram publicados em 2014 os dois primeiros volumes do *Atlas Linguístico do Brasil*, com os próximos volumes já em andamento.

Os primeiros atlas publicados exibem um intervalo de tempo mais longo entre uma publicação e outra e se concentram na região Nordeste do país. Recentemente, diminui-se esse período de publicação entre os atlas, mesmo porque outras regiões brasileiras também já possuem atlas publicados, a exemplo da região Sul e da região Norte.

Os Atlas Regionais (AR) servindo de suporte para o conhecimento mais efetivo das particularidades sociolinguísticas de cada região contribuíram para a elaboração do Atlas Linguístico do Brasil. Este recobre toda a área nacional e proporciona uma visão global do português brasileiro, mas não permite de *per si* uma visão detalhada de cada área particular. Assim, o ALiB apoia a elaboração de atlas linguísticos regionais, por considerar que podem oferecer um conhecimento mais detalhado e circunstanciado de cada região.

O Brasil já conta com vários atlas linguísticos estaduais, concluídos e publicados, outros concluídos e não publicados, um regional publicado e mais o *Atlas Fonético do Entorno da Baía da Guanabara*, recentemente concluído e nove atlas em andamento. A realização desses

atlas se deu, inicialmente, como consequência das dificuldades para elaboração de um atlas nacional em um país de extensão territorial tão grande como o nosso. O mapeamento regional do país foi apontado pelos primeiros dialetólogos brasileiros como uma alternativa e uma etapa necessária para a execução do objetivo maior: um atlas linguístico de todo o Brasil.

A experiência resultante desses atlas tem sido imprescindível para a execução do Projeto ALiB, valendo ressaltar, hoje, a consciência dos pesquisadores brasileiros de que a elaboração do atlas nacional não invalida a realização de outros atlas estaduais ou regionais, uma vez que o primeiro visa a uma descrição geral da situação linguística do Brasil e os últimos permitem uma investigação de aspectos mais específicos de cada área, Estado ou região.

Seguem abaixo em ordem de publicação e elaboração, os Atlas concluídos e publicados:

(i) O Atlas Prévio dos Falares Baianos – APFB/1963

Primeiro atlas linguístico brasileiro, o APFB foi elaborado entre os anos 1960 e 1962 e publicado em 1963. Realizado sob a direção do professor Nelson Rossi (UFBA), contou com a colaboração de Dinah Maria Montenegro Isensee e Carlota da Silveira Ferreira. Contou com uma rede de pontos de 50 localidades, distribuídas pelas 16 zonas fisiográficas do Estado, e com um extrato de um questionário de 164 perguntas, selecionadas com base em uma versão de questionário mais ampla com 3.000 questões, divididas nas áreas semânticas *terra, vegetais, homem e animais*.

O atlas teve um total de 100 informantes, 57 mulheres e 43 homens, com idade variando entre 25 e 60 anos. Com relação à escolaridade, todos eram analfabetos ou semialfabetizados. O APFB apresenta, como inovação em trabalhos dessa natureza, a aplicação de um teste de reconhecimento ou teste de identificação: após a aplicação do inquérito, não obtidas certas respostas, indagou-se ao informante o conhecimento de determinadas expressões, obtidas em sondagens iniciais.

Recobrindo todo o Estado da Bahia, constitui-se de 209 cartas, compreendendo 11 de identificação, 154 fonéticas e léxicas e 44 cartas resumo. Os termos vêm transcritos no interior da própria carta ou através de legendas e símbolos, em preto e branco e colorido. Algumas cartas apresentam dados etnográficos, muitas delas acompanhadas de ilustrações de objetos.

O Atlas Prévio dos Falares Baianos constitui um marco nos estudos da Geografia Linguística no Brasil, não só por ter sido o primeiro trabalho a ser publicado mas por sua fundamental importância para o conhecimento do falar regional da Bahia e, por extensão, de grande parte do falar nordestino. Outro marco importante na realização deste Atlas é a sua

natureza metodológica: a coleta dos dados foi feita em base auditiva, sem auxílio de gravadores com os quais se realizam, atualmente, trabalhos de campo.

(ii) O Esboço de um Atlas Lingüístico de Minas Gerais – EALMG/1977

Os professores Mário Roberto Lobuglio Zágari, José Ribeiro, José Passini e Antônio Gaio foram os mentores deste atlas. O EALMG teve seu primeiro volume publicado em 1977, sendo assim o segundo atlas linguístico publicado no Brasil. Alguns pesquisadores colaboraram com a realização do EALMG: Cláudia Coutinho, Edimilson Pereira, José Dionísio Ladeira e Núbia Magalhães Gomes.

Constitui-se de uma rede de 116 localidades e 83 informantes, dentre eles analfabetos, pessoas com primário incompleto (ALiB – Documento 2, 2006 p.38) e de nível superior. O EALMG une métodos tradicionais da pesquisa geolinguística com outros modernos, oriundos da sociolinguística norte-americana, não se restringindo, assim, mais ao informante do tipo "HARAS" (homem, adulto, rurícola, analfabeto e sedentário), permitindo, portanto, a análise de outros níveis de variação linguística. As pesquisas provenientes do atlas apontam para a confirmação da existência de três falares distintos no território mineiro: o falar baiano ao Norte, o falar paulista no Sul-Sudeste e o falar mineiro no Centro-Leste (PEREIRA, 2007).

(iii) O Atlas Lingüístico da Paraíba – ALPB/1984

De autoria de Maria do Socorro Silva de Aragão e Cleusa Bezerra, foi o terceiro atlas regional publicado no Brasil, o qual faz parte de um projeto mais amplo do Curso de Pós-Graduação em Letras da Universidade Federal da Paraíba, a saber: o Levantamento Paradigma-Sintagmático do Léxico Paraibano. O ALPB contou com uma exaustiva etapa de preparação de sua equipe de pesquisadores e com um intenso trabalho de levantamento bibliográfico relativo à Dialetologia e à Geolinguística no Nordeste, no Brasil e no exterior.

A rede de pontos recobre 25 municípios base, mais três municípios satélites para cada base (localidades satélite não constam da cartografia). O questionário aplicado⁶, fruto do aperfeiçoamento de três outras versões, compõe-se de duas partes: uma geral, com 289 questões, e uma específica, com 588.

A parte geral compreende os campos semânticos Terra, Homem, Família, Habitação e Utensílios Domésticos, Aves e Animais, Plantação e Atividades Sociais, modelo adotado pela

⁶ Publicado pela Universidade Federal da Paraíba.

equipe do APFB. A específica se refere aos cinco principais produtos agrícolas do Estado: mandioca, cana-de-açúcar, agave, algodão e abacaxi.

(iv) O Atlas Lingüístico de Sergipe – ALS I/1987

Escolheu-se o Estado de Sergipe para dar prosseguimento ao trabalho feito na Bahia. O ALS I constitui um passo a mais no mapeamento linguístico da área dos falares baianos pela equipe de Dialetologia do Instituto de Letras da Universidade Federal da Bahia, iniciado com a elaboração do *Atlas Prévio dos Falares Baianos* (APFB – 1963). O ALS foi o quarto atlas linguístico brasileiro a ser publicado, embora já estivesse concluído há mais de dez anos. Seus autores são: Nelson Rossi, Carlota Ferreira, Judith Freitas, Nadja Andrade, Suzana Cardoso, Vera Rollemburg e Jacyra Mota. Em Sergipe, foi possível aperfeiçoar os instrumentos metodológicos utilizados na Bahia. Nesse sentido, o questionário definitivo do ALS, resultado de uma seleção de duas versões preliminares testadas, teve um total de 686 questões: 164 retiradas do Extrato de Questionário aplicado para o APFB e 522 selecionadas dos questionários preliminares.

Examinando o conjunto de perguntas e transcritas as informações, verificou-se, no entanto, que nem todos os itens ofereciam respostas que justificassem um tratamento cartográfico. Assim, das perguntas consideradas, apenas 105 têm os seus resultados cartografados.

As questões recobriam as mesmas áreas semânticas do APFB. Também em conformidade com a metodologia do APFB, foi aplicado um "teste de identificação", referente tanto a formas obtidas no APFB quanto a outras identificadas preliminarmente na área do inquérito.

A distribuição de pontos contou com um total de quinze localidades, distribuídas por cinco zonas fisiográficas do Estado. Das quinze localidades, sete coincidem com pontos da proposta de Nascentes (1958). O perfil dos informantes segue as mesmas características adotadas nos atlas anteriores.

(v) O Atlas Lingüístico do Paraná – ALPR/1994

Esse atlas foi resultado da tese de Doutorado da Professora Vanderci de Andrade Aguilera, pela Universidade Estadual Paulista (UNESP), defendida em novembro de 1990; foi o quinto atlas linguístico do País.

O ALPR apresenta como objetivo, além da documentação cartográfica da variação lexical e da variação fonética e a delimitação de isoglossas, a organização de um glossário, no qual se registra "todo vocabulário cuja forma e/ou sentido" não pertence" ao vocabulário ativo de um falante da norma padrão urbana" (AGUILERA, 1994, p 12).

Como metodologia do questionário, seguiu-se, basicamente, a mesma do Atlas Linguístico do Estado de São Paulo (ALESP), contendo 325 questões e abrangendo dois campos semânticos Terra e Homem, subdivididos em: Terra: (a) natureza, fenômenos atmosféricos, astros, tempo; (b) flora; (c) plantas medicinais; (d) fauna; Homem: (a) partes do corpo, funções, doenças; (b) vestuário e calçados; (c) agricultura, instrumentos agrícolas; (d) brinquedos, jogos infantis; (e) lendas e superstições.

Para seleção da rede de pontos, adotou-se, inicialmente, a proposta de Nascentes (1958, p.21) para o atlas nacional, que continha 24 pontos para o Estado do Paraná. As tais localidades foram acrescentadas outras 41, totalizando 65 localidades, que contemplam 24 microrregiões: são os pontos de número 364 a 387, no rol dos seiscentos e seis municípios brasileiros a fazerem parte da rede de investigação para o Atlas Linguístico do Brasil. Os informantes têm entre 27 e 62 anos, tendo sua escolaridade variando entre analfabetos e com primário completo. Destacaram-se como auxiliares para a pesquisa: Ivone Alves de Lima, Rita de Cássia Paulino e Elaine Cristina Fabris, alunas do Curso de Letras da UEL.

(vi) O Atlas Lingüístico-Etnográfico da Região Sul do Brasil – ALERS/2002

O ALERS foi o sexto atlas brasileiro e inovador ser o primeiro a não se limitar ao mapeamento de um Estado. Coordenado pelo Professor Walter Koch, este atlas abrange aspectos etnográficos da variação espacial a partir da análise dos dados do QSL. A rede de pontos do ALERS conta com 294 localidades, sendo 106 no Paraná, 86 em Santa Catarina e 102 no Rio Grande do Sul.

O ALERS adotou três tipos de questionários, com um total de 711 questões: (a) 26 questões no Questionário Fonético-Fonológico (QFF), além de outras 24 questões para as áreas de colonização não-lusa; (b) 75 no Questionário Morfossintático (QMS); e (c) 610 no Questionário Semântico-Lexical (QSL).

Os informantes têm idade entre 28 e 58 anos e pouca escolaridade, tendo sido escolhidos dois por localidade nas áreas rurais e três nas áreas urbanas. Seus cooperadores foram: José Luiz da Veiga Mercer, Basílio Agostini, Hilda Gomes Vieira, Felício Wessling Margotti, Mário Silfredo Klassmann e Cléo Vilson Altenhofen.

(vii) Atlas Lingüístico Sonoro do Pará – ALISPA/2004

É o primeiro atlas do Brasil que inclui mapeamento automático fonético, além dos arquivos sonoros dos próprios informantes; foi o oitavo dos atlas regionais publicados, tendo como autor e coordenador o professor Abdelhak Razky, da Universidade Federal do Pará. O ALISPA é um dos atlas mais modernos no Brasil por utilizar uma tecnologia digital e informatizada para o mapeamento de dados sonoros no campo da Geolinguística e da Dialetologia.

A digitalização de dados e a elaboração de uma interface eletrônica facilitam o acesso a informações dialetais de forma interativa e instantânea, garantindo um nível de aproximação bem maior das realizações fonéticas do falante-informante do que o permitido pela transcrição fonética, além de assegurar a identificação das variações no tempo e no espaço físico-social. (RAZKY, 2013, p.225)

O ALiSPA permite a visualização automática de 636 cartas fonéticas. A meta desse empreendimento foi a construção de uma fonética dos falares paraenses, a partir da recolha de dados em 10 cidades do Estado, com um total de 40 informantes, sendo quatro por localidade (um homem e uma mulher de 18 a 33 anos e um homem eu uma mulher de 40 a 70 anos). O programa apresenta um mapa das dez cidades escolhidas e permite ao usuário consultar de maneira instantânea as variações fonéticas no Estado, ou seja, “o usuário pode ver e ouvir como uma determinada variável é realizada em cada uma das cidades pesquisadas” (RAZKY, 2013, p. 225). Aplicado em 2001, o Questionário Fonético-Fonológico é constituído de 159 perguntas, retiradas do ALiB. Alguns professores atuaram como colaboradores, a saber: Orlando Cacique, Marilúcia Oliveira e Raquel Lopes.

(viii) Atlas Lingüístico de Sergipe II – ALS II/2005

Publicado em 2005, o ALS II foi desenvolvido como tese de Doutorado da Professora Suzana Alice Marcelino Cardoso, pela Universidade Federal do Rio de Janeiro, em 2002. Foi o sétimo atlas concluído e é o segundo volume do *Atlas Linguístico de Sergipe* (ALS). Utiliza um *corpus* não explorado no primeiro volume. Centrado na área semântica Homem, o ALS-II busca explorar também outras variáveis linguísticas já contempladas no ALS-I, como a diassexual.

Ao lado do caráter pluridimensional impresso nas cartas linguísticas, o ALS-II traz comentários às cartas, nos quais se destacam aspectos diatópicos e sociolinguísticos. Constitui-

se de um conjunto de 108 cartas, 105 das quais são semântico-lexicais e três introdutórias. A rede de pontos constitui-se de 15 localidades, distribuídas por todas as microrregiões do Estado, contando, cada ponto, com dois informantes, identificados com A-mulheres e B-homens e escolhidos conforme os critérios configurados como básicos para os estudos dialetais – nascidos na cidade objeto de estudo, filhos, preferivelmente, de pais da mesma localidade, não-alfabetizados ou semialfabetizados, com afastamento nulo ou por pouco tempo do ponto de residência.

Como o ALS-II apresenta comentários às cartas, segundo uma perspectiva sócio-antropológico-linguística, sobre aspectos salientados pelas informações cartografadas, pode-se classificar esse volume do *Atlas Linguístico de Sergipe* entre aqueles chamados atlas de segunda geração, isto é, aqueles que fornecem os dados e atentam para sua interpretação (PEREIRA, 2007).

(ix) O Atlas Linguístico do Mato Grosso do Sul (ALMS/2007)

A partir de fevereiro de 1996, a coordenação do projeto foi feita por Dercir Pedro de Oliveira. Também fazem parte da equipe desse trabalho as professoras Aparecida Negri Isquierdo, Maria José Gomes, Ana Maria Pinto Pires e Albana Xavier Nogueira. O projeto almeja oferecer aos estudiosos e professores subsídios para incorporar ao ensino de língua a diversidade linguística, apontando as características linguístico-culturais das diferentes regiões do Estado. Nesse sentido, o estudo propõe-se a fazer uso dos fundamentos da Dialetologia e da Geolinguística contemporâneas, tendo como foco principal o estudo da variação diatópica, sem descuidar-se dos aspectos sociais e culturais da linguagem.

Para efetivação da metodologia, foram consideradas na rede de pontos 33 localidades, tendo sido escolhidos apenas os municípios criados até a década de 1980. Os informantes são distribuídos pelas seguintes faixas etárias: Faixa I – de 18 a 30 anos; Faixa II – de 45 a 70 anos. A escolarização está estratificada em: analfabeto e até a 4^a série do Ensino Fundamental. Os indivíduos devem ser nascidos no município ou nele residido desde os 8 anos de idade. Após a realização dos inquéritos, vêm sendo concluídas as transcrições grafemáticas do Questionário Lexical e das questões que visam à coleta de opiniões e de relatos pessoais.

Através da aplicação de dois questionários, um lexical, com 506 perguntas, e um fonético, com 46 questões, tem-se uma estimativa de que sejam documentados quatro informantes por localidade, conferindo as variáveis sociais escolhidas.

(x) O Atlas Linguístico do Estado do Ceará (ALECE)

O ALECE, além de registrar o falar cearense, teve a preocupação de ser “um instrumento capaz de fornecer dados para a reformulação do ensino de língua portuguesa, sobretudo no que se refere ao primeiro segmento do primeiro grau...” (BRANDÃO, 1991, p. 64).

Dirigido por José Rogério Fontenele Bessa, já se encontra concluído e à espera de publicação. Foi distribuído em três volumes: volume I, com introdução, descrição dos procedimentos metodológicos e critérios de seleção de localidades e informantes; volume II, com apresentação das 108 cartas lexicais e das 132 cartas fonéticas; e volume III, contendo um glossário com 908 itens.

Constitui-se de 69 pontos e conta com documentação recolhida de 268 informantes de faixa etária entre 30 e 60 anos, de ambos os性os e analfabetos ou de instrução primária. O questionário tem 306 perguntas.

(xi) O Atlas Linguístico do Estado do Amapá (ALAP)

Publicado em 2017, o *Atlas Linguístico do Amapá* é de autoria da professora Celeste Maria da Rocha Ribeiro (UNIFAP), do professor Abdelhak Razky (UFPA/UNB) e do professor Romário Duarte Sanches (UFPA/UEPA).

Para constituição do *corpus* da pesquisa, utilizou-se dez pontos de inquérito: (01) Macapá, (02) Santana, (03) Mazagão, (04) Laranjal do Jarí, (05) Pedra Branca do Amaparí, (06) Porto Grande, (07) Tartarugalzinho, (08) Amapá, (09) Calçoene e (10) Oiapoque. Foram selecionados 40 informantes, sendo 04 por localidade, sendo 1 homem e 1 mulher de 18 a 30 anos; 1 homem e 1 mulher de 50 a 75 anos, de escolaridade fundamental.

Os questionários utilizados foram: fonético-fonológico (QFF) e semântico lexical (QSL): o QFF será composto por 159 questões fechadas e o QSL por 202 perguntas distribuídas em 14 campos semânticos, com perguntas abertas. O resultado do questionário ajudou na formulação de 16 cartas fonéticas, 73 cartas lexicais e 30 cartas estratificadas.

2.2.4.1 Atlas em andamento (incluindo teses)

Em sua maioria, os atlas que estão em andamento são inspirados na metodologia dos atlas publicados. Esses projetos incorporam também diferentes concepções dos estudos dialetológicos, das mais tradicionais às contemporâneas. Uma análise comparativa da metodologia adotada por alguns desses projetos demonstra os avanços evidenciados nos estudos

geolinguísticos no Brasil, que, por sua vez, estão em sintonia com as tendências desses estudos nos continentes europeu e americano (AGUILERA, 1994).

Como é sabido, ao lado dos atlas já publicados e dos concluídos (à espera de publicação) estão em curso outros atlas de caráter estadual, que se encontram em diferenciados estágios de elaboração. Apresentamos, por região, os atlas já concluídos e não publicados e também os que se encontram em andamento ou em fase inicial de suas atividades.

(i) O Atlas Linguístico do Amazonas – ALAM/2004

Foi resultado da tese de Doutorado, defendida em 2004 na UFRJ, de Maria Luiza de Carvalho Cruz. Penúltimo atlas regional elaborado no Brasil, o de número 10, o ALAM tem uma rede de pontos com nove localidades distribuídas pelas diferentes microrregiões do Estado, considerando-se as bacias hidrográficas dos principais afluentes do Rio Amazonas. Os informantes, num total de 54, distribuem-se seis por cada ponto, sistematicamente três homens e três mulheres, que se enquadram em três faixas etárias — 18 a 35 anos, 36 a 55 anos e a partir de 56 anos. São todos alfabetizados, tendo no máximo até a 4ªsérie do nível fundamental.

Para a coleta de dados, utilizou-se um questionário elaborado segundo alguns parâmetros do Projeto ALiB, incluindo questões que visam a apurar aspectos regionais específicos, e está assim organizado: Questionário Fonético-Fonológico — 156 perguntas; Questionário Semântico-Lexical — 327 perguntas; e mais a orientação para conversa semidirigida. A partir da análise desse material, foram obtidas 107 cartas fonético-fonológicas e 150 cartas semântico-lexicais.

(ii) O Atlas geo-sociolinguístico do Pará (ALiPA)

O ALISPA, que é parte de um projeto mais amplo, o do *Atlas Geo-Sociolinguístico do Pará*, iniciado em 1996, configura-se como o primeiro atlas sonoro sobre a variedade brasileira da língua portuguesa.

Organizado por Abdelhak Razky, tem uma rede prevista de 57 localidades distribuídas pelas seis mesorregiões do Estado do Pará. O questionário compõe-se de duas partes: uma geral, que contempla as áreas semânticas Terra e Homem, e, ainda, lendas e superstições e narrativa pessoal; e outra parte vinculada a aspectos específicos da região investigada. Os informantes são de ambos os sexos e se distribuem entre a zona rural e a urbana.

Parte desse Projeto foi publicada em formato digital (CD), o Atlas Linguístico Sonoro do Pará (ALiSPA), que tem como meta a construção de uma fonoteca dos falares paraenses, a partir de amostras recolhidas em 10 cidades do Estado (PEREIRA,2007).

(iv) O Projeto de um Atlas Etnolinguístico do Acre (ALAC)

Organizado por Luiza Galvão Lessa Karlberg, o ALAC tem uma rede de pontos constituída por 18 localidades situadas em três áreas de pesquisa: Vale do Acre, Vale do Juruá e Vale do Purus. Foram 1250 inquéritos, dos quais os informantes se distribuem entre os dois sexos e em três faixas etárias (16-25 anos; 26-35 anos e 36-80 anos⁷). Os questionários, com um total de 1205 perguntas, estão organizados em dois tipos, geral e específico, para elaborar 220 Cartas Léxicas.

A coleta de dados já foi concluída e está reunida em 4.025 horas de gravação, com entrevistas entre documentador e locutor, no Questionário Palavra e Coisa e no Questionário Semântico-Lexical, nos três Vales da pesquisa (KARLBERG,2018).

(v) O Atlas Linguístico do Maranhão (ALiMA)

Assessorado por Conceição de Maria de Araújo Ramos, o ALiMA teve início em 2002, está em sua fase de desenvolvimento, adotou a mesma metodologia do ALiB e está assim estruturado: 15 localidades pertencentes às microrregiões do Estado do Maranhão, 64 informantes, nas faixas etárias de 18 a 30 e de 45 a 60 anos, de nível de instrução entre alfabetizados até a 4^a série do 1º grau e nível superior, homens e mulheres.

Na sua metodologia, utiliza-se dos mesmos questionários do ALiB, acrescidos de 50 novas questões sobre as principais manifestações artístico-culturais e dos produtos agrícolas do Estado do Maranhão.

(vi) O Atlas Linguístico do Piauí – ALiPI

Dirigido por Luiz Egito de Souza, o projeto está em fase inicial de elaboração. Adotou a mesma metodologia do Atlas Linguístico do Brasil e organizou-se, metodologicamente, da seguinte forma: 12 localidades pertencentes às microrregiões do Estado do Piauí, 52 informantes, homens e mulheres, nas faixas etárias de 18 a 30 e de 45 a 60 anos, com nível de

⁷ Para os estudos sociolinguísticos, a distribuição da faixa etária de 36-80, no ALAC, toma um espaço de tempo muito longo, o que pode comprometer a análise dos dados. Salienta-se que o estudo do ALAC, de natureza lexical, objetiva fazer análise linguística e apresentar a presença de fronteiras lexicais ou até mesmo de isoglossas na região em destaque, à luz da Geolinguística.

instrução entre a alfabetização, a 4^a série do Ensino Fundamental e nível superior. Os questionários utilizados são os mesmos do ALiB, porém há sugestões de outros que contemplam as principais manifestações artístico-culturais e os produtos agrícolas do Estado do Piauí.

(vii) O Projeto de um Atlas Linguístico do Rio Grande do Norte (ALiRN)

O ALiRN recobre uma rede constituída de dez localidades representativas das diversas regiões do Estado. É coordenado por Maria das Neves Pereira (Universidade Potiguar – UnP) e encontra-se em fase de implementação. Definida a rede de pontos, conta com 11 localidades e já fez aplicação de inquéritos experimentais em 80% da rede de pontos.

(viii) O Projeto Atlas Linguístico do Espírito Santo (APES)

Dirigido por Catarina Vaz Rodrigues, da UFES, está em fase inicial de realização. Na definição da sua rede de pontos foram consideradas as localidades indicadas por Antenor Nascentes (1958). Conta com 30 localidades pertencentes às microrregiões do Estado. Os informantes, homens e mulheres, em número de 60, estão divididos por faixa etária de 30 a 55 anos, com grau de instrução da alfabetização até, no máximo, a 4^a série do Ensino Fundamental.

Para a constituição do *corpus*, o questionário foi estruturado seguindo a metodologia de outros questionários utilizados em atlas estaduais e no ALiB, além de questionários com aspectos específicos do Estado do Espírito Santo. Na composição da equipe do APES, participam: Valéria Cristina Barbosa Gabriel e Valdeciliana da Silva Ramos Andrade, ambas pertencentes a Faculdades de Letras de Vitória.

(ix) O Atlas Fonético do Entorno da Baía de Guanabara (AFEBG)

O AFEBG adota como metodologia um Questionário Fonético-Fonológico, especialmente para esse fim, adaptado à realidade das comunidades investigadas, constituído de 279 questões para entrevistas com 24 informantes de ambos os sexos. O *corpus* constitui-se de 7400 dados dos quais resultaram 308 cartas fonéticas. Esse atlas é o mais recente atlas linguístico elaborado no Brasil. Consiste em um estudo dialetológico de cunho fonético-fonológico, com vistas à descrição da fala de quatro comunidades fluminenses, segundo os princípios atuais da Geolinguística e da Sociolinguística.

É resultado de uma dissertação de Mestrado defendida em 2006 por Luciana Gomes de Lima, sob a orientação da Professora Doutora Sílvia Brandão.

(x) O Projeto Atlas Linguístico Sonoro do Estado do Rio de Janeiro (ALiSon-Rio)

Dirigido por Cláudia de Souza Cunha, da UFRJ, o ALiSon-Rio será divulgado por meio de CD – Rom. As respostas aos questionários serão apresentadas em cartas em que se integrarão texto e som.

Concebido como uma pesquisa que está em fase de implantação, abarcará 24 localidades, representando as oito microrregiões do Estado do Rio de Janeiro. Em cada ponto de inquérito, serão entrevistados quatro informantes com, no máximo, a 4^a série do Ensino Fundamental, um homem e uma mulher, distribuídos por duas faixas etárias: de 18 a 30 anos e de 50 anos em diante.

Salienta-se que os inquéritos têm como base a metodologia do ALiB.

(xi) O Projeto Micro Atlas Fonético do Estado do Rio de Janeiro (MicroAFERJ)

O Micro AFERJ foi elaborado na UFRJ como tese de Doutorado de Fabiana da Silva Campos Almeida, sob a orientação da Professora Sílvia Brandão. O Questionário Fonético-Fonológico, o número de informantes por localidade e os critérios para sua escolha são os mesmos adotados para a elaboração do AFeBG. Foram estipulados 12 pontos de inquérito, que contemplam áreas litorâneas e interioranas: Barra do Itabapoana, Porciúncula, Cantagalo, Cabo Frio, Santa Maria Madalena, Cachoeiras de Macacu, Quissamã, Itaguaí, Três Rios, Valença, Resende e Parati. Todas as recolhas já foram finalizadas e os dados estão em fase de transcrição fonética.

(xii) O Atlas Linguístico do Estado de São Paulo (ALESP)

O pesquisador Pedro Caruso é o precursor desse trabalho. O ALESP teve início nos anos 1980. Na sua metodologia, considerou 60 pontos sugeridos por Nascentes (1958), fazendo, porém, alguns acréscimos e desprezando localidades já não mais pertinentes. Numa primeira fase de constituição desse atlas, foram realizados inquéritos por correspondência em todo o Estado, com utilização de um questionário reduzido de 50 perguntas. Distribuídos 1.382 questionários, o pesquisador responsável recebeu respostas de 853 deles, número que lhe permitiu uma visão preliminar da realidade paulista.

Já na segunda fase, o questionário utilizado contou com 310 perguntas, distribuídas em duas grandes áreas semânticas –Terra e Homem –, havendo uma terceira parte, constituída de

seis questões destinadas a apurar relatos referentes a lendas e superstições, e uma quarta e última parte, destinada a colher relatos de experiência pessoal.

(xiii) O Projeto Atlas Linguístico do Mato Grosso (ALiMAT)

Constitui-se numa pesquisa em fase de implantação. No ALiMAT foram selecionadas 22 localidades para constituição da amostra. Contou-se com 92 informantes, distribuídos 4 no interior e 8 na capital. Adotou-se a metodologia do ALiB para a elaboração dos questionários e estão assim organizados: QSL com 256 perguntas, QFF com 159 e QMS com 124 perguntas, além de um discurso livre com 3 perguntas. Fazem parte da equipe de pesquisadores: José Leonildo Lima, Judith Albuquerque, Antônio Tadeu de Azevedo, Cássia Regina Tomanin, Liliane Batista Barros, Maria José Marques, Valéria Cardoso, Wellington Quintino e Zeneide Souza, com coordenação de Dercir Oliveira.

Percebemos, portanto, que há uma preocupação, de certo modo nacional, com a geografia linguística no Brasil e “afigura-se como um prenúncio da vontade de que é urgente uma descrição acurada da realidade linguística brasileira, para que se alcance o pleno conhecimento do português do Brasil (CARDOSO, 2010, p. 168).

Observemos no Quadro 6 a relação de Atlas que foram concluídos e publicados, por ordem de publicação, e Atlas que ainda se encontram em andamento.

Quadro 6: Atlas que já foram concluídos/publicados e Atlas que estão em andamento

| ATLAS CONCLUÍDOS/PUBLICADOS | ATLAS EM ANDAMENTO |
|---|--|
| Atlas Prévio dos Falares Baianos (1963) | Atlas Linguístico do Espírito Santo |
| Esboço de um Atlas Linguístico de Minas Gerais (1977) | Atlas Linguístico do Mato Grosso |
| Atlas Linguístico da Paraíba (1984) | Atlas Linguístico de Rondônia |
| Atlas Linguístico de Sergipe (1987) | Atlas Linguístico do Pará |
| Atlas Linguístico do Paraná (1994) | Atlas Linguístico do Maranhão |
| Atlas Linguístico-Etnográfico da Região Sul (2002) | Atlas Linguístico do Rio Grande do Norte |
| Atlas Linguístico Sonoro do Estado do Pará (2004) | Atlas Linguístico do Acre |
| Atlas Linguístico de Sergipe II (2005) | Atlas Linguístico de Goiás |
| Atlas Linguístico do Mato Grosso do Sul (2007) | Atlas Linguístico do Piauí. |
| Atlas Linguístico do Estado do Ceará (2010) | Atlas Linguístico de Pernambuco |
| Atlas Linguístico do Amapá (2017) | |

Fonte: Elaborado pela autora.

Do que vimos até aqui, fica claro que os Atlas linguísticos desenvolvidos e os projetados, conforme Quadro 6, em sua grande maioria, seguem, na sua metodologia, os princípios da Dialetologia contemporânea, que tem como objetivo o estudo da variação espacial vinculada à variação social. São, portanto, os mesmos parâmetros metodológicos que seguimos para a efetivação desta pesquisa. Assim, podemos notar o interesse de muitos pesquisadores em descrever e caracterizar a variedade brasileira da língua portuguesa no intuito de mapear a variação linguística através de atlas.

É evidente que os registros dos diferentes falares regionais apontam marcas socioculturais peculiares das diferentes regiões brasileiras. Nessa sequência, as pesquisas geolinguísticas no Brasil podem contribuir para a desconstrução de ideias cristalizadas e equivocadas sobre os falares regionais, bem como pode, também, revelar áreas de isoglossas significativas para o estudo da língua portuguesa no Brasil.

2.3 DIALETOLOGIA E SOCIOLINGUÍSTICA: UM DIÁLOGO POSSÍVEL

Vários ramos do conhecimento humano direcionam-se para o mesmo objeto de estudo, a língua. Porém, mesmo tendo em comum o mesmo material de estudo, tais ramos se diferenciam pela forma como analisam a língua, isto é, “uma disciplina se diferencia das demais que lhes são correlatas muito mais pelo objeto formal do que pelo objeto material” (MONTEIRO, 2000, p. 27).

Por outro lado, sabemos que não é tão fácil diferenciar os campos ou as perspectivas adotadas. Falando da Sociolinguística, tendo seu conceito tão ampliado, houve uma certa invasão de outros domínios e, por esse motivo, se torna muitas vezes trabalhoso distingui-la de disciplinas afins: Sociologia da Linguagem, Etnologia, Dialetologia e outras. Muito embora a Dialetologia seja a disciplina mais próxima da Sociolinguística, a ponto de muitas vezes os campos se entrecruzarem, os linguistas entendem que a Dialetologia ocupa-se de descrever os dialetos, tarefa também incumbida à Sociolinguística, considerada, por esse motivo, como parcialmente herdeira da Dialetologia (MONTEIRO, 2000).

A atribuição dada à Dialetologia é estudar a variação diatópica e diastrática e à Sociolinguística, o trabalho de estudar esses fenômenos no contexto social, tendo em vista suas realizações, ou seja, uma realiza pesquisas de cunho horizontal; a outra, investigações de caráter vertical.

Para melhor explicar a atribuição da Sociolinguística, Brigh (1974) argumenta:

Se tentarmos ser mais exatos, observaremos que a Sociolinguística difere de algumas preocupações anteriores com as relações língua – sociedade pelo fato de que, seguindo novas perspectivas da própria linguística, considera tanto a língua quanto a sociedade como sendo uma estrutura e não uma coleção de itens. [...] A tarefa da Sociolinguística é, portanto, demonstrar a covariância sistemática das variações linguística e social e, talvez, até mesmo demonstrar uma relação causal em uma ou outra direção. (BRIGHT, 1974, p. 17)

Assim, definir a Sociolinguística, exclusivamente como a relação entre língua e sociedade ou como a disciplina responsável pelo estudo dos fenômenos linguísticos relacionados a fatos sociais, torna-se vago, uma vez que não engloba seus objetivos (RIBEIRO, 2012). A correlação da língua com os fatores sociais não pode ser ignorada quando se estudam fenômenos linguísticos, porém só passaram a ser objeto de estudo com maior transparência, a partir do advento da Sociolinguística.

As temáticas da Dialetologia atual se confundem de certa forma com as da Sociolinguística; a separação é muito mais de caráter metodológico, concedendo à Sociolinguística o estudo das variações por intervenção de um tratamento quantitativo, na tentativa de medir o peso de cada grupo de fatores favorecedores ou inibidores da aplicação de uma determinada regra variável. Desse modo, não se pode negar que o conhecimento da realidade linguística brasileira teve início no âmbito da Dialetologia e atingiu o ápice com a Sociolinguística, mais especificamente com a Sociolinguística laboviana e os sofisticados métodos de análise estatística.

Vale destacar que a Dialetologia já examinava os fatores sociais, embora não necessariamente os utilizasse para explicar a variação linguística. Contribuições metodológicas da Sociolinguística, como a utilização de métodos quantitativos de análise, e o estudo da variação diatópica associada à variação diagenérica, diageracional ou diastrática⁸ são exemplos dos elos constituídos entre as disciplinas (RIBEIRO, 2012).

Callou (2010) chama atenção para o livro clássico de Labov sobre a linguagem⁹ de Nova York (1966), editado um ano após a publicação de *Aspects*, de Chomsky (1965), fatos que contribuíram e marcaram, significativamente, os estudos no campo da variação linguística bem

⁸ Desde que Thun (2000) conjugou critérios sociolinguísticos aos estudos geolinguísticos, os atlas adquiriram uma feitura pluridimensional ao inserir fatores como escolaridade, profissão, renda familiar e sexo aos fatores já existentes.

⁹ Pensamos que a língua, por um lado, é provida de uma dimensão imanente, aquela própria do sistema em si mesmo, algo pronto para ser ativado pelos indivíduos, quando preciso. Por outro lado, a língua está vinculada à realidade histórico-social de um povo, é um *sistema em uso*. A linguagem é o suporte, a mediação pela qual tudo passa de um indivíduo a outro, de um grupo a outro, de uma geração a outra. Não podemos reduzir a linguagem à questão de falar certo ou de falar errado (ANTUNES, 2009, p. 21-21).

como aproximaram as duas disciplinas. Labov, representante da Sociolinguística Variacionista atual, era considerado, naquela época, como dialetólogo, motivação que o levou a ser vítima de muitas críticas pelos formalistas, os quais faziam restrições à utilização de dados não-intuitivos em análise linguística. “A segunda edição do livro é de 2006 e nela o autor faz um retrospecto e uma avaliação de quarenta anos de pesquisa Sociolinguística” (CALLOU, 2010, p. 33).

Callou (2010) assim se manifesta sobre a Dialetologia tradicional:

A intersecção dos enfoques dialetológico e sociolinguístico deve-se também ao fato de (i) ter crescido a partir da década de 60, o interesse pela descrição da linguagem urbana, (ii) ter sido necessário pensar em criar nova metodologia para o estudo das línguas românicas [...]. Os métodos da dialetologia tradicional rural sofrem uma adaptação, a fim de possibilitar o levantamento da linguagem de grandes centros urbanos, mais diversificados, em princípio, vindo, consequentemente, a Dialetologia urbana a confundir-se com a Sociolinguística. (CALLOU, 2010, p. 34)

Como já mencionado anteriormente, foi a partir de novas metodologias, inclusive de medida quantitativa de dados linguísticos que foi possível confrontar fatos linguísticos a fatos socioculturais e compreender melhor as diferenças dialetais. A metodologia quantitativa, ao utilizar programas computacionais, permite que o pesquisador tenha acesso a um número infinitamente maior de dados e possa compreender não só a variação estável como também mudanças em curso, através da distribuição de variáveis.

É válido reafirmar que a Sociolinguística define como seu objeto de estudo, a língua falada, em correlação com o contexto social no qual ela é utilizada. Para efetivação de uma pesquisa sociolinguística, é preciso levantar uma hipótese de trabalho, estabelecer um método para a realização do mesmo, selecionar os informantes, fichar os dados, quantificá-los e, por fim, analisá-los.

Segundo Tarallo (1985, p. 20), a pesquisa sociolinguística, de modelo quantitativo, prevê que a representatividade do *corpus* seja “sempre avaliada em função da variável estudada e com base nos objetivos centrais do estudo em questão”. Definindo as hipóteses de trabalho e tendo um *corpus* para análise, o pesquisador passa a avaliar quais os fatores (linguísticos e extralinguísticos) que podem ser examinados para caracterizar o fenômeno linguístico escolhido.

Cardoso (1999) salienta que não se trata da substituição de uma teoria por outra, mas do redimensionamento e da revisão dos novos rumos metodológicos dos estudos dialetais e, para tanto, propõe que

A Geolinguística hoje [...] deve continuar a priorizar a variação diatópica, abrindo espaço para o controle de outras variáveis como sexo, 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. [...] Tomar diferentes tipos de falantes, diversificados socialmente, significa, por outro lado, ampliar o espectro da fotografia. (CARDOSO, 1999, p. 152)

A Sociolinguística nasceu de certa forma dentro da Dialetologia. São, assim, duas perspectivas de estudo da língua que não se opõem, e sim se complementam. Como exemplo de pesquisa que utiliza as duas metodologias temos o Projeto de Estudo da Norma Linguística Urbana Culta (NURC), que surgiu a partir das novas orientações dos estudos dialetólogos. O desenvolvimento do Projeto NURC, inspirado na Dialetologia urbana e na Sociolinguística, teve início no final da década de 1960 e início da década de 1970 (CALLOU, 2010, p. 35).

O Projeto abarcou as seguintes capitais: Salvador, Recife, Rio de Janeiro, São Paulo e Porto Alegre. Tinha como objetivo estudar a norma urbana culta, encontrada em falantes com curso universitário completo. Buscava-se comprovar a pluralidade de normas e responder a algumas questões, a saber: como se fala hoje em determinadas comunidades? que variações acontecem, sejam sociais, regionais, combinatórias ou distribucionais?

Combinando as duas perspectivas, dialetológica e sociolinguística, esse projeto pode juntar o regional ao social, o rural ao urbano e proporcionou a convivência com a diversidade, não só linguística bem como teórico-metodológica.

Concebendo as realidades da Dialetologia e da Sociolinguística e os objetivos da nossa pesquisa, esta tese alinhou os dois enfoques teóricos: a Dialetologia, quando investiga o uso das formas de tratamento *tu* e *você* em cinco Estados do Nordeste, a partir da análise de um extrato do *corpus* do ALiB, cujo objetivo é mapear essas formas pronominais nessas localidades bem como representá-las em cartas linguísticas; a Sociolinguística, quando averiguamos a relevância das variáveis sociais (sexo, idade, escolaridade) e linguísticas (referenciação, função sintática, paralelismo, dentre outras) na variação das formas pronominais *tu* e *você*.

Frente ao exposto, depreende-se que a Dialetologia e a Sociolinguística estão bem próximas, pois utilizam a língua falada como objeto de estudo e observam as variáveis sociais que atuam nos contextos de fala. Por outro lado, apesar da proximidade teórica entre a Sociolinguística e a Dialetologia, as mesmas dão um tratamento particular ao seu objeto de estudo, sendo prioritária para a Dialetologia a variação diatópica ou geográfica, sem, no entanto, desconsiderar os fatores sociais. Já a Sociolinguística ocupa-se, prioritariamente, da correlação

entre fatores linguísticos e sociais, ainda que isso não anule um olhar para o ponto de vista espacial.

Acreditamos que a associação da Dialetologia com a Sociolinguística facilitou o conhecimento da realidade espacial do domínio do português, evidenciando as diferenças e convergências que se registram no território brasileiro. Assim pensamos que a nossa pesquisa “possa ser mais um recurso para discutir com os alunos em sala de aula sobre o tema variação linguística como fenômeno peculiar a toda e qualquer língua bem como fornecer dados para equacionar os problemas decorrentes do preconceito linguístico” (PAIM, 2019, p. 113).

2.4 DA DIALETOLOGIA À SOCIOLINGUÍSTICA: CONTRIBUIÇÕES PARA O ENSINO DE LÍNGUA PORTUGUESA

Sendo um dos objetivos desta tese fornecer aos docentes de Língua Portuguesa subsídios para o aprimoramento do ensino-aprendizagem da língua vernácula, examinaremos de que forma a Dialetologia e a Sociolinguística podem auxiliar com novas estratégias de ensino a partir do reconhecimento e da identificação da realidade linguística de um território.

Cardoso e Mota (2006, p. 87) esboçam algumas ideias da interface entre a Dialetologia e o ensino:

- (i) a Dialetologia como instrumento para reconhecer-se, identificar-se e definir-se a realidade de cada região e, consequentemente, levantar a base linguística do estudante que chega à escola;
- (ii) a Dialetologia como meio de se estabelecerem as relações entre as diversas modalidades de uso de uma língua num território circunscrito e de apreciar a natureza dos usos na perspectiva do comportamento social do falante diante da sua língua.

É fato, portanto, que a Dialetologia propõe oferecer, através dos trabalhos dialetais, inclusive do atlas linguísticos, a visão pluridimensional de fenômenos linguísticos, situando o aluno no seu contexto linguístico, levando-o a entender, sem estigmas, a diversidade que o rodeia.

Cardoso e Mota (2006) afirmam que:

O estudo da variação linguística, quer do ponto de vista social, quer do diatópico, traz para a sala de aula o conhecimento do falar do usuário ali presente, ajudando a determinar as características do seu vernáculo, antes que lhe seja ensinada uma outra modalidade linguística, seja ela literária, ou,

simplesmente, culta; fornece as características dessa forma literária ou culta da língua materna, desfazendo equívocos e eliminando preconceitos. (CARDOSO; MOTA, 2006, p. 103)

Nesse sentido, considera-se necessário e importante que o estudante de Letras realize pesquisas sobre variação linguística ao longo de sua formação, tanto Sociolinguística quanto Geolinguística (CARDOSO; MOTA, 2006). Os professores que conhecem a diversidade linguística estão habilitados a investigar e a descrever as variedades do português em uso nas comunidades de onde seus alunos residem, podendo, portanto, posicionar-se de maneira crítica em relação ao ensino da língua. Dessa maneira, poderão contribuir para o desenvolvimento de uma pedagogia sensível às diferenças sociolinguísticas e culturais dos alunos.

De acordo com Bortoni-Ricardo (2005), pesquisas têm mostrado que professores sensíveis às diferenças sociolinguísticas e culturais desenvolvem intuitivamente estratégias interacionais em sala de aula que são bastante positivas. “Nessas condições, o aluno é ratificado pelo professor e pelos colegas como um falante legítimo e começa a aprender a alternar seu dialeto vernáculo e a língua de prestígio” (BORTONI-RICARDO, 2005, p. 132).

É importante mencionar o que dizem os PCNs de Língua Portuguesa em relação à variação linguística:

A língua portuguesa, no Brasil, possui muitas variedades dialetais. Identificam-se geográfica e socialmente as pessoas pela forma como falam. Mas há muitos preconceitos decorrentes do valor social relativo que é atribuído aos diferentes modos de falar: é muito comum se considerarem as variedades linguísticas de menor prestígio como inferiores ou erradas. (BRASIL, 1997, p. 26)

O documento debate a questão do preconceito linguístico propondo que a escola, enquanto formadora do sujeito-cidadão, seja incentivadora das igualdades sociais e do respeito às diferenças: “O problema do preconceito disseminado na sociedade em relação às falas dialetais deve ser enfrentado, na escola, como parte do objetivo educacional mais amplo de educação para o respeito à diferença” (BRASIL, 1997, p. 26).

As universidades de formação de professores necessitam de projetos que permitam que os futuros professores tenham total segurança em lidar com a heterogeneidade da língua, “cuja variação estrutural no espaço geográfico, na ordem social e em suas funções (da língua) está em grande parte relacionada à história, às crenças e aos valores culturais e ideológicos da comunidade que a fala” (CARDOSO; MOTA, 2006, p. 122).

Outras ações podem ser desenvolvidas pelos estudantes de Letras no intuito de habilitá-lo a lidar com as diferenças linguísticas, a saber: 1) projetar um Banco de Dados informatizado com a descrição das variantes encontradas em determinado município; 2) realizar trabalhos de campo nos cursos de graduação e pós-graduação; 3) efetuar análises do *corpus* gravado nas entrevistas, dentre outras.

Inferimos, portanto, que o professor de língua portuguesa deve conhecer os postulados teóricos da Sociolinguística acerca da linguagem em seu funcionamento social para poder atuar, de maneira eficiente, na orientação da aprendizagem e na formação contínua do aluno-cidadão. E deve também (re)conhecer e trabalhar a diversidade dialetal da língua a partir, por exemplo, do acesso aos dados veiculados nos atlas linguísticos.

2.5 SÍNTESE DO CAPÍTULO

Neste capítulo, apresentamos sobre a Teoria da Variação Linguística, a qual forneceu-nos o aparato metodológico para lidar com a variação *tu/você*, possibilitando-nos saber em que medida essas formas variam e quais fatores condicionam essa variação. Forneceu-nos também o embasamento teórico que nos permitiu selecionar os fatores que atuam sobre o fenômeno pesquisado.

Para o tratamento da diversidade linguística, existem duas áreas de estudos afins: a Dialetologia e a Sociolinguística. A primeira, tem por objeto de estudo os dialetos, considerados como quaisquer variedade de uma língua, e a segunda, ciência que também estuda as diferentes formas de uso de uma língua no seio das comunidades de fala. Lembramos que cada uma delas opera com um método de investigação científico particular.

Como já registramos, a abordagem da língua pode se dar numa dimensão interna e/ou numa dimensão externa. Neste capítulo, comentamos sobre a dimensão externa da variação linguística. Tratamos dos seguintes tipos:

- Variação *diatópica, regional ou geográfica*;
- Variação *social ou diastrática*;
- Variação *estilística ou diafásica*;
- Variação na *fala e na escrita*, também chamada de variação *diamésica*.

Em relação à variação estilística, discorremos sobre as propostas de mudança de estilo segundo Labov e Bell. Essas duas visões sobre o estilo diferem quanto à abordagem metodológica, o que não nos impede, no entanto, de achar um ponto de congruência entre elas:

a necessidade de saber, sobretudo, o que leva um falante a mudar de estilo em uma dada interação. Ou melhor, o que leva um mesmo falante no nosso caso, falantes nordestinos, a mudar de um estilo mais *casual* para um mais *formal*, em uma entrevista, ou vice-versa.

Ficou evidente que a Dialetologia utiliza como método de estudo a Geografia Linguística ou Geolinguística, que consiste na elaboração de atlas linguísticos. Enquanto a Sociolinguística se serve de método científico que opera especificamente com números e tratamento estatístico, voltando sua atenção para um tipo de investigação que confronta aspectos linguísticos e sociais.

Ainda, demonstramos no capítulo em pauta, os atlas concluídos e aqueles que estão em andamento no nosso território, os mesmos revelam a realidade linguística e, juntamente com os estudos sociolinguísticos desenvolvidos no Brasil, constituem hoje um acervo importante para estudiosos da língua portuguesa e diferentes áreas do conhecimento que se dedicam ao homem, seu meio e sua fala.

3 ASPECTOS METODOLÓGICOS

Neste capítulo, evidenciamos os procedimentos metodológicos utilizados na confecção deste trabalho, ou seja, tentamos, de forma clara e detalhada, especificar o *como* da pesquisa, o *que se quer* e *onde acontece*. O capítulo está organizado em oito seções, descritas sumariamente a seguir.

As duas primeiras seções (3.1 e 3.2) são dedicadas ao Projeto ALiB, que forneceu os dados para a pesquisa do fenômeno analisado. Trata-se de informações sobre as etapas, organização do Projeto, histórico, comissão científica, objetivos, metodologia e instrumentos utilizados para a coleta dos dados.

Na terceira seção (3.3), apresentamos as regiões escolhidas como *locus* da nossa pesquisa. Para melhor circunstanciar as áreas investigadas, partimos das propostas de delimitação de espaços no Nordeste, situando as localidades nos cinco Estados analisados, que perfazem um total de trinta pontos. Assim, detalhamos cada ponto, apresentando questões pertinentes às vivências sociais, históricas e culturais, bem como ano de fundação, localização geográfica, nomes anteriores, população e outros dados relevantes.

Na seção 3.4, divulgamos os questionários do Projeto ALiB, delimitando o conjunto de documentos analisados. Salientamos que o Questionário Fonético-Fonológico (QFF) não foi utilizado, uma vez que a nossa pesquisa é de cunho morfossintático.

Na seção 3.5, apresentamos o quadro com os critérios gerais de escolha dos informantes, segundo a metodologia do ALiB, a qual adotamos neste estudo.

Na seção seguinte, 3.6, mostramos o processo de recolha e levantamento de dados, tanto do Projeto ALiB, quanto da pesquisa que deu origem a esta tese.

Na seção 3.7, descrevemos as variáveis dependentes e independentes sob análise e as hipóteses aventadas na pesquisa.

Na seção 3.8, discutimos a metodologia empregada, baseada na Sociolinguística quantitativa, também conhecida como Sociolinguística laboviana. A análise quantitativa dos dados se deu a partir do uso do pacote computacional GoldVarb X, que é constituído por um conjunto de programas para análise estatística de dados linguísticos em variação. Nessa mesma seção, mostramos como foi feita a codificação das ocorrências e como procedemos à análise dos dados.

Na seção 3.9, abordamos o trabalho geolinguístico em relação à confecção de cartas: é necessário saber o que representar, quais variantes são válidas, qual extensão de legenda, a natureza da carta (léxica, fonética, morfossintática, isoglóssica) e o tipo de representação, entre outros fatores.

Por fim, fazemos uma síntese retomando alguns aspectos que foram mencionados neste capítulo.

3.1 PROJETO ALiB: ANTECEDENTES HISTÓRICOS

Em 1952, quando se mostrou o desejo de elaboração de um atlas linguístico brasileiro, a Europa já havia produzido alguns atlas e o Brasil evidenciava condições precárias. Boa parte da população, cerca de 63%, residia no campo e somente 37% nas grandes cidades. A utilização dos meios de comunicação, rádio, televisão, telefone, era bem reduzida. Entretanto, a partir da segunda metade do século passado, a nova realidade econômica e social do país resultou na mobilidade social e no aumento da população nos centros urbanos. E os avanços tecnológicos facilitaram a interação entre as pessoas.

No entanto é preciso ressaltar que a motivação para a criação de um atlas linguístico do Brasil teve seu começo na primeira metade do século XX. Desde então, filólogos e linguistas buscaram sensibilizar o governo brasileiro a se manifestar favorável à realização de um atlas linguístico do Brasil. Contudo, só na década de 1950, por meio do Decreto nº 30.643, de 20 de março de 1952, é que a proposta se efetiva no campo jurídico:

A Comissão de Filologia promoverá pesquisas em todo o vasto campo de filologia portuguesa – fonológicas, morfológicas, sintáticas, léxicas, etimológicas, métricas, onomasiológicas, dialetológicas, bibliográficas, históricas, literárias, problemas de texto, de fontes, de Autoria, de influências, sendo sua finalidade principal a elaboração do Atlas Linguístico do Brasil. (BRASIL, 1952a, apud CARDOSO, 2014, p. 17)

A regulamentação deste Decreto se dá através da Portaria nº 536, de 26 de maio do mesmo ano, do Ministério da Educação e Saúde (BRASIL, 1952). Dentre várias pesquisas a serem planejadas, estabeleceu-se como meta principal a elaboração do atlas linguístico do Brasil. Dessa maneira, Silva Neto (1957) define as tarefas que considerava urgentes para a concretização dos estudos dialetais no Brasil, instituindo o *Guia para estudos dialectológicos* (1957), entre as quais se lê no inciso f) “inquérito nacional feito com questionário uniforme e distribuição geográfica dos materiais recolhidos, em mapas” (CARDOSO et al., 2014a). Outros pesquisadores, a exemplo de Celso Cunha e Antenor Nascentes, acolheram a ideia. Na ocasião do *III Colóquio Internacional de Estudos Luso-Brasileiros*, que aconteceu em Lisboa, no ano de 1957, os estudiosos salientaram sobre a necessidade de elaboração de atlas linguísticos regionais.

Ainda assim, inúmeras foram as dificuldades para a realização dessa empreitada, a saber: falta de investimento financeiro, inexistência de pesquisadores preparados para esse tipo de pesquisa, estradas de difícil circulação que dificultavam o acesso à investigação, dentre outras, que contribuíram para a não realização de um atlas nacional.

A obra de Antenor Nascentes, *Bases para elaboração do atlas linguístico do Brasil*, publicada em dois volumes (1958), auxilia, fundamentalmente, os trabalhos de pesquisa nessa área. Logo no primeiro capítulo, o autor deixa claras as vantagens de um atlas feito ao mesmo tempo para todo o país. Por outro lado, admite a impossibilidade de sua concretização:

Embora seja de toda vantagem um atlas feito ao mesmo tempo para todo o país, para que o fim não fique muito distanciado do princípio, os Estados Unidos, país vasto e rico e com excelentes estradas, entregou-se à elaboração de atlas regionais, para mais tarde juntá-los no atlas geral. Assim também devemos fazer em nosso país, que é também vasto e, ainda, mais pobre e sem fáceis vias de comunicação. (NASCENTES, 1958, p. 7)

O trabalho de Nascentes (1958) apresenta os instrumentos elementares para a efetivação da pesquisa de campo: questionário organizado em áreas semânticas (caráter provisório) e rede de pontos selecionada em diferentes Estados. Desse modo, é importante lembrar que poderá

estar “sujeita ao parecer do técnico encarregado de dirigir os trabalhos em cada região, na qualidade de melhores conhecedores do terreno” (NASCENTES, 1958, p. 19). Numa perspectiva evolutiva, a linguagem apresenta-se de forma dinâmica, acompanhando as mudanças na história do país, o que constitui razão maior para justificar a importância de descrever a realidade linguística brasileira.

3.2 CONSTRUÇÃO DO PROJETO ATLAS LINGUÍSTICO DO BRASIL

A nova configuração da realidade brasileira aspirava à construção de um atlas linguístico geral do Brasil no tocante à língua portuguesa. Passados cerca de cinquenta anos, a ideia foi retomada: (i) pela urgência de descrever o português brasileiro, antes que se percam dados e fatos capazes de esclarecer aspectos da história linguística do país; e (ii) pelo papel que a Geografia Linguística continuava a ter, apesar da implementação dos estudos sociolinguísticos.

Graças à iniciativa de um grupo de pesquisadores brasileiros da área, que contou com a presença do Prof. Dr. Michel Contini, do Centre de Dialectologie de Grenoble, diretor do *Atlas linguistique roman* e membro do Comitê Diretor do *Atlas linguarum Europae*, que se reuniram no Instituto de Letras da Universidade Federal da Bahia, foram consolidadas as bases para a concretização do Projeto Atlas Linguístico do Brasil (ALiB).

As bases para a elaboração desse projeto foram lançadas por ocasião do Seminário “Caminhos e Perspectivas para a Geolinguística”, no ano de 1996, na Bahia. Inicialmente, o Projeto envolveu dezessete universidades brasileiras e, a partir daí, formou-se o Comitê Nacional¹⁰, que se encarregaria da estruturação do projeto e da implantação da pesquisa, formado por representantes do atlas até então publicados e de um representante dos atlas em andamento. O Comitê instituiu os objetivos do Projeto ALiB e considerou a documentação de 250 localidades distribuídas por todo o território nacional e representativas das diversas regiões. O projeto visa à “descrição acurada da realidade linguística brasileira, para que se alcance o pleno conhecimento do português do Brasil” (CARDOSO, 2010, p. 168).

A seguir, seguem outros objetivos do Projeto ALiB:

10 Constitui-se esse Comitê, na sua segunda composição, dos seguintes integrantes: Jacyra Andrade Mota (Diretora Presidente); Silvana Soares Costa Ribeiro (Diretora Executiva) e alguns Diretores Científicos: Abdelhak Razky, Aparecida Negri Isquierdo, Conceição Maria de Araújo Ramos, Fabiane Cristina Altino, Felício Wessling Margotti, Marcela Moura Torres Paim, Maria do Socorro Silva de Aragão, Marilúcia Barros de Oliveira, Regiane Coelho Reis, Valter Pereira Romano e Vanderci de Andrade Aguilera.

- a) Descrever a realidade linguística do português do Brasil com vistas a identificar fenômenos fonéticos, morfossintáticos, lexicais, semânticos e prosódicos característicos da diferenciação ou definidores da unidade linguística no território nacional.
- b) Estabelecer isoglossas com vistas a traçar a divisão dialetal do Brasil, tornando evidentes as diferenças regionais através de resultados cartografados em mapas linguísticos [...].
- c) Registrar, com base na análise em tempo aparente, processos de mudança.
- d) Identificar fenômenos linguísticos localizados e específicos de áreas com vistas a estudar as suas repercussões no ensino-aprendizagem da língua materna.
- e) Examinar os dados coletados na perspectiva de sua interface com outros ramos do conhecimento – história, sociologia, antropologia –, com vistas a fundamentar e definir posições teóricas sobre a natureza da implantação e desenvolvimento da língua portuguesa no Brasil.
- f) 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. (CARDOSO, 1989, p. 171-172)

O Projeto ALiB impulsionou, significativamente, a produção de atlas linguísticos observada no país nos dias atuais, uns já elaborados e outros em andamento (ver Anexo A). Germinada a ideia e estabelecidos os objetivos, chegava a hora de convocar os interessados em pôr em prática o projeto. A esse respeito, Cardoso (2012) afirma que

Era preciso ter a presença de alguém de muita experiência na direção de atlas atuais, mas era preciso também que fosse alguém provido de entusiasmo pelo trabalho no campo da Geolinguística e, sobretudo, capaz de transmiti-lo a um público de dialetólogos diversificado em idade, em experiência e em motivação. (CARDOSO, 2012, p. 16-17)

O ALiB propõe um atlas voltado à pluridimensionalidade, adotando fundamentos da Geolinguística contemporânea. O desenvolvimento do Projeto ALiB configura-se, assim, numa grande contribuição para o entendimento dos diferentes usos da língua, “eliminando visões distorcidas que privilegiam uma variante tida como culta e estigmatizam as demais variantes, causando, desse modo, ao ensino-aprendizagem da língua materna consideráveis prejuízos” (CARDOSO, 2010, p. 169).

Tendo em vista a grande dimensão territorial a ser coberta pela investigação, o Comitê Nacional organizou-se de modo que cada Diretor Científico tivesse sob a sua responsabilidade uma região específica, ficando sob seu controle o planejamento e a execução da pesquisa, assim como a ordenação da equipe regional de pesquisadores¹¹.

11 Para o desenvolvimento das atividades do Comitê Nacional, realizam-se reuniões em diferentes locais do País, o que permite avaliações do desenvolvimento dos trabalhos e definem-se etapas a serem cumpridas. Os Workshop-

Cardoso (2012, p. 20) apresenta um quadro com a distribuição dos responsáveis pela execução da pesquisa de campo nos Estados, conforme o que ficou decidido na XXXV Reunião do Comitê Nacional, ocorrida em julho de 2012, em Salvador:

Quadro 7: Regionais do Projeto ALiB

| REGIONAL | COMPOSIÇÃO | DIRETOR CIENTÍFICO RESPONSÁVEL | COLABORAÇÃO A OUTROS ESTADOS |
|--------------------|---|---|---|
| PARÁ | Pará | Abdelhak Razky | Amapá |
| CEARÁ | Paraíba, Rio Grande do Norte, Piauí, Ceará, Maranhão | Maria do Socorro Silva de Aragão | Amazonas |
| BAHIA | Pernambuco, Alagoas, Sergipe, Bahia, Rio de Janeiro | Jacyra Andrade Mota | Piauí, Minas Gerais, Espírito Santo |
| MATO GROSSO DO SUL | Rondônia, Roraima, Acre, Mato Grosso, Mato Grosso do Sul, Goiás | Aparecida Negri Isquierdo | Amazonas, Pará, Espírito Santo, Rio de Janeiro, São Paulo |
| MINAS GERAIS | Minas Gerais | Ana Paula Rocha, Mário Roberto Zágari | - |
| PARANÁ | Paraná, São Paulo, Amapá | Vanderci de Andrade Aguilera | Amazonas, Roraima, Mato Grosso, Goiás |
| RIO GRANDE DO SUL | Santa Catarina, Rio Grande do Sul | Cléo Altenhofen, Felício Wessling Margotti, Valter Koch | Minas Gerais, Espírito Santo, Rio de Janeiro |

Fonte: Cardoso (2012, p. 20).

Em 2014, os volumes 1 e 2 do Atlas Linguístico do Brasil foram publicados como frutos do trabalho desenvolvido pelo Projeto ALiB, na sua íntegra “um projeto linguístico porque busca documentar, descrever e interpretar a realidade do português brasileiro, tendo, exatamente por esse caráter, uma evidente interface com diferentes ramos do conhecimento organizado”, sucedendo, dessa maneira, a existência de uma relação intrínseca e inseparável entre língua e sociedade (OLIVEIRA; PAIM; RIBEIRO, 2018, p. 215).

ALiB e os Seminários, realizados ao longo da sua implementação, têm definido novas estratégias e uma política de trabalho para a consumação de “um compromisso acadêmico que se planejou” (CARDOSO, 2006, p. 34), há meio século, para com a nossa língua portuguesa no Brasil.

O volume 1 (Introdução) evidencia explicações a respeito da trajetória de construção do Atlas Linguístico do Brasil, descrevendo a metodologia utilizada, com destaque para a rede de pontos, os questionários e os informantes, aos quais se junta a informação sobre a cartografia dos dados.

O volume 2 (Cartas linguísticas 1), com 368 páginas, apresenta resultados das 25 capitais brasileiras, exceto Palmas e Brasília, que, por critérios metodológicos, não integram a rede de pontos. As cartas inclusas nesse volume fornecem resultados referentes a dados fonéticos, morfossintáticos e semântico-lexicais que retratam dos diferentes usos da fala nas capitais dos Estados. Encontramos também no volume 2 o nome de todos os inquiridores e auxiliares de pesquisa, sinalizando a participação de cada um no que tange à aplicação do questionário nos pontos do ALiB.

A publicação dos dois volumes do ALiB foi feita pela Eduel, editora da Universidade Estadual de Londrina. Além desses volumes, pode-se contar com uma série de publicações denominadas *Documentos* (ver Anexo B), volumes que apresentam relatórios, artigos, análises preliminares, memórias e avaliações. Há ainda as publicações dos alunos de graduação em CDs (ver Anexo C), coleção intitulada *Descrevendo a língua, formando jovens pesquisadores*, além de diversas dissertações, monografias e teses concluídas que utilizaram o *corpus* do ALiB.

A publicação do Atlas Linguístico do Brasil vem corroborar, de forma somatória, os estudos dialetais brasileiros, apresentando dados sobre o português brasileiro e suas diversidades de uso.

Segundo Paim (2016),

Os volumes 1 e 2 do atlas fornecem 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. Assim, os atlas permitem a atualização de dicionários da língua portuguesa bem como auxiliam a construção de uma gramática voltada para a realidade do português brasileiro. (PAIM, 2016, p. 78)

Isto posto, o Atlas Linguístico do Brasil valida o trabalho de pesquisadores da língua portuguesa e de áreas afins, no sentido de contribuir para o aprimoramento do ensino-aprendizagem da língua.

O trabalho com os atlas linguísticos na sala de aula propiciará um melhor equacionamento do ensino-aprendizagem à realidade de cada região, uma vez que, descritas as peculiaridades de cada área e caracterizada a variedade de uso da língua ali dominante, pode-se construir um modelo de ensino do vernáculo mais eficaz com o uso dos dados linguísticos que venham a possibilitar a adequação de material didático à realidade linguística de cada região e o entendimento do caráter multidiáletal do Brasil. (PAIM, 2019, p. 114)

Enfim, o trabalho com variação linguística em sala de aula ajuda no entendimento em relação à convivência com as diferenças bem como auxilia na compreensão da variação linguística como peculiar à língua, desmistificando, portanto, a discriminação linguística e os juízos de valores acerca do que é “feio” ou “bonito” na fala.

Algumas ações são desenvolvidas graças à interinstitucionalidade do Projeto ALiB. Por exemplo, o **Congresso Internacional de Dialetologia e Sociolinguística – CIDS**, congresso pensado e organizado pelos pesquisadores do ALiB, contempla interesses das áreas da Sociolinguística e da Dialetologia. O evento acontece a cada dois anos e homenageia pesquisadores dessas esferas. Outra ação resultante dessa interinstitucionalidade foi o convênio Capes – COFECUB, parceria da UFBA com a Université Paris XIII. Como consequência desse convênio, podemos usufruir de dois outros Projetos:

- a) *Dicionário Dialetal Brasileiro* com dados do Projeto ALiB, desenvolvido pelo professor Américo Venâncio Machado Filho;
- b) *Projeto Valextra*, coordenado pela professora Marcela Paim, que estuda os fraseologismos no *corpus* do ALiB (MOTA, 2020).

Outra contribuição de grande valia para o Projeto ALiB veio da parceria com a ciência da computação, que evidencia o seu caráter interdisciplinar. Elencamos algumas ações:

- (i) automatização da geração das cartas linguísticas;
 - (ii) desenvolvimento do banco de dados;
 - (iii) informatização dos inquéritos;
 - (iv) criação de mapas sonoros;
 - (v) melhoria no acesso de informações do ALiB para pesquisadores e para a sociedade;
 - (vi) extração de informações para auxílio nas tomadas de decisões inteligentes.
- (CLARO; OLIVEIRA, 2020)

Percebemos, portanto, que, mesmo com muitas dificuldades enfrentadas pelo Projeto, desde 1996, tem-se um caminho sólido percorrido no decorrer desses anos. Para a efetivação

desse projeto, já fora alcançado o marco dos dois volumes publicados e prevê-se o lançamento dos demais volumes.

3.3 O *CORPUS*

O *corpus* analisado nesta tese é composto pelas respostas dadas aos seguintes questionários: Questionário Semântico-Lexical (QSL); Questionário Morfossintático (QMS), Questões de Pragmática (QP); Temas para Discurso Semidirigido (TDS); Perguntas Metalingüísticas (PM). Assim, para a constituição do *corpus*, partimos da leitura e do exame das transcrições de todas as questões dos referidos questionários, complementadas pela audição dos registros dos inquéritos linguísticos considerados.

3.3.1 Divisão territorial do Nordeste

Para a constituição da rede de pontos do Projeto ALiB, foram levados em consideração alguns critérios, como: aspectos demográficos de cada região e Estado do Brasil; aspectos culturais e históricos; extensão de cada Estado/região; e natureza do povoamento e desenvolvimento socioeconômico de cada localidade. “Para tanto, contou-se com o assessoramento de antropólogos, historiadores e geógrafos e com a colaboração de indigenistas” (ISQUERDO; TELES, 2014, p. 39). A rede de pontos do ALiB conta com um total de 250 localidades distribuídas por todo o país, perfazendo as cinco regiões: Norte, Nordeste, Centro-Oeste, Sudeste e Sul.

Os atlas linguísticos tradicionais, de características rurais, não incluíam os grandes centros urbanos na rede de pontos porque almejava-se o caráter arcaico da linguagem. Por outro lado, os estudos geolinguísticos contemporâneos incluem grandes centros urbanos na rede de pontos, em virtude do reconhecimento das novas configurações da sociedade moderna e da necessidade de incluir tais aspectos às pesquisas dialetológicas (ISQUERDO; TELES, 2014).

Ressalta-se que Palmas, capital do recém-criado Estado de Tocantins, cidade ainda em formação, ou seja, de recente criação (1989) e que não tem, por consequência, informantes nativos filhos de pais nativos, e Brasília, que, inaugurada em 1960, vive ainda uma miscigenação de falares, não foram incluídas no Projeto ALiB. Sobre Brasília, alude-se que a capital recebeu imigrantes de todas as partes e, por conta disso, do ponto de vista linguístico, não apresenta ainda uma norma linguística própria, ou seja, não há um nivelamento linguístico.

Foram consideradas as indicações do dialetólogo Antenor Nascentes para a escolha das localidades através da obra *Bases para a elaboração do Atlas Linguístico do Brasil* (1958), tendo sido mantidas as localidades que ainda se mostravam pertinentes para os objetivos do Projeto. Em sua obra, o autor sugere 605 localidades para a formação do atlas nacional, porém admite a dificuldade de realização de uma pesquisa de tamanha amplitude, motivo pelo qual se torna mais viável a produção de vários atlas regionais, de menor configuração, antes da criação de um atlas nacional, orientação seguida pela dialetologia brasileira até antes do ALiB.

Para a definição da rede de pontos, as informações concedidas pelo IBGE (Instituto Brasileiro de Geografia e Estatística) foram de suma importância, aliadas à “ampla leitura da história do Brasil, dos primeiros viajantes e suas narrativas até teses recentes sobre desbravamento e povoação do país” (ISQUERDO; TELES, 2014, p. 77).

O Projeto ALiB recobre toda a área do território brasileiro e, para esta pesquisa, elegemos como foco de análise a Região Nordeste do país, com o objetivo de mapear cinco Estados (apresentados posteriormente) da referida região com relação aos usos de pronomes *tu* e *você*.

De acordo com Isquierdo e Teles (2014),

A rede de pontos tem, pois, a finalidade de assegurar a representatividade da documentação da variação espacial da língua, a comparação posterior dos dados e a sua respectiva distribuição num determinado espaço geográfico, por meio de cartas linguísticas. (ISQUERDO; TELES, 2014, p. 37)

Acerca do Nordeste, afirma-se que a língua portuguesa aportou primeiro na referida região através dos donatários das capitâncias, quando a antiga Terra de Vera Cruz era apenas uma faixa estreita, limitada pelo Tratado de Tordesilhas. Em meio à catequese jesuítica, eles preferiam usar o latim a aprender a língua geral falada na costa – o tupi – para melhor divulgar a fé cristã.

Segundo Carvalho (2000, p. 22), o início do processo de colonização do Brasil se deu justamente entre os Estados de Pernambuco e Bahia, enquanto outras partes do país só viriam a receber a influência lusitana bem mais tarde. A autora justifica:

[...] uma vez que a língua e a cultura são indissociáveis, no Nordeste, encontramos nessa região uma cultura rica em termos, ritmos e expressão plástica, com um traço eminentemente popular, que não se aprende na escola, nem é valorizado em época de globalização. (CARVALHO, 2000, p. 22)

Até os dias atuais, a região carrega em sua identidade características peculiares desse processo de culturação/aculturação presenciado nos costumes e linguagem da população.

Destaca-se, nessa região, a cidade de Salvador por ter sido capital do Brasil até 1808, quando houve a mudança da sede para o Rio de Janeiro. Depois da transferência da capital para o Rio de Janeiro e da chegada da família real para o Brasil, todo o crescimento foi levado para a grande metrópole, dando origem, portanto, a dois mundos: um rural, sinônimo de atraso; o outro, urbano, sinônimo de moderno e próspero.

Faz-se necessário mencionar a história social e diferentes ciclos econômicos da região Nordeste, uma vez que foram fatores determinantes para o povoamento da região. O ciclo do pau-brasil (1500-1530) facilitou o comércio entre portugueses e indígenas, caracterizado por troca de bugigangas pelo trabalho do indígena no corte e carregamento de madeira até os portos. A cana-de-açúcar, com seus produtos derivados, também foi bastante aceita no mercado europeu. O cacau, planta natural da região amazônica, introduzido no sertão nordestino no final do século XIX, configurou-se como um grande ciclo econômico brasileiro. Nessa perspectiva, percebemos que alguns ciclos econômicos foram específicos de determinadas regiões em detrimento de outras (MUNIZ, 2019).

A rede de pontos do Nordeste se delineou conforme mostra a Figura 4.

Figura 4: Distribuição dos pontos da rede do Nordeste



Fonte: <https://alib.ufba.br/content/rede-de-pontos>.

No que concerne à distribuição espacial dos pontos do ALiB, é sabido que se buscou evitar a proximidade entre eles, mantendo-se um distanciamento que garantisse a representatividade das diferentes localidades escolhidas. A Região Nordeste apresenta 18% da área territorial do Brasil e 27,8 % da população nacional, sendo contemplada com 78 pontos da rede de pontos do Projeto. A Figura 1, Carta VII, mostra as 78 localidades que pertencem à rede de pontos da região Nordeste.

Os cinco Estados que forneceram os dados para a nossa análise estão localizados ao longo do território nordestino, como mostra o mapa da Figura 2. A seguir, são apresentadas informações mais detalhadas dessas localidades, obtidas através de pesquisas em páginas variadas na internet, a saber: *site* do Instituto Brasileiro de Geografia e Estatísticas (IBGE), páginas oficiais das prefeituras de algumas cidades, *site* do Wikipédia, *blogs* atualizados por historiadores locais, páginas referentes ao turismo, dentre outras.

3.3.2 Localidades que perfazem a amostra

A amostra constituída para esta tese foi composta por 30 (trinta) pontos pertencentes a cinco Estados nordestinos, como mostra o Quadro 8:

Quadro 8: Localidades da Amostra

| ESTADOS | PONTOS | TOTAL |
|---------------------|--|-------|
| Rio Grande do Norte | Mossoró, Angicos, Natal, Pau dos Ferros, Caicó | 5 |
| Paraíba | Cuité, Cajazeiras, Itaporanga, Patos, Campina Grande, João Pessoa | 6 |
| Pernambuco | Exu, Salgueiro, Limoeiro, Olinda, Afrânio, Cabrobó, Arcoverde, Caruaru, Recife, Floresta, Garanhuns, Petrolina | 12 |
| Alagoas | União dos Palmares, Santana do Ipanema, Arapiraca, Maceió | 4 |
| Sergipe | Propriá, Aracaju, Estância | 3 |
| | | 30 |

Fonte: Elaborado pela autora.

Para cada localidade investigada, apresentamos informações pertinentes às vivências sociais, históricas e culturais bem como o ano de fundação, localização geográfica, nomes anteriores, população, aspectos demográficos e outros dados relevantes, uma vez que é necessário conhecer a história da população para, então, saber sobre sua língua.

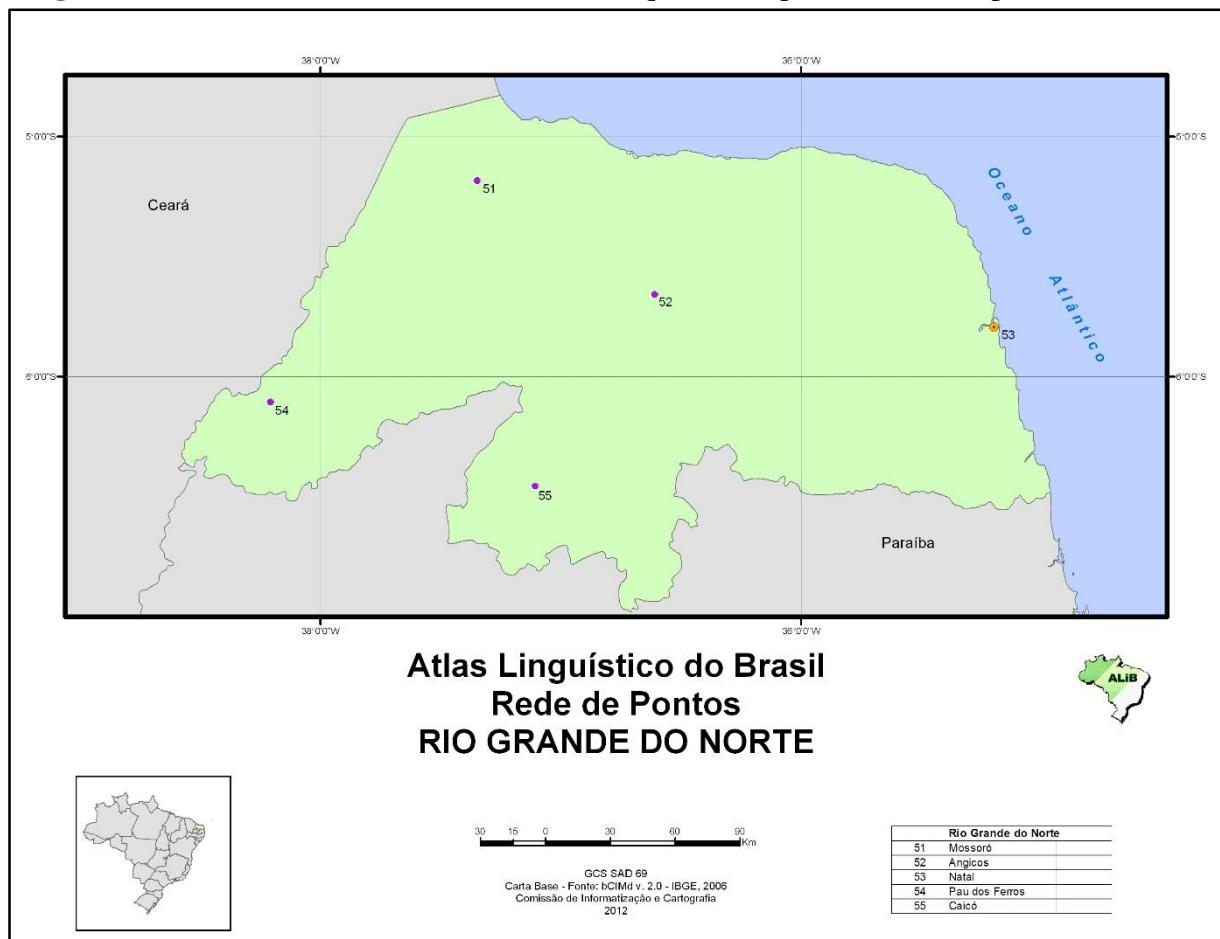
As cidades são apresentadas a seguir, conforme a ordem da rede de pontos do ALiB.

3.3.2.1 Rio Grande do Norte

O Estado do Rio Grande do Norte possui 167 municípios, em uma área de 52.811.047 km². Com, aproximadamente, 3.375.959 habitantes, apenas sete cidades concentram quase a metade da população geral do Estado, a saber: Natal, Mossoró, Parnamirim, São Gonçalo do Amarante, Ceará-Mirim, Macaíba e Caicó.

No Projeto ALiB, há um total de cinco localidades na rede de pontos pertencentes a esse Estado, incluindo a capital, como mostra a Figura 5.

Figura 5: Localidades do Rio Grande do Norte que fazem parte da rede de pontos do ALiB



Fonte: <https://alib.ufba.br/content/rede-de-pontos>.

3.3.2.1.1 Mossoró (Ponto 51 do ALiB)

O município de Mossoró está situado no oeste potiguar, entre duas capitais, Natal e Fortaleza, que são ligadas pela BR-304. Mossoró ocupa uma área de aproximadamente 2.100 km², sendo o maior município do Estado em área, estando distante 281 quilômetros da capital do Estado, Natal. Em 2017, sua população foi estimada pelo IBGE em 295.619 habitantes, posicionando-se como o segundo município mais populoso do Rio Grande do Norte (ficando atrás somente da capital).

De acordo com o site Wikipédia¹², há controvérsias em relação à origem do topônimo “Mossoró”. Dizem que o nome provém do termo *monxoró*, denominação atribuída aos

¹² Algumas informações sobre as cidades nordestinas, aqui selecionadas, foram consultadas no Wikipédia, site polêmico, pois, mesmo havendo um controle editorial, promulgam que os conteúdos são demasiadamente abertos. Utilizamos também o site do IBGE, sites oficiais das prefeituras de algumas cidades, sites oficiais de agências turísticas, dentre outros.

primeiros indígenas que habitavam a região; outros dizem que o nome advém de *mororó*, que significa árvore resistente e flexível.

Mossoró é uma das principais cidades do interior nordestino e atualmente vive um intenso crescimento econômico e de infraestrutura, sendo considerada uma das cidades brasileiras de médio porte mais atraentes para investimentos no país. O município é o maior produtor em terra de petróleo no país como também de sal marinho.

Segundo informações do IBGE, a fruticultura irrigada, voltada em grande parte para a exportação, também possui relevância na economia do município, tendo um dos maiores PIB *per capita* da região. As festividades realizadas na cidade anualmente atraem uma enorme quantidade de turistas, como o Mossoró Cidade Junina, um dos maiores “arraiás” do Brasil, e o Auto da Liberdade, o maior espetáculo brasileiro em palco ao ar livre.

Afirma-se que as primeiras manifestações de povoamento foram por volta de 1600. O município possui todo o seu território situado na Bacia Hidrográfica do Rio Apodi/Mossoró. Os principais rios que cortam o município são o Apodi/Mossoró e o do Carmo. Os principais riachos são o Bonsucesso, o do Cabelo Negro, o de São Raimundo e o do Pai Antônio.

Sobre a vegetação do município, a mesma é formada pela caatinga, apresentando plantas de baixo porte adaptadas à seca, como o faveleiro¹³ e a jurema-preta, além da vegetação com plantas diversificadas, como a carnaúba e a palmeira como espécies predominantes. A cidade de Mossoró, junto com Baraúna, abriga o Parque Nacional de Furna Feia, área de preservação ambiental (*site* Wikipédia).

3.3.2.1.2 Angicos (Ponto 52 do ALiB)

O município de Angicos está localizado na zona do sertão, centro-norte do Estado. Está posicionado à margem esquerda do rio Pataxó ou Angicos, distante a 156 km da capital estadual. Seu clima é seco, ameno e salubre. O município teve essa denominação por decorrência de uma árvore de grande aspecto muito comum no Nordeste.

Conforme o *site* IBGE, os primeiros habitantes da região foram os índios pataxós, pertencentes à nação Gê ou Tapuia. Vários municípios do Rio Grande do Norte originaram-se da fixação de algumas famílias nas caatingas, onde estabeleciam fazendas de criação de gado. Assim aconteceu com Angicos.

13 O mesmo que mandioca-brava.

Segundo o *site* oficial da cidade, *Angicos*, as primeiras moradias na região foram estabelecidas em 1760 e seu fundador foi o tenente Antônio Lopes Viegas, descendente de uma família de nome Dias Machado. É citado que, na época em que foi criada a Vila Nova da Princesa – hoje cidade do Açu – abrangendo os municípios de Açu, Angicos, Macau e Santana do Matos, já se encontravam no território de Angicos várias fazendas.

De acordo com o IDEMA¹⁴, o solo da região é do tipo *litólicos-eutróficos*. A maior parte da área tem aptidão regular e restrita para pastagem natural. Uma pequena área ao norte tem aceitabilidade para lavouras e para culturas de ciclo longo como algodão arbóreo, sisal, caju e coco (*site* IBGE).

Angicos ficou bastante conhecida devido à implantação do projeto de alfabetização, experiência pioneira de Paulo Freire que foi intitulada como “Quarenta horas de Angicos”. Em 2009, criou-se a Universidade Federal Rural do Semiárido (UFERSA), funcionando provisoriamente em locais improvisados. Só em 2011, começou a funcionar em sua sede própria. Angicos passou então a ser conhecida como Cidade Universitária. É conhecida também como “Terra do Pelo” e “Coração do Rio Grande do Norte”.

3.3.2.1.3 Natal (Ponto 53 do ALiB)

Natal é também conhecida como *Cidade do Sol*, devido ao fato de o sol brilhar durante o ano todo; somente entre os meses de março e julho é que o sol é mais ameno, quando ocorrem as chuvas. Afirma-se que a cidade de Natal possui o ar mais puro da América do Sul. A cidade está localizada à margem direita do Rio Potengi, próximo à foz, no Oceano Atlântico. O solo é arenoso, motivação para o aparecimento de dunas. A costa, entrecortada por baías, é protegida por recifes que aparecem ao longo de boa parte do litoral.

Lê-se no *site* Wikipédia que em dezembro de 1597 uma frota comandada por Jerônimo de Albuquerque desembarcou no rio Potengi com a missão de fundar uma nova cidade e construir um forte para proteger os colonos portugueses das incursões dos piratas franceses, que estavam tentando negociar com os potiguares, índios nativos. Um ano depois, esse mesmo grupo começou a construir o Forte dos Três Reis Magos. Assim, no dia 25 de dezembro de 1599, foi fundada uma pequena vila a pouco mais de 2 km de distância do Forte. Essa vila foi batizada de Natal, em homenagem à data de fundação.

14 Instituto de Desenvolvimento Sustentável e Meio Ambiente do Rio Grande do Norte.

O desenvolvimento da cidade, em seus primórdios, foi lento devido à arenosidade do solo e, portanto, não apropriado para o cultivo de cana-de-açúcar. Em 1633, os holandeses invadiram Natal; o Forte foi renomeado para Fort Keulen e assim ficou até 1654, quando os portugueses o reconquistaram. Salienta-se que os holandeses, bem como os portugueses, não tiveram muito entusiasmo pelo desenvolvimento da cidade de Natal, preferiram investir nas cidades de Recife e Olinda.

Durante o Brasil Colônia, após o ciclo da cana-de-açúcar, iniciou-se o ciclo do ouro, também não trazendo muitos benefícios para a cidade de Natal, capital do Rio Grande do Norte. A região não sentiu, como outras, a situação de ser colônia sujeita a uma metrópole; este fato provavelmente contribuiu para fazer de Natal um dos municípios mais liberais do Brasil. O Estado foi o primeiro a abolir a escravidão, dez anos antes dos outros Estados do país.

No início do século XIX, Natal era ainda uma pequena vila, dividida em duas áreas principais: a Cidade Baixa, perto do cais do porto, que concentrava o comércio; e a Cidade Alta, no topo dos morros ao redor do porto, onde se localizavam igrejas e prédios do governo.

Segundo informações retiradas do *site* Wikipédia, em 1922, Natal iniciou um acelerado processo de urbanização. A cidade não cresceu em demasia, houve um planejamento. Os principais produtos do Estado foram xelita, minério de tungstênio, algodão e cera de carnaúba. Outros produtos influenciadores para o setor econômico norte-rio-grandense são: sal e petróleo. O Estado do Rio Grande do Norte é o maior produtor de sal do Brasil.

3.3.2.1.4 Pau dos Ferros (Ponto 54 do ALiB)

O município de Pau dos Ferros está localizado na região do Alto Oeste. Está distante 389 quilômetros da capital do Estado, Natal, ocupando uma área de aproximadamente 260 km². Sua população é de 30.183 habitantes, segundo o IBGE/2018, tornando-se o décimo oitavo município mais populoso do Estado.

Em relação às origens e emancipação, no *site* Wikipédia, afirma-se que, durante muito tempo, a região do atual município de Pau dos Ferros foi habitada pelos índios panatis. No início do século XVIII, vaqueiros e viajantes descobriram o rio Apodi e ali paravam para descansar. Assim, foram surgindo pontos comerciais nas proximidades do rio e das árvores que serviam de descanso e marcação de gado, nos seus troncos.

Na zona oeste do Rio Grande do Norte só existiam três povoados: Apodi, Portalegre e Pau dos Ferros. Somente este último, devido à sua localização estratégica e privilegiada entre

duas grandes serras, tinha um crescimento regular. A cidade foi emancipada de Portalegre e o nome “Pau dos Ferros” é referência a uma árvore, que, pela sua grande dimensão, oferecia sombra e, consequentemente, um local para repouso dos vaqueiros que deram origem ao povoamento da região.

Pau dos Ferros é cortado pelo rio Apodi/Mossoró e possui todo o seu território situado na bacia hidrográfica de nome homônimo. Em relação à vegetação, a cidade é formada pela caatinga hiperxerófila¹⁵, típica do sertão nordestino, sem folhas na estação seca e com a predominância de plantas de baixo porte.

O artesanato, manifestação cultural pau-ferrense, é uma das formas mais espontâneas de uma parte da população. É possível encontrar uma produção feita com matérias-primas regionais, como o bordado e a madeira, além da culinária típica local.

3.3.2.1.5 Caicó (Ponto 55 do ALiB)

A cidade de Caicó está posicionada na região do Seridó, distante 282 km da capital estadual. Sua população, de acordo com estimativas do IBGE/2017, era de 68.222 habitantes. Reconhecida como a sétima cidade mais populosa do Estado, seu território ocupa uma área de 1.228.574 km², o equivalente a 2,33% da superfície estadual.

Situada na confluência dos rios Seridó e Barra Nova, na microrregião do Seridó Ocidental, é lembrada devido à sua principal festa, a festa de Nossa Senhora Santana. Em relação à cultura, a cidade é reconhecida pelos famosos bordados.

Encontramos no *site* oficial do município algumas explicações para a origem do nome Caicó, cuja derivação deve-se à língua cariri, significando “mato ralo”. Acredita-se também que a região foi habitada pelos índios caicó, da família dos cariri e que os mesmos denominaram a região de Caicó, que significaria “macaco esfolado” por causa dos serrotes pelos quais a vegetação era desmatada. Outra versão, defendida por Câmara Cascudo (1955) é que o topônimo provém dos termos “Acauã” e “Cuó”, que servem à designação de acidentes geográficos (rio e serra, respectivamente).

Sobre a economia de Caicó, devemos lembrar que a plantação de algodão cresceu de maneira significativa nas terras do Seridó, em que, até então, dominava a pecuária. Caicó, assim como toda a região do Seridó, se orgulhava de produzir uma das melhores variedades de algodão do mundo, o algodão mocó ou algodão seridó, variedade que resistia às secas. “Foi em

15 Paisagem esbranquiçada que aparece no período seco.

Caicó no ano de 1984, que se deu o primeiro registro da colheita de algodão de fibra colorida. Até os anos 70, a base da economia de Caicó deveu-se ao setor primário" (FEMENICK, 2010, p. 58).

Caicó está entre os municípios mais quentes do Rio Grande do Norte, ao lado de Mossoró e Pau dos Ferros, ambos na região oeste do Estado, com temperaturas que atingem facilmente 37°C.

3.3.2.2 Paraíba

O Estado da Paraíba possui 223 municípios, com o total de 3.914.421 habitantes, em uma área de 56.469,778 km². O Estado conta com seis localidades na rede de pontos do Projeto ALiB, como mostra a Figura 6.

Figura 6: Localidades da Paraíba que fazem parte da rede de pontos do ALiB



3.3.2.2.1 Cuité (Ponto 56 do ALiB)

A localidade de Cuité está situada na região centro-oeste do Estado da Paraíba, mesorregião do Agreste Paraibano e microrregião do Curimataú Ocidental, a 235 km de distância da capital, João Pessoa. A cidade limita-se com as cidades de: Cacimba de Dentro, Damião, Barra de Santa Rosa, Sossego, Nova Floresta, Picuí, Baraúna e o Estado do Rio Grande do Norte. Possui uma área de 741.840 km² e altitude de 750m acima do nível do mar. Sua população foi contabilizada em 19.978 habitantes pelo censo de 2010, com estimativa de 20.348 habitantes, em 2017.

O nome Cuité é proveniente de um fruto denominado coitezeiro, utilizado pelos índios que habitavam a região para confecção de cuias, gamelas e cochos. No dialeto indígena, *cuy* significava vasilha e *eté*, grande, real.

Conta-se, no *site* oficial da Câmara de Cuité, que no Curimataú habitavam as tribos Janduí e Canindé, especialmente, no local onde hoje está localizada a cidade de Cuité. O povoamento da Serra de Cuité iniciou-se após a guerra de expulsão dos bárbaros, na qual o homem civilizado, de forma cruel e revoltante, eliminou os primeiros habitantes da região. Os primeiros povoadores vieram das margens do Rio São Francisco e da Zona da Mata de Pernambuco em busca de terras propícias à lavoura e à criação de gado.

Sobre à hidrografia, o município de Cuité está inserido nos domínios da bacia hidrográfica do Rio Jacu. Os principais cursos de água são os rios: Jacu, Japi, Pinta Cachorro, Campo Comprido, do Paturá, de Trapiá e Bonsucesso.

Em relação à economia, a Capital do Curimataú e Seridó da Paraíba tem sua economia em forte desenvolvimento. A cidade tem uma localização estratégica para os negócios, fazendo fronteira com vários municípios do Rio Grande do Norte. Por ser uma cidade maior que suas vizinhas, as pessoas de outros municípios se deslocam até Cuité em busca de serviços, entretenimento e produtos, que são disponibilizados na cidade (cf. *site* oficial da Câmara de Cuité).

Um fator que ajudou no desenvolvimento da cidade foi a chegada da Universidade Federal de Campina Grande, impulsionando o comércio e a indústria, principalmente a indústria da construção.

3.3.2.2 Cajazeiras (Ponto 57 do ALiB)

A origem do nome faz referência a uma fazenda fundada no século XVIII por Luiz Gomes de Albuquerque, onde existiam plantadas várias cajazeiras, árvores que produzem o fruto cajá. É o sétimo município mais populoso da Paraíba.

Economicamente, o setor terciário é sua principal fonte de renda, sendo o comércio e os serviços importantes atividades econômicas no município. As indústrias mais abundantes são: a alimentícia, a de construção civil e a têxtil.

O clima é definido como semiárido, quente e seco, caracterizado pelas baixas nebulosidades e elevadas temperaturas. As chuvas são distribuídas entre os meses de janeiro e maio, de maneira irregular.

Situada na extremidade ocidental do Estado da Paraíba, está distante 468 km da capital do Estado, João Pessoa. Ocupa uma área de aproximadamente 566 km² e sua população, conforme dados do IBGE/2018, era de 61.776 habitantes, sendo o sétimo município mais populoso da Paraíba. São cidades limites: São João do Rio do Peixe, Nazarezinho, São José de Piranhas, Cachoeira dos Índios, Bom Jesus e Barro, localizada no Ceará. Pertence à região geográfica Intermediária de Souza-Cajazeiras e à região geográfica Imediata de Cajazeiras.

De acordo ao *site* Wikipédia, Cajazeiras possui todo o seu território inserido na sub-bacia do rio do Peixe, na bacia hidrográfica do rio Piranhas-Açu, e abriga dois importantes cursos de água da Paraíba: o açude Engenheiro Ávidos, com capacidade para 255 milhões de metros quadrados, e a Lagoa do Arroz. Sobre a vegetação, a caatinga predomina no município, com plantas de médio e grande porte; a espécie xerofítica é a espécie em destaque.

Em relação à cultura, o artesanato é uma das formas mais espontâneas da expressão cultural cajazeirense. É comum, em diversas partes do município, encontrar uma produção artesanal diferenciada, criada de acordo com a cultura e o modo de vida local.

3.3.2.3 Itaporanga (Ponto 58 do ALiB)

Município localizado na região metropolitana do vale do Piancó, apresenta uma população estimada em 24.653 habitantes, conforme dados do IBGE/2018. Está distante 420 km da capital do Estado.

Misericórdia, como assim era chamada, nasceu ao lado do rio Piancó e próximo a uma capela. Sua história começa quando, em 1765, Antônio Vilela de Carvalho, um português

pioneiro, comprou um pedaço de terra, onde construiu uma casa de morada e um curral para a criação de gado, à margem do rio. Naquele local, anos depois, se iniciou uma pequena povoação, que depois passou a ser conhecida por Misericórdia Velha, já que os primeiros habitantes do lugar atravessaram o rio e foram se fixar na outra margem. A cidade permaneceu por 63 anos com o seu nome de origem. Em 1938, passou a se chamar Itaporanga, que, em tupi, significa “pedra bonita” (cf. *site* Wikipédia).

Em relação à emancipação política, a vila a ganhou desligando-se de Piancó em 1863, através da Lei Provincial 104, adquirindo o nome de Freguesia de Nossa Senhora da Conceição de Misericórdia.

3.3.2.2.4 Patos (Ponto 59 do ALiB)

Localizada no Vale do rio Espinharas, é a quarta maior cidade do Estado, com uma população de 106.984 habitantes, e a vigésima oitava cidade mais populosa do interior do Nordeste. Está distante 307 km da capital do Estado.

Lê-se no *site* Wikipédia que, até meados do século XVII, toda a região que abrange o território do atual município de Patos era habitada pelos índios pegas e panatis. Logo depois desse período, as pessoas que povoaram a região foram os membros da família Oliveira Ledo, que fundaram algumas fazendas de gado. Depois disso, outras fazendas foram sendo habitadas por colonizadores portugueses que ali se estabeleceram com seus escravizados.

Inicialmente, o local chamava-se Itatiunga, nome dado pelos gentios que significa “pedra branca”. Mais tarde, passou a se chamar Patos. Conforme informações apresentadas no *site* oficial da Fundação Ernani Satyro, a denominação Patos se originou do nome de uma lagoa, hoje aterrada, situada às margens do rio Espinharas, conhecida por Lagoa dos Patos, em consequência da grande quantidade dessas aves ali encontrada.

A vegetação da região é predominantemente constituída pela caatinga; perfeitamente adaptada ao clima quente e seco, bastante sofrida pela degradação e intempéries. Na estiagem, a maioria das plantas perde as folhas e os troncos tornam-se esbranquiçados, enquanto no inverno o verde reaparece. O clima é semiárido, quente e seco, com poucas chuvas.

O município faz parte da Bacia Hidrográfica do Piranhas, umas das nove existentes no Nordeste, e é cortado por três rios: o Farinha, que nasce no município de Salgadinho e percorre Areia de Baraúnas; o rio Passagem; e o rio Cacimba de Areia.

No que concerne aos recursos minerais, encontra-se na região de Patos um grande potencial extrativista, onde se localizam as principais jazidas do sertão paraibano. Destacam-se, no município, a cultura do algodão e do feijão. As principais indústrias são as de calçado, de extração de óleos vegetais e de beneficiamento de algodão e cereais.

3.3.2.2.5 Campina Grande (Ponto 60 do ALiB)

O município tem origem na ocupação do sítio de Campina Grande pelos índios ariús. A cidade tornou-se destaque pelas feiras de gado e cultura do algodão. O posicionamento geográfico, integrando sertão com litoral, facilitou o seu desenvolvimento, uma vez que era passagem dos viajantes. Assim, devido ao progresso do local, o povoado transformou-se em vila em 1790, passando a ser chamado de Vila Nova da Rainha.

Segundo informações do *site* Wikipédia, Campina Grande está localizada no agreste paraibano, na parte oriental do Planalto da Borborema, na serra do Boturité. Está situada a 125 km da capital e possui 407.754 habitantes, segundo estimativas do IBGE/2016. Em relação à hidrografia, o município não possui rios de proporção significativa, apenas açudes.

É a segunda cidade mais populosa da Paraíba, um importante centro universitário, que conta com 21 universidades. Destaca-se ainda por ser um centro de capacitação para o nível médio e técnico.

No início do século XX, a produção de algodão tornou-se a atividade responsável pelo crescimento da cidade, atraindo especuladores de todo o Nordeste, fato que contribuiu para a cidade tornar-se a segunda maior exportadora de algodão do mundo. Curiosamente, a cidade nunca produziu algodão, mas se sobressaía nessa atividade porque, naquela redondeza, ela era a única cidade do interior do Brasil que possuía uma máquina de beneficiamento de algodão, sendo a matéria-prima necessária para a produção proveniente de cidades produtoras vizinhas (cf. *site* Wikipédia).

Campina Grande aparece entre as 100 melhores cidades para se trabalhar e fazer carreira do Brasil; é considerada a cidade mais dinâmica do Nordeste e a 6^a mais dinâmica do Brasil, segundo “A Gazeta Mercantil”. Nos dias atuais, é referência em desenvolvimento de *softwares* e de indústrias de informática e eletrônica.

3.3.2.2.6 João Pessoa (Ponto 61 do ALiB)

João Pessoa é a oitava cidade mais populosa do Nordeste, com população estimada, em 2018, de 800.323 habitantes. Fundada em 1585 com o nome de “Cidade Real de Nossa Senhora das Neves”, João Pessoa é a terceira capital de Estado mais antiga do Brasil, já fundada com o título de cidade. É conhecida como “Porta do Sol”, uma vez que o município está localizado na Ponta do Seixas, ponto mais oriental das Américas, o que faz a cidade ser reconhecida como aquela “onde o sol nasce primeiro nas Américas”.

Sobre a história de João Pessoa, salienta-se que, em 1588, a cidade adquiriu o nome de “Filipeia de Nossa Senhora das Neves”, em homenagem ao rei Filipe, o qual acumulava os tronos da Espanha e de Portugal naquele período. Depois, a cidade passou a ser chamada de Frederikstadt e, logo em seguida, adquiriu o nome de “Cidade da Parahyba”. O nome atual, João Pessoa, é uma homenagem ao político paraibano João Pessoa Cavalcanti de Albuquerque.

É importante lembrar que o município de João Pessoa nasceu às margens do rio Sanhauá; a expansão urbana ocupou a antiga área rural. Na metade dos anos de 1960, com a ocupação da orla marítima, a economia da área perdeu um pouco as características antigas.

Em conformidade com o *site* oficial do município, a cidade já recebeu o título de “cidade mais verde do Brasil”, graças ao Jardim Botânico Benjamim Maranhão, situado na área central da cidade, com 515 hectares de Mata Atlântica preservada, constituindo a maior floresta semiequatorial nativa plana densamente cercada por área urbana do mundo. Dessa maneira, torna-se uma das capitais de melhor qualidade de vida do Nordeste.

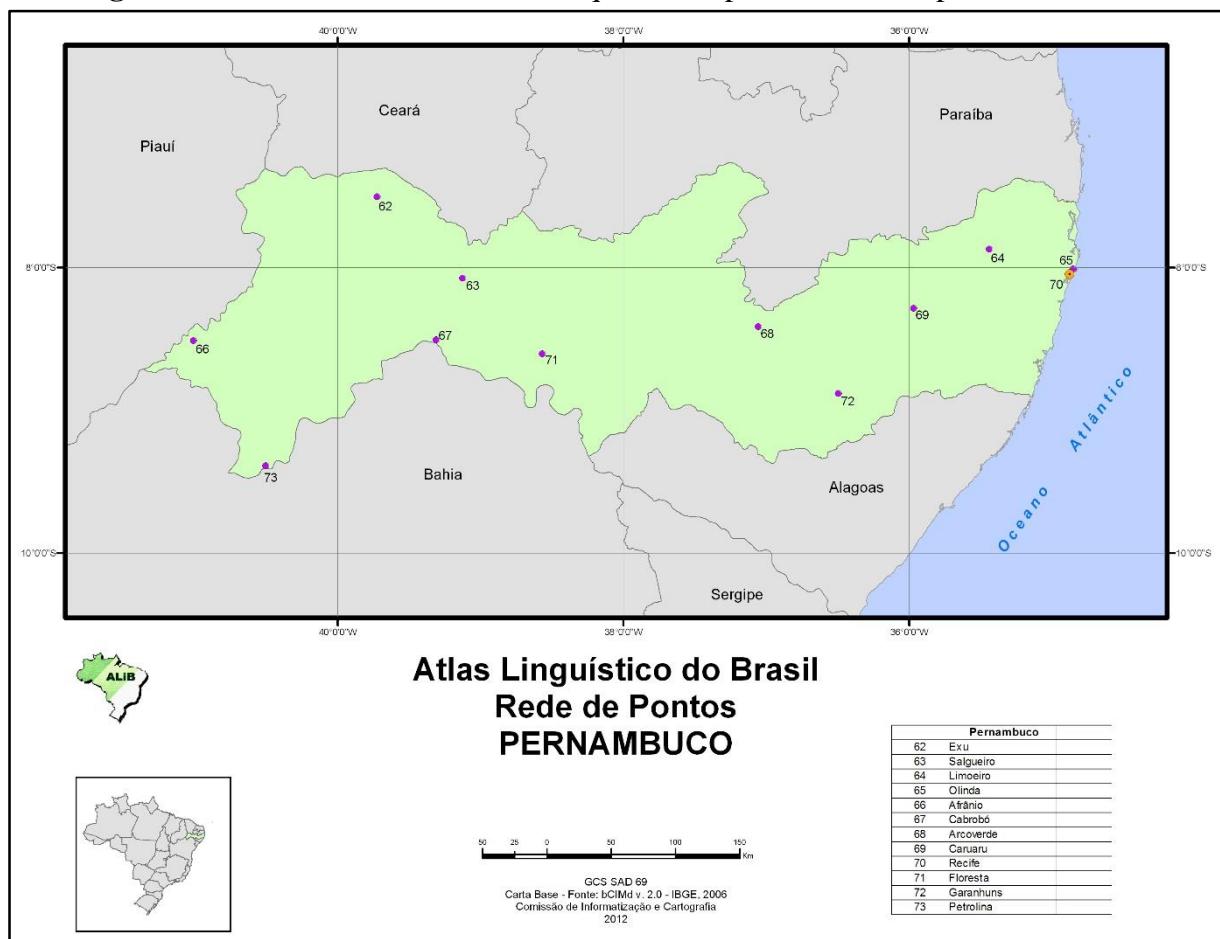
Reconhecida como patrimônio nacional do mundo, a cidade teve seu centro histórico tombado pelo Instituto do Patrimônio Histórico e Artístico Nacional (IPHAN) em 2007, inscrito nos seguintes Livros do Tombo: Histórico e Arqueológico, Etnográfico e Paisagístico. Seu tombamento foi em circunstância de ser uma das primeiras cidades fundadas no Brasil. Entre as construções que estão presentes no centro histórico da capital paraibana, destacam-se vários prédios de diferentes períodos da história.

O litoral paraibano tem cerca de 24 quilômetros de extensão, nove praias no município, além das praias da Região Metropolitana, a exemplo das cidades de Cabedelo, Lucena e do distrito de Jacumã. As praias urbanas ainda possuem Mata Atlântica preservada, uma barreira natural a cerca de 6 km da costa que protege grande parte do litoral pessoense.

3.3.2.3 Pernambuco

O Estado de Pernambuco está dividido em 185 municípios, com um total de 9.208.550 habitantes, em uma área de 98.148.323 km². Há um total de doze localidades na rede de pontos do Projeto ALiB, como mostra a Figura 7.

Figura 7: Localidades de Pernambuco que fazem parte da rede de pontos do ALiB



Das 12 localidades que fazem parte da rede de pontos do ALiB, depois de Recife, a cidade de Caruaru é a terceira maior do Estado.

3.3.2.3.1 Exu (Ponto 62 do ALiB)

Inicialmente, a região onde está o município de Exu, foi habitada pelos índios ançus, da família dos cariris. Mais tarde, essa região começou a receber fazendeiros e logo em seguida

vieram as missões jesuíticas, fatos que contribuíram, de modo significativo, para a formação da cidade.

Com base nos dados do IBGE, há duas versões prováveis para o nome do município: uma defende que o prenome é resultado de um eufemismo de *ançu*, uma tribo indígena que habitava nas proximidades; a outra argumenta que o nome tenha vindo da abelha *enxu* (*inxu*), inseto comum naquela época. É importante lembrar que dificilmente as origens do nome sejam no orixá de mesmo nome, uma vez que naquele período a vila, que fora criada por colonos, tinha como primeiros habitantes os índios.

Exu está localizada no polígono da seca, na Mesorregião do sertão pernambucano e Microrregião de Araripina. Está a 630 km da capital, Recife, e abriga um museu em homenagem ao seu filho mais ilustre, Luiz Gonzaga. O município se estende por 1.337,5 km² e somava 31.636 habitantes no último censo.

A maior parte da cidade está inserida na unidade geoambiental dos Maciços e Serras Baixas, com altitudes entre 300 e 800 metros. Ao norte, uma porção está inserida na unidade geoambiental das Chapadas Altas.

No que diz respeito ao relevo de Exu, encontramos terrenos planos e acidentados, característicos da Chapada do Araripe, com terras altas e planas. A vegetação é, predominantemente, de floresta caducifólia e caatinga. O município também apresenta vegetação do cerrado no sopé da chapada, com espécies diferenciadas, como ameixa, pequi, amburana, entre outras. O clima é tropical com estação seca, frio no inverno e quente no verão.

3.3.2.3.2 Salgueiro (Ponto 63 do ALiB)

O município de Salgueiro está localizado na Mesorregião do sertão Pernambucano e na Microrregião de Salgueiro. Tem ligação fácil com o Sul e o Sudeste do país, através da principal rodovia, a BR-116.

Salgueiro é a principal cidade da região do sertão central pernambucano, detendo, a nível regional, um comércio diversificado. Situa-se no cruzamento das BR-232 e BR-116. Desse modo, ficou conhecida como a “Encruzilhada do Nordeste”, oferecendo, portanto, uma localização estratégica do ponto de vista logístico (cf. *site* Wikipédia).

O povoamento do município antecede a época em que chegaram os primeiros colonizadores, uma vez que as terras de Salgueiro já eram habitadas por índios cariris. A região

atraiu grandes fazendeiros para criação de gado, devido à fertilidade do solo e, assim, edificaram-se várias propriedades rurais.

No tocante ao clima, a cidade tem verão quente e chuvoso, típico do clima semiárido. O inverno é seco e ameno. Devido ao fato de o relevo e o clima serem variados, a região apresenta áreas de sequeiro com chuvas escassas e mal distribuídas.

Em relação à economia, Salgueiro tem como atividades econômicas de bastante relevância a agricultura e o comércio varejista. Logo os produtos agrícolas em destaque na cidade de Salgueiro são cebola, algodão, milho, feijão, arroz, entre outros.

3.3.2.3.3 Limoeiro (Ponto 64 do ALiB)

Localizada na Mesorregião do agreste pernambucano, na unidade geoambiental do Planalto da Borborema, Limoeiro possui uma área de 273.739 km². O nome da cidade justifica-se por existirem, na região, muitos limoeiros. O município foi uma aldeia de índios tupis e sua localização estratégica contribuiu para o seu povoamento, pois era pouso para tropeiros em direção ao porto do Recife.

Limoeiro destacou-se pelo cultivo de algodão há várias décadas; toda a região de Limoeiro viu impulsionar o comércio desse item agrário, o que gerou um grande desenvolvimento econômico à cidade. O município destaca-se também pela produção de óleo comestível, além de ser região de pecuária. Atualmente, é produtor de banana, maracujá, limão, mandioca e mamão.

Limoeiro apresenta clima com características do tipo tropical sub-úmido, com período chuvoso nos meses de outono–inverno, entre março e agosto. O período seco ocorre nos meses de primavera–verão, entre setembro e fevereiro. O município está inserido na bacia do rio Capibaribe e sua vegetação é a caatinga arbórea.

3.3.2.3.4 Olinda (Ponto 65 do ALiB)

Uma das mais antigas cidades brasileiras, fundada em 1535, Olinda foi a primeira capital do Estado e da cultura, declarada em 1982 pela UNESCO. Destacam-se, na cidade, excepcionais exemplos de arquitetura religiosa dos séculos XVI e XVII, como o Convento e Igreja de Nossa Senhora do Carmo e o Convento de Nossa Senhora das Neves, que integra o conjunto arquitetônico do Convento de São Francisco.

Os primeiros habitantes foram os índios tapuias, que logo foram expulsos para outras localidades pela chegada da etnia tupi dos caetés.

Olinda está localizada no litoral do Estado de Pernambuco, na Mesorregião Metropolitana do Recife e Microrregião do Recife, distante 6km da capital pernambucana.

Segundo o IBGE, Olinda tinha, em 2016, uma população estimada em 390.144 habitantes numa área de 37,9 km². É um município essencialmente habitacional, comercial e turístico. Durante todo o ano, em especial no sítio histórico de Olinda, há eventos culturais, como feirinhas de artesanato, *reggaes*, sambas, maracatus e afoxés.

Olinda possui relevo formado por planícies e colinas, algumas delas íngremes. A cidade possui a maior parte de seu território na bacia hidrográfica do rio Paratibe (55,13%) e o restante na bacia do rio Capibaribe (44,87%). A vegetação é formada por Mata Atlântica, com espécies de grande porte.

Consoante o *site* Wikipédia, a toponímia do nome se deve a um fidalgo português, Duarte Coelho, primeiro donatário da Capitania de Pernambuco a ter utilizado a frase exclamativa: “Oh, linda situação para se construir uma vila!”, referindo-se à cidade. Daí o nome Olinda.

Em tempos remotos, a cidade já foi comparada com uma “Lisboa pequena”, talvez pela grandeza dos casarões antigos com fachadas de azulejos, que lembram as obras neoclássicas. Enfim, apresenta um perfil urbano com características portuguesas do período medieval. Em meados do século XIX, a cidade deixou de ser a capital de Pernambuco.

3.3.2.3.5 Afrânio (Ponto 66 do ALiB)

O município de Afrânio está localizado no extremo oeste do Estado de Pernambuco. O início da sua população deu-se por conta da construção da estrada de ferro Petrolina–Teresina e ali inaugurou-se, em 1926, a Estação Inveja, nome dado ao povoado. Mais tarde, a denominação dada ao pequeno povoado foi mudada para São João, por Frei Fortunato. Logo depois, passou à categoria de vila, depois a distrito de Petrolina, sendo comumente chamado de São João de Afrânio, em referência ao engenheiro da estrada de ferro. Enfim, só em 1963, Afrânio foi eleito à categoria de município autônomo.

De acordo ao censo de 2013 do IBGE, Afrânio possui uma população de 18.625 habitantes, distribuídos numa área de 1.490.596 km².

A cidade situa-se na bacia do rio Pontal. Seus principais riachos são: do Barreiro, do Caboclo, Cachoeirinha, Cachoeira do Roberto, Baixa do Morro, Barra da Cabeceira, Satisfeito, dentre outros. Em relação ao clima, o município tem o clima semiárido, do tipo clima de estepe, ou seja, os verões são quentes e úmidos, e é neste período que se concentra, praticamente, quase toda a chuva do ano; os invernos são mornos e secos.

A economia afraniense deve muito à produção leiteira, uma vez que o município é um dos maiores produtores de leite do Estado e o maior da região São Francisco pernambucano. Em relação à cultura, produz trabalhos com barro, palha e cabaça.

3.3.2.3.6 Cabrobó (Ponto 67 do ALiB)

Cabrobó está localizada no sertão do São Francisco, posicionada nas demarcações da RIDE (Região Integrada de Desenvolvimento) polo Petrolina e Juazeiro. É ponto de partida do eixo norte da Transposição do Rio São Francisco e tem a agricultura como principal fonte de renda. Sua área é de 1.658 km² e a população conta com 33.856 habitantes, conforme IBGE/2017. Está situada a 531km da capital.

Segundo informações do *site* Wikipédia, inicialmente, viviam na região índios das etnias Truká e Pancararus. A toponímia Cabrobó é de origem indígena e significa “árvore ou mato de urubus”, resultado de *caa* (árvore) e *orobó* (urubu). A chegada dos primeiros moradores para Cabrobó teve início na segunda metade do século XVIII, em torno de uma aldeia de índios.

A vegetação predominante é a caatinga hiperxerófila. Cabrobó é um dos municípios com maior risco de desertificação¹⁶ do país. A caatinga dessa região tem sido bastante modificada pelo homem. Os solos vêm sofrendo um processo intenso de desertificação, devido à substituição da vegetação natural por campos de cultivos.

No que concerne à hidrografia da cidade, o município encontra-se na bacia do rio São Francisco. Seu principal açude é Barra do Chapéu, com 1.600.000 m³. O clima do município é o semiárido quente, com verões quentes e úmidos e invernos mais amenos e secos, com a diminuição das chuvas. Sobre a economia, a cebola é o principal produto primário do município.

¹⁶ Empodrecimento e diminuição da umidade em solos arenosos, localizados em regiões de clima árido e semiárido.

Sobre o turismo do município, um dos principais pontos que atrai muitos turistas são suas cachoeiras e as águas verdes do rio São Francisco. Há, no município, eventos como a Festa da Cerveja e as Vaquejadas, e festejos de São João.

3.3.2.3.7 Arcoverde (Ponto 68 do ALiB)

A cidade está localizada na Microrregião do sertão do Moxotó, a oeste de Recife, distante 256 km da capital. Sua população está estimada em 73.844 habitantes. Possui características urbanas consistentes, reflexo de sua taxa de 90% de urbanização, possuindo a maior população urbana da Mesorregião do sertão pernambucano.

O surgimento da cidade deu-se graças ao povoado de Olho d'Água, que foi se desenvolvendo com o aparecimento de algumas casas e, mais tarde, passou à categoria de vila. O topônimo foi alterado para Arcoverde em homenagem ao cardeal Arcoverde, natural do município. Depois da inauguração da ferrovia, outro fator que muito contribuiu para a consolidação econômica do local foi a criação da feira de gado, em 1916.

No site oficial do município, confirma-se que Arcoverde é também um dos principais acessos para o *Parque Nacional do Catimbau*, localizado a 40 km. A reserva é famosa pelas atrações arqueológicas, apreciadas através de oito trilhas. O município tem características de grandes centros urbanos; é um importante polo comercial, de serviços e de entidades governamentais do interior do Estado. Atende cerca de 25 municípios circunvizinhos, garantindo ao município a marca de mais importante centro expedidor e distribuidor de mercadorias para o sertão pernambucano. É também um grande polo na área médica.

Quanto à hidrografia, Arcoverde está inserida nos domínios das bacias hidrográficas dos rios Ipanema e Moxotó. Sobre o relevo, o município está inserido na unidade geoambiental do Planalto da Borborema, formada por maciços e outeiros altos, com altitude variadas. A cidade faz parte da Região de Desenvolvimento do Moxotó, cuja economia é baseada na agropecuária e nas atividades pastoris.

No que tange à cultura de Arcoverde, o município conta com o cinema Rio Branco, o mais antigo em funcionamento do Brasil.

3.3.2.3.8 Caruaru (Ponto 69 do ALiB)

Município do Vale do Ipojuca, conhecido como a *Princesa do Agreste* e a *Capital do Forró*, Caruaru é o município mais populoso do interior de Pernambuco. Está localizado no oeste da capital do Estado, distante desta cerca de 130 km. Ocupa uma área de 920.611 km², sendo que 16.65 km² estão em perímetro urbano e os 903.961 km² restantes formam a zona rural. De acordo as estimativas do IBGE/2018, a cidade possui cerca de 356.872 habitantes.

Atesta-se no *site* Wikipédia que a toponímia Caruaru é resultado da junção oriunda do dialeto dos índios cariris, habitantes da região na época do desbravamento. *Caru* equivaleria a alimento e fartura e *aru* à abundância. Logo Caruaru significaria “terra da fartura”. Outra explicação para o nome é: planta vulgarmente conhecida por caruru (espécie de bredo) e que outrora cobria um poço na margem do rio Ipojuca; assim, com o acréscimo apenas do A, resultou em Caruaru.

Geograficamente, a cidade está situada na unidade geoambiental da Província da Borborema, sendo formada por maciços e outeiros altos. Localizada no Planalto da Borborema, seu ponto culminante é o Morro do Bom Jesus. Situa-se na bacia hidrográfica do rio Ipojuca e do rio Capibaribe.

A cidade ficou mundialmente conhecida devido ao artesanato de barro produzido pelas mãos de Vitalino Pereira dos Santos, conhecido como mestre Vitalino, que representou Pernambuco na exposição de Arte Primitiva e Moderna Brasileira no ano de 1955, na Suíça.

Houve um desenvolvimento maior do município quando foi construída a linha férrea, unindo a cidade à capital, Recife, uma vez que esse transporte levava a produção agrícola e as mercadorias de sua tradicional feira.

Segundo o IBGE, a cidade é responsável pela maior festa junina do mundo, com mais de 1,5 milhões de visitantes, o que a levou a ganhar o registro de *Guinness World Records* como a maior festa *country* regional ao ar livre do mundo.

3.3.2.3.9 Recife (Ponto 70 do ALiB)

Capital do Estado de Pernambuco, Recife é a nona cidade mais populosa do país e a quarta maior rede urbana do Brasil em população. Por estar situada próximo à linha do Equador, apresenta clima ensolarado e temperaturas médias elevadas na maior parte do ano.

Informações colhidas no *site* Wikipédia confirmam que a cidade de Recife tem sua origem vinculada ao município de Olinda, uma vez que o Recife permaneceu pertencente aos portugueses até a independência do Brasil e Olinda era sede do Estado, local escolhido pela aristocracia para moradia.

Em relação à sua localização, o município pertence às regiões geográficas Intermediária e Imediata do Recife, segundo a divisão regional vigente desde 2017, divulgada pelo IBGE. Faz divisa ao sul com os municípios de Jaboatão e Guararapes; a oeste com São Lourenço da Mata e Camaragibe; e ao norte com os municípios de Paulista e Olinda.

Ainda no *site* Wikipédia, declara-se que Recife possui uma vasta rede de rios e canais, o que a faz conhecida como “Veneza Brasileira”, devido à semelhança fluvial com a cidade de Veneza, na Europa. Destacam-se as bacias dos rios Capibaribe, Beberibe e Tejipió. O município conta com dezenas de pontes, entre elas a mais antiga da América Latina, a Ponte Maurício de Nassau. Há uma quantidade significativa de ilhas e mangues que fazem parte da riqueza hídrica recifense.

Quanto ao clima, a cidade apresenta um clima tropical úmido, típico do litoral leste nordestino. As precipitações surgem sob a forma de chuvas, podendo vir acompanhadas de raios e trovões com certa intensidade, ocasionando, portanto, alagamentos.

3.3.2.3.10 Floresta (Ponto 71 do ALiB)

O município de Floresta é o ponto de partida do Eixo Leste da Transposição do rio São Francisco. Está distante 433 km da capital pernambucana. Possui uma população de 32.152 habitantes, distribuídos numa área de 3.644.168 km² (sendo o segundo maior município pernambucano em extensão territorial), segundo o IBGE/2016. Localiza-se no sertão pernambucano, na Mesorregião do São Francisco, Microrregião de Itaparica.

Floresta surgiu a partir do povoamento das fazendas Curralinho e Paus Pretos, porém foi na Fazenda Grande, à margem direita do rio Pajeú, que foram surgindo as primeiras moradias. A fazenda servia de curral temporário para o gado que vinha da Bahia abastecer os engenhos de açúcar pernambucanos.

O município está inserido na bacia do Rio São Francisco e do Rio Pajeú. Seus principais riachos são: do Capim Grosso, da Lagoinha, do Navio, das Porteiras, do Papagaio, do Toco, da Pedra Branca, da Salina, Poço da Areia, da Travessa, da Várzea, do Carcarazeiro, do Mari, dentre outros. Quanto ao clima, enquanto os verões são quentes e úmidos, os invernos

são mornos e secos. A vegetação do município é composta por caatinga hiperxerófila e com trechos de Mata Atlântica (*site Wikipédia*).

O *site* oficial da prefeitura informa que o município de Floresta destaca-se no Estado de Pernambuco pela sua liderança no tamanho de seu rebanho de caprinos e ovinos (um dos maiores do país). Floresta também é um dos maiores produtores de tomate e melancia. Além da pecuária, desenvolve-se a agricultura irrigada, presente nas margens do rio Pajeú e do riacho do Navio e nas proximidades do lago de Itaparica, como mais uma alternativa de crescimento econômico, resultante da implantação de projetos irrigados, a partir do Canal Eixo Leste/PTSF¹⁷.

3.3.2.3.11 Garanhuns (Ponto 72 do ALiB)

De acordo com informações do *site Wikipédia*, Garanhuns está situada no planalto da Borborema, a 842 metros acima do nível do mar. No ponto mais elevado, a altitude chega a 1.030 metros. É o principal município do Agreste Meridional, distante 230 km da capital do Estado. O IBGE estimou sua população em aproximadamente 137. 810 habitantes, sendo o nono município mais populoso de Pernambuco.

A toponímia Garanhuns é explicada através de algumas versões: primeiro, refere-se a *guirá-nhum* (os pássaros pretos), nome de uma tribo existente no local; segundo, o termo vem do nome de uma tribo cairu, da nação cariri ou quiriri; e, por último, é resultante de uma palavra de origem indígena que significa “sítio de *guardás* e *anuns*”, formado de *guardás* (espécie de cão selvagem) e *anuns* (ave tida como agoureira) (cf. *site Wikipédia*).

Intitulada de “Suíça Pernambucana”, “Cidade das Flores” e “Cidade do Clima Maravilhoso”, possui um clima com baixas temperaturas e ares europeus, mesmo no verão em pleno Nordeste brasileiro, graças às sete colinas que abraçam a região. A paisagem permanentemente verde e a quantidade e diversidade de flores espalhadas pelas praças também são responsáveis pelos diferentes títulos.

Lê-se no *site* oficial da cidade que a mesma está situada no Planalto da Borborema e é cercada pelas colinas: Antas, Columbino, Ipiranga, Magano, Monte Sinai, Quilombo e Triunfo. Pertence à bacia hidrográfica do rio Mundaú.

Vários ciclos econômicos marcaram o desenvolvimento da cidade ao longo do século XX: o primeiro deles estava ligado ao traço mais forte de sua identidade, o cultivo do café;

17 Ponto de Transposição do rio São Francisco.

depois, veio a agropecuária; e, logo em seguida, a partir de 1966, iniciou-se um novo ciclo, tendo como destaques o comércio, o turismo e serviços no município. Por fim, a região de Garanhuns se destaca também pela produção artesanal, semiartesanal e industrial de derivados do leite.

3.3.2.3.12 Petrolina (Ponto 73 do ALiB)

Sexto município mais rico de Pernambuco, Petrolina possui 769.544 habitantes e é integrante da Região Administrativa Integrada de Desenvolvimento do polo Petrolina e Juazeiro.

Encontra-se no *site* do IBGE que os primeiros habitantes do município foram os frades franciscanos que trabalhavam na catequese dos índios da região. Uma das explicações para o nome de Petrolina está relacionada a uma homenagem ao imperador D. Pedro II, que ocupava, então, o trono do Brasil. Outra explicação para o topônimo seria uma dupla homenagem, com a junção do nome do imperador, em sua forma latina (*Petrus*), ao da imperatriz Tereza Cristina, resultando em Petrolina. Outra versão sugere que o topônimo teria sido derivado de “pedra linda”, expressão dada a uma pedra que havia na margem do rio.

O município se localiza na unidade geoambiental da Depressão Sertaneja, unidade que é formada pelas principais características do semiárido nordestino. Petrolina está inserida na macrobacia hidrográfica do Rio São Francisco, do Rio Pontal e do grupo de Bacias de Pequenos Rios Interiores (cf. *site* Wikipédia).

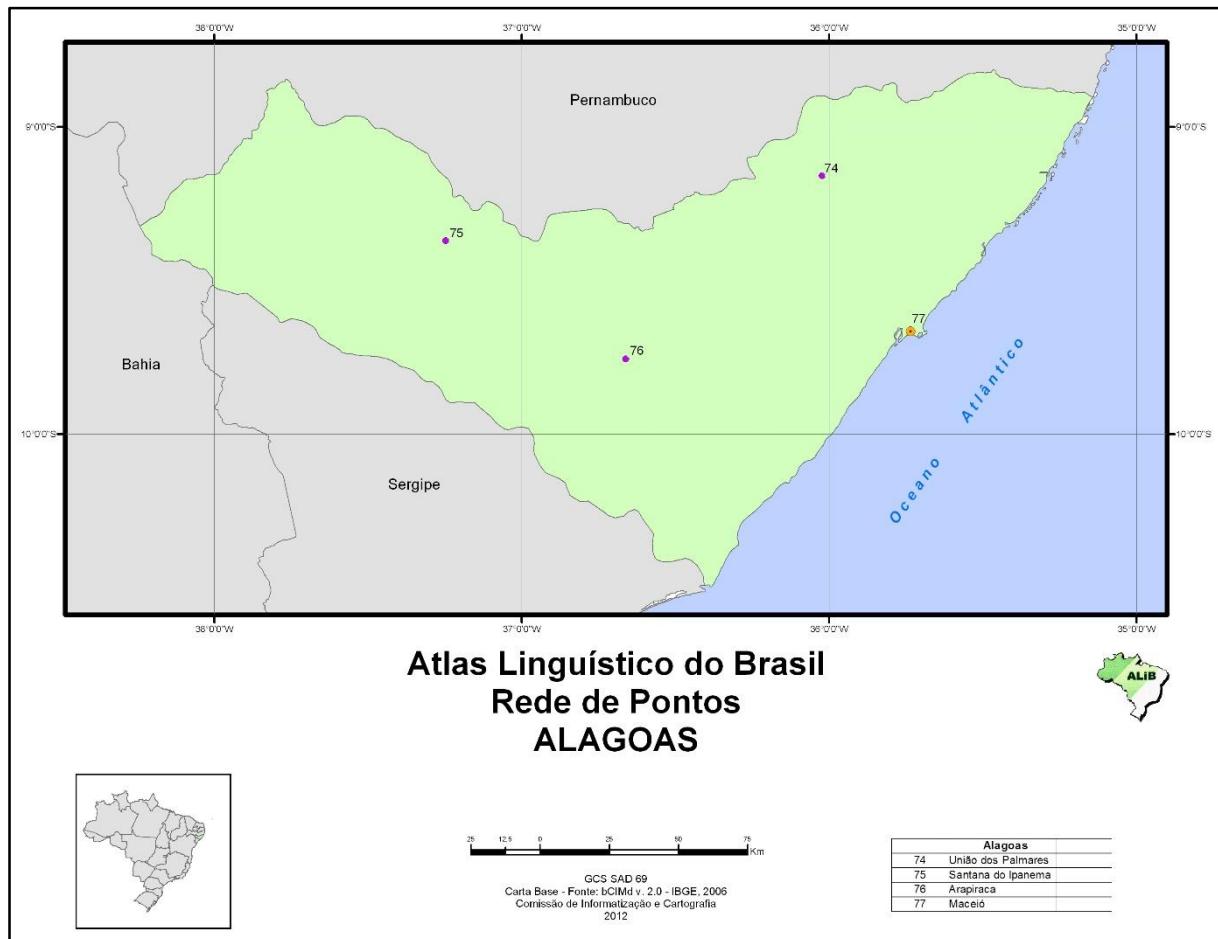
O clima petrolinense apresenta-se como semiárido quente, com chuvas de primavera-verão, também caracterizado pela escassez e irregularidade de chuvas. A vegetação nativa e predominante do município é a caatinga, exibindo arbustos com galhos retorcidos e com raízes profundas.

Muito embora a cidade pertença a uma região semiárida, destaca-se por sua agricultura irrigada, sendo reconhecida por ter o terceiro maior PIB agropecuário, ser o segundo maior produtor de vinho e o maior exportador de frutas do país.

3.3.2.4 Alagoas

O Estado de Alagoas tem 102 municípios, que somam um total de 3.300.935 habitantes, distribuídos em 27.778,506 km². É o segundo menor Estado da região Nordeste. No Estado há quatro localidades da rede de pontos do Projeto ALiB, como podemos observar na Figura 8.

Figura 8: Localidades de Alagoas que fazem parte da rede de pontos do ALiB



3.3.2.4.1 União dos Palmares (Ponto 74 do ALiB)

Localiza-se a norte da capital do Estado, Maceió, distando desta cerca de 73 km. Sua população foi estimada em 2015 pelo IBGE em 66.017 habitantes, sendo assim uma das dez cidades mais populosas do Estado de Alagoas.

O povoamento do município está relacionado com a fuga dos negros dos engenhos de açúcar dos Estados de Alagoas e Pernambuco para a sede do Quilombo dos Palmares. Quando

sede do Quilombo, a localidade chamava-se Cerca Real dos Macacos, provavelmente em referência ao Riacho dos Macacos.

A terminologia “União”, que surgiu por meio do Decreto nº 46, de 25 de setembro de 1890, teve origem devido ao fato de a cidade ter sido palco da estrada de ferro que liga Alagoas a Pernambuco. Em 1944, ocorreu a alteração definitiva para “União dos Palmares”, homenageando o Quilombo, que permaneceu na região por quase um século.

União dos Palmares é banhada pelo rio Mundaú e está localizada a 76 km da capital do Estado. A cidade faz parte da microrregião Serrana dos Quilombos, fazendo limites com Santana do Mundaú, São José da Laje, Ibateguara, Branquinha e Joaquim Gomes. O município ficou conhecido por ser “A Terra da Liberdade”, uma vez que o primeiro grito de liberdade por Zumbi dos Palmares aconteceu na referida cidade. A Serra da Barriga, principal palco do famoso Quilombo dos Palmares, hoje declarada Patrimônio Histórico do Brasil, está a cerca de 6 km da sede do atual município de União dos Palmares.

De acordo ao *site* Wikipédia, a economia do município tem as suas bases no binômio agricultura-pecuária, destacando-se como um dos maiores produtores de cana-de-açúcar de Alagoas, além de ser conhecido também como um dos maiores produtores de banana do Estado. Possui usina de açúcar e álcool, indústria de laticínios e plásticos, de cerâmicas em barro (olaria), piscicultura, suinocultura, avicultura (a mesma com as instalações mais modernas do país). A pecuária também tem contribuído de maneira relevante para a economia do município.

3.3.2.4.2 Santana do Ipanema (Ponto 75 do ALiB)

Seu povoamento deu-se graças a chegada do padre Francisco José Correia à região, onde ali, no arraial, já habitavam índios e mestiços e, com a ajuda do padre, fundou-se uma igreja. Era notória a chegada de várias pessoas, inclusive portugueses, que se instalaram nas fazendas.

Inicialmente, o povoado passou a ser chamado de Sant’Ana da Ribeira do Ipanema, por estar situado à margem do rio Ipanema. Mais tarde, foi elevado à categoria de vila com a denominação de Santana do Ipanema. Segundo pesquisa do IBGE de 2018, a estimativa demográfica é de 47.486 habitantes, sendo a principal cidade do sertão alagoano. Situa-se a 207 km da capital do Estado.

O principal recurso hídrico de Santana do Ipanema é o rio Ipanema, que tem sua nascente em Pernambuco e foz na Barra do Ipanema, formada pelo rio São Francisco.

3.3.2.4.3 Arapiraca (Ponto 76 do ALiB)

O município de Arapiraca é o segundo mais populoso do Estado de Alagoas e o primeiro de sua microrregião.

Pertencente à Mesorregião do Agreste alagoano, localiza-se a oeste da capital do Estado, distante desta cerca de 136 km. Sua população foi estimada em 232.671 habitantes pelo IBGE/2016. Sua localização central permite que seja rota para as cidades circunvizinhas e outras cidades.

As árvores robustas na região foram responsáveis pelo topônimo Arapiraca: com o passar do tempo, a árvore Arapiraca começou a atrair pessoas e, assim, o povoado foi se formando. Arapiraca era um distrito subordinado, sucessivamente, a Penedo, Porto Real do Colégio, São Brás e Limoeiro. Somente em 1924, foi elevado à categoria de município.

Em relação ao relevo de Arapiraca, o mesmo faz parte da unidade dos Tabuleiros Costeiros. Essa unidade acompanha o litoral de todo o Nordeste, os solos são profundos e de baixa fertilidade natural. O clima é do tipo tropical chuvoso com verão seco. O período chuvoso começa no outono, tendo início em fevereiro e término em outubro.

O município encontra-se inserido na hidrografia do rio São Francisco, banhado pelos afluentes rio Jurubeba e rio Coruripe. Geograficamente, o município está inserido na Província Borborema.

Lê-se no *site* da prefeitura que foi a partir de 1970 que começou o desenvolvimento econômico da cidade, principalmente quando a cultura da produção de fumo, o conhecido “*Ouro Verde*”, gerou grande produtividade no beneficiamento do tabaco. Nos dias atuais, a cidade arapiraquense vem se destacando por ser uma das que mais tem gerado empregos em todo o território nacional; muitos negócios, ligados ao comércio, vêm sendo criados na cidade.

3.3.2.4.4 Maceió (Ponto 77 do ALiB)

Maceió está situada na microrregião homônima e na mesorregião do Leste de Alagoas. Segundo dados do IBGE, a capital possui 1.012.387 habitantes. Posiciona-se na faixa costeira do nordeste oriental, inserida nos domínios da Mata Atlântica. Estende-se por uma área de aproximadamente 500 km².

A terminologia do nome “Maceió” tem origem no termo tupi *maçayó* ou *maçaio-k*, que significa “o que tapa o alagadiço”. O Dicionário Aurélio (FERREIRA, 1975) diz que o termo “maceió” designa uma lagoa temporária e cíclica que se localiza na beira do mar, na foz de um curso de água pequeno (cf. *site* Wikipédia).

Maceió identifica-se com um arquipélago, o qual é constituído por um conjunto de nove ilhas: Bora Bora, Irineu, Almirante, Fogo, Um Coqueiro, Só, Cabras, Santa Marta, e Andorinhas.

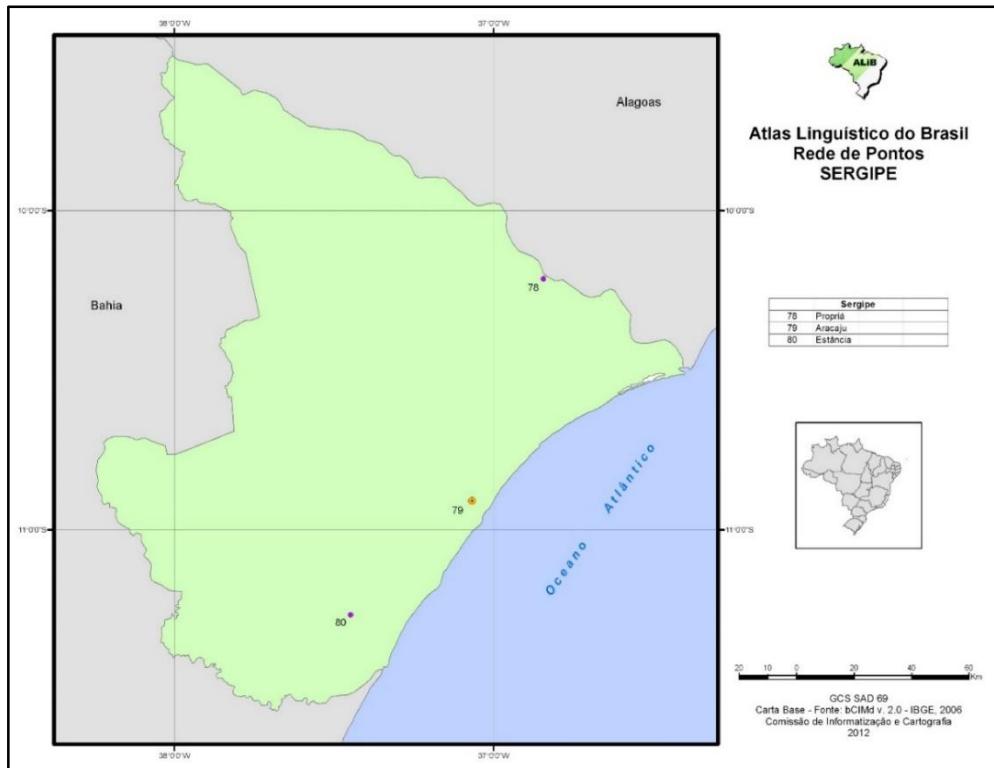
Em relação à economia, a cidade é a maior produtora brasileira de sal-gema. Seu setor industrial diversificado é composto de indústrias químicas, açucareiras, de álcool, de cimento e alimentícias. Destaca-se também na extração de gás natural e petróleo.

Nos últimos anos, a cidade de Maceió vem recebendo um grande número de imigrantes chineses, os quais foram responsáveis pelo crescimento do comércio local, ofertando produtos de 1,99 reais, importados da China.

3.3.2.5 *Sergipe*

Sergipe é o menor Estado da região Nordeste e o menor Estado brasileiro, com um total de 21.915,116 km². Tem uma população total aproximada de 2.195.662 habitantes. Tem 75 municípios e, por ser o menor, o Estado também tem o menor número de localidades do Nordeste na rede de pontos do Projeto ALiB, três, como podemos observar na Figura 9.

Figura 9: Localidades de Sergipe que fazem parte da rede de pontos do ALiB



Fonte: <https://alib.ufba.br/content/rede-de-pontos>.

3.3.2.5.1 Propriá (Ponto 78 do ALiB)

A cidade de Propriá, no passado, foi conhecida como “urubu de baixo”. Era considerada uma das principais cidades do Estado, chegando a ser a segunda mais importante na década de 1950. Porém, com a construção da Ponte da Integração Sergipe-Alagoas, na BR-101, iniciou-se o seu declínio econômico.

O município já foi destaque em relação ao cinema sergipano, quando possuía três salas de projeções, uma delas com um restaurante que servia guloseimas em mesas, enquanto as pessoas assistiam a filmes da época. No Cine Fernandes, ocorriam shows musicais de grandes nomes da música brasileira.

Lê-se no site Wikipédia que Propriá já foi a segunda economia do Estado de Sergipe (a primeira era Aracaju) e liderava o comércio atacadista do Baixo São Francisco (Sergipe e Alagoas). Hoje, é famosa pela fabricação de doces típicos, sendo o doce de batata conceituado como o melhor da região.

No que concerne à economia, Propriá é a terra da rizicultura; milhares de toneladas de arroz saem dos perímetros da Codevasf no município. Outra forma de economia do município

é o cultivo de diferentes frutas em grande quantidade, além da piscicultura, que é desenvolvida em larga escala no perímetro irrigado do município.

Convém salientar um dos símbolos da cidade de Propriá, a ponte sobre o Velho Chico, construída em meados da década de 1970, medindo 793 metros de comprimento. Por ela, as embarcações fazem a travessia entre os Estados de Sergipe e Alagoas.

3.3.2.5.2 Aracaju (Ponto 79 do ALiB)

A cidade de Aracaju localiza-se no leste do Estado e é cortada pelos rios Sergipe e Poxim. Segundo pesquisa do IBGE/2017, a estimativa da população é de 648.939 habitantes. É a menos populosa das capitais nordestinas. Destaca-se por ser a capital com menos desigualdade social do Nordeste.

A região da atual cidade de Aracaju teve, inicialmente, as suas terras habitadas por índios tapuias, que foram expulsos pelos tupis. Assim, quando vieram os europeus para colonização, o território já era ocupado pelos tupis.

A lexia Aracaju pode ter origem no termo tupi *arákaíu*, que significa “cajueiro das araras” (*ará*, arara + *akaíu*, cajueiro). A cidade foi a segunda planejada do Nordeste (a primeira foi Teresina), portanto todas as suas ruas foram projetadas geometricamente para desaguarem no rio Sergipe (cf. *site Wikipédia*).

O município pertence às regiões geográficas Intermediária e Imediata de Aracaju. O solo da cidade era principalmente composto por areia e algumas zonas eram área de manguezal, constantemente inundadas. Atualmente, essas áreas de manguezal foram soterradas. O relevo é bastante plano, o que possibilita, nos dias atuais, a prática do ciclismo.

O clima de Aracaju é quente e úmido; os meses mais quentes são janeiro, fevereiro e março, com temperatura média de 27 °C. O período mais chuvoso é entre março e julho, pois o vento forte, devido às temperaturas mais baixas no Sul e Sudeste do país, nesses meses, traz várias nuvens carregadas. Nos meses mais secos, entre setembro e fevereiro, o vento fica mais fraco; por conseguinte, chove menos.

3.3.2.5.3 Estância (Ponto 80 do ALiB)

A cidade está estabelecida no litoral sul sergipano e sua população estimada em 2018 era de 68.804 habitantes. O município é conhecido como o “jardim de Sergipe”. É considerada

a cidade dos azulejos, das festas juninas e do barco de fogo. Caracteriza-se como a terceira maior economia do Estado, perdendo apenas para Aracaju e Nossa Senhora do Socorro. É também o segundo maior polo industrial do Estado e tem um dos maiores comércios da região.

Estância era subordinada à Vila de Santa Luzia do Real. Era conceituada durante boa parte do século XX, pois foi a segunda maior cidade de Sergipe. Em 1831, a sede da Vila de Santa Luzia foi transferida para Estância e, em 1848, foi elevada à categoria de cidade.

O município pertence às regiões geográficas Intermediária de Aracaju e Imediata de Estância. Em relação ao clima, predomina o clima tropical, com meses de maior calor (janeiro, março e dezembro) e meses mais chuvosos (maio, junho, julho, agosto e setembro). Os estancianos contam com os rios Piauitinga e Piauí.

O acervo cultural do município ainda conserva exemplares remanescentes de antigos engenhos de cana-de-açúcar. O ciclo da cana-de-açúcar e o engenho foram bastante significativos na formação econômica e social do povo sergipano; o sul do Estado também participou ativamente da economia açucareira.

Na economia, destaca-se a cultura do coco e da mangaba. Já no setor secundário, as indústrias – alimentícias, têxteis, de cerveja, de perfumaria, dentre outras – contribuem para o crescimento do município.

3.4 QUESTIONÁRIOS

O questionário linguístico do Projeto ALiB foi organizado pelos membros do Comitê Nacional, seguindo como base os questionários linguísticos utilizados nos atlas estaduais e regionais publicados ou em andamento no Brasil e os questionários do *Atlas linguistique roman (ALiR) e do Atlas linguístico-ethnográfico de Portugal e da Galiza (ALEPG)*. Tais questionários foram examinados e, com base neles e na experiência dos pesquisadores, foram elaboradas versões experimentais do questionário. Após aplicações visando a testagem do instrumento, seu refinamento e sua adequação à realidade do país, gerou-se o questionário definitivo, publicado em 2001, que se aplicou na constituição do *corpus* em todo o país.

O questionário, publicado em 1998 pela Universidade Estadual de Londrina, foi organizado em três partes:

- a) Questionário Fonético-Fonológico (QFF), com 159 perguntas;
- b) Questionário Semântico-Lexical (QSL), com 207 questões; e
- c) Questionário Morfossintático (QMS), com 80 questões.

Devido à grande extensão demográfica do Brasil, cuja distribuição geográfica vai do Oiapoque (ponto 1) ao Chuí (ponto 250), estabeleceram-se, inicialmente, como primeira proposta, 235 pontos de inquérito: 23, na Região Norte; 71, no Nordeste; 79, no Sudeste; 41, no Sul; e 21, no Centro-Oeste. Mais tarde, o número de pontos foi ampliado para 250.

Na última edição, versão de 2001, os questionários linguísticos sofreram alterações, após a realização do *III Workshop para inquiridores do ALiB*, na Universidade Estadual de Londrina. Com a intenção de torná-lo mais completo, a atual equipe do ALiB acrescentou à aplicação já existente outros tipos de questionários e algumas outras maneiras de coleta de dados, que ajudaram na complementação dos resultados obtidos (COMITÊ NACIONAL DO PROJETO ALiB, 2001). Nesse sentido, o Projeto elaborou os chamados Questionários do ALiB, que constam de sete partes distintas: (1) QFF – Questionário Fonético-Fonológico (com 159 perguntas, às quais se juntam 11 questões de prosódia); (2) QSL – Questionário Semântico-Lexical (com 202 perguntas); (3) QMS – Questionário Morfossintático (com 49 perguntas); (4) QP – Questões de Pragmática (com 4 perguntas); (5) TDS – Temas para Discurso Semidirigido (com 4 temas – relato pessoal, comentário, descrição e relato não pessoal); (6) PM – Perguntas Metalinguísticas (com 6 perguntas); e (7) LE – Texto para Leitura (“Parábola dos sete vimes”).

O QSL comprehende os seguintes campos semânticos: I. Acidentes geográficos; II. Fenômenos atmosféricos; III. Astros e tempo; IV. Flora; V. Atividades agropastoris; VI. Fauna; VII. Corpo Humano; VIII. Cultura e convívio; IX. Ciclos da Vida; X. Religião e Crenças; XI. Festas e Divertimentos; XII. Habitação; XIII. Alimentação e Cozinha; XIV. Vestuário; e XV. Vida Urbana.

Esse Questionário é composto de 202 questões, por meio das quais não se busca a realização de um vocábulo específico, mas, sim, a variação lexical. São questões de cunho onomasiológico. Prestam-se também para estudos da variação diastrática, diageracional e diassexual, pois, quando os informantes são incitados a apresentarem “outras denominações” para o item semântico-lexical em questão, muitas vezes são registradas colocações como: “os mais antigos falavam assim”, “ouvia minha avó dizer isto”, “antigamente não se falava assim, não”, “o pessoal ignorante fala assim” ou “ouço dizer”. Buscam-se documentar formas linguísticas empregadas na localidade, as formas de uso comum e geral. O QSL está organizado em 14 áreas semânticas, apresentadas no Quadro 9.

Quadro 9: Áreas semânticas do QSL

| ÁREAS SEMÂNTICAS | Nº. DE PREGUNTAS |
|------------------------------------|------------------|
| 1. Acidentes geográficos | 06 |
| 2. Fenômenos atmosféricos | 15 |
| 3. Astros e tempo | 17 |
| 4. Atividades agropastorais | 25 |
| 5. Fauna | 25 |
| 6. Corpo humano | 32 |
| 7. Ciclos da vida | 15 |
| 8. Convívio e comportamento social | 11 |
| 9. Religião e crenças | 08 |
| 10. Jogos e diversões infantis | 13 |
| 11. Habitação | 08 |
| 12. Alimentação e cozinha | 12 |
| 13. Vestuário e acessórios | 06 |
| 14. Vida urbana | 09 |
| TOTAL | 202 |

Fonte: Elaborado pela autora.

O QMS está assim organizado: I. Artigo; II. Substantivo (gênero, casos de plural e concordância nominal); III. Adjetivo (comparativo de *grande/pequeno*, de *bom/mau*); IV. Pronomes pessoais (sujeito – *tu/você*, *senhor/senhora*, *vocês*; pronomes possessivos – *meu/teu/seu*, *seu/dele*; e pronomes indefinidos – *menos/meninas*); V. Verbo (flexão modo-temporal – presente indicativo, pretérito indicativo, futuro presente e futuro subjuntivo; uso de participios passados reduzidos; concordância verbal – com o verbo *fazer* e com os verbos *haver/ter*); VI. Advérbios (colocação do *não* em respostas negativas).

É a parte do questionário de maior dificuldade de aplicação, sendo necessário, às vezes, reformular a pergunta, de modo a assegurar um bom entendimento por parte do informante. Em alguns momentos utilizaram-se gravuras para vencer as dificuldades, como, por exemplo, para o plural dos substantivos. Em outros, foram solicitados pequenos relatos, descrição de receitas, simulações de ocorrências da vida cotidiana, para analisar o uso das diferentes formas linguísticas.

Salientamos que já foram realizados inquéritos em todas as regiões do Brasil, discutidos os problemas encontrados quanto aos questionários e atuações dos inquiridores e reformuladas algumas questões, a partir dos problemas levantados.

Observamos, no questionário, que todas as perguntas são acompanhadas de um texto instrutivo sobre “como devem ser formuladas, evitando, dessa forma, possíveis distorções na(s)

resposta(s) obtida(s), em decorrência do modo como se apurou a forma fornecida pelo informante” (PAIM; RIBEIRO, 2018, p. 113).

Ainda demonstrando o cuidado com o modo de aplicação do questionário e a consequente preservação da regularidade na obtenção dos dados, quando nacionalmente considerados, notamos que, além da formulação inicial, estão sinalizados outros recursos complementares para a obtenção de respostas, como: gestos, mímicas, figuras impressas ou reálias (miniaturas de objetos ou exemplares originais de alguns objetos). Existem também indicações para reformulação de perguntas (caso necessário) “e de itens para os quais são demandadas ampliações da pergunta para se obter contexto de utilização (exemplo: FAZENDA – QFF 53) ou descritivo detalhado da forma em uso na localidade (exemplo: AMARELINHA – QSL 167)” (RIBEIRO, 2012, p. 121).

A variedade de tipo de questão é um dos procedimentos que podem contribuir para amenizar o caráter artificial do diálogo informante-documentador, minimizando, assim, o que Labov (1983) intitula como “paradoxo do observador”.

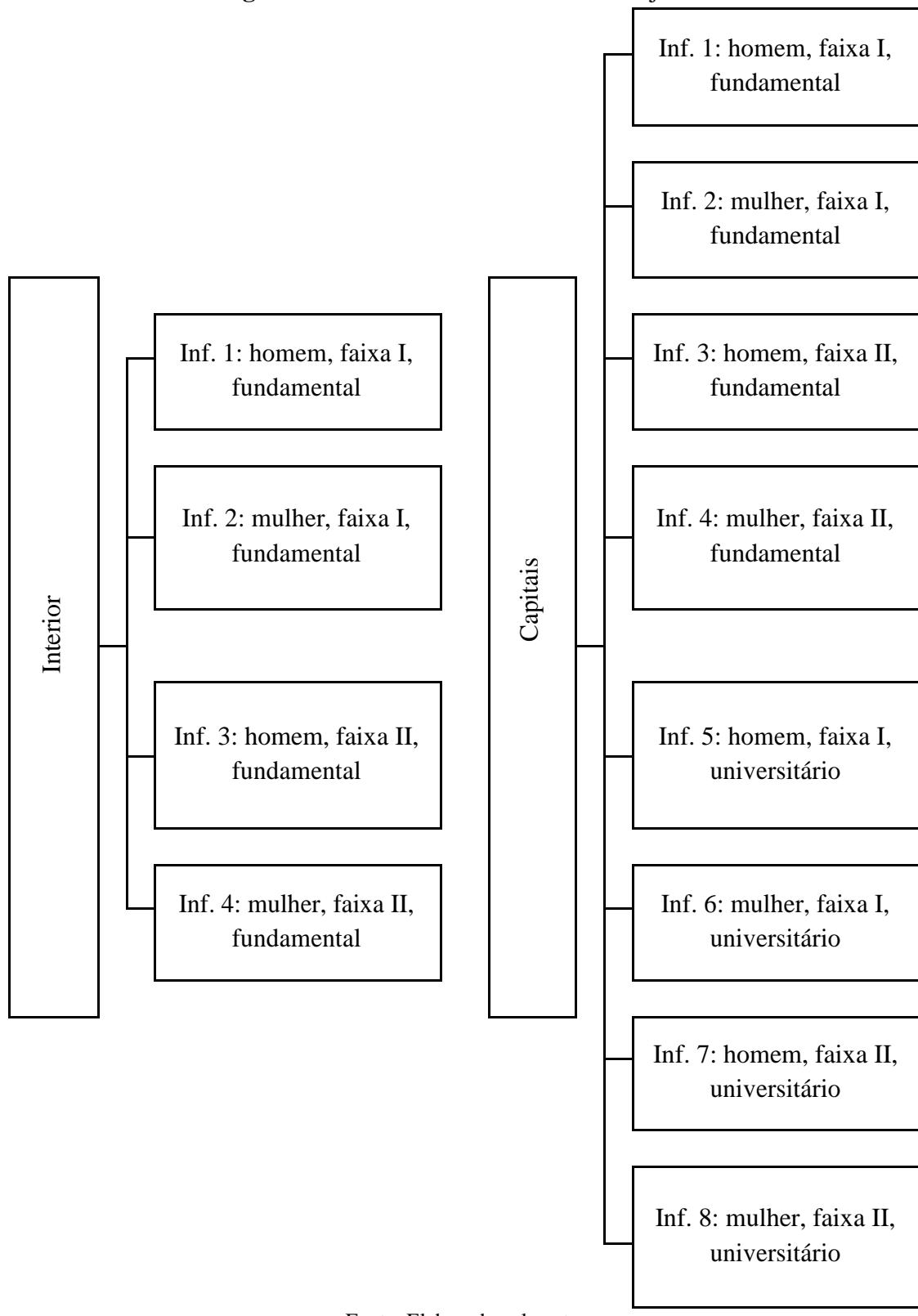
O questionário é aplicado em todas as localidades constituintes da Rede de Pontos do ALiB. As entrevistas, fruto da aplicação do questionário, duram em média três horas e são totalmente gravadas para posteriores transcrições grafemática e fonética.

O Projeto iniciou, efetivamente, os inquéritos a partir de 2001, período em que a Editora da Universidade Estadual de Londrina, publicou a versão definitiva dos questionários.

3.5 INFORMANTES

Para a realização dos inquéritos do Projeto ALiB, foram selecionados quatro informantes por localidade do interior, falantes de língua portuguesa e representativos do falar local, sendo dois do sexo masculino e dois do feminino, distribuídos em duas faixas etárias – de 18 a 30 anos (faixa I) e de 50 a 65 anos (faixa II) –, com profissão definida (variável que não incluímos nesta tese), escolarizados, no máximo até o Ensino Fundamental II (9º ano/8ª série). Nas capitais, foram acrescentados mais quatro com nível universitário de escolaridade. Assim, considerando os 250 pontos da rede do ALiB, foram gravados 1.100 informantes, oito em cada capital de Estado (com exceção de Brasília e Palmas, cidades de formação recente) e quatro em cada cidade do interior, como mostra a Figura 10.

Figura 10: Perfil dos informantes do Projeto ALiB



Fonte: Elaborada pela autora.

Para constituição do *corpus* desta tese, foram selecionadas as entrevistas de 140 informantes (quatro de cada ponto do interior e oito de cada capital), que integram o corpo de

informantes do Projeto ALiB, com o seguinte perfil: (i) são naturais da localidade pesquisada, têm pais dessa mesma localidade e não se ausentaram dela por períodos longos durante sua vida; (ii) os mais jovens têm idade compreendida entre 18 e 30 anos e os mais velhos entre 50 e 65 anos; (iii) são, em representação equitativa, dos sexos masculino e feminino, como mostra o Quadro 10.

Quadro 10: Quantidade de informantes por Estado

| ESTADO | INFORMANTES | | | TOTAL |
|---------------------|-------------|---------------|-------------|------------|
| | Capital | | Interior | |
| | Fundamental | Universitário | Fundamental | |
| Rio Grande do Norte | 04 | 04 | 16 | 24 |
| Paraíba | 04 | 04 | 20 | 28 |
| Pernambuco | 04 | 04 | 44 | 52 |
| Alagoas | 04 | 04 | 12 | 20 |
| Sergipe | 04 | 04 | 8 | 16 |
| TOTAL | 40 | | 100 | 140 |

Fonte: Elaborado pela autora.

3.6 RECOLHA E LEVANTAMENTO DE DADOS

A seguir, descrevemos as etapas sequenciais por que passou, metodologicamente, a presente pesquisa. Abordamos, inicialmente a pesquisa de campo do Projeto ALiB, tendo em vista o fato de os dados trabalhados terem sido coletados por membros desse projeto. Acrescentamos também os passos seguidos para constituição da amostra utilizada nesta pesquisa.

3.6.1 Pesquisa de campo do ALiB

Considerando o modelo laboviano, a pesquisa acontece com grupos de indivíduos, observando-se os aspectos sociais que interferem nas suas falas. Labov (1983, p. 256) salienta que o vernáculo é “propriedade de um grupo, não de um indivíduo”. Nessa direção, a preocupação de um investigador, na pesquisa de campo, deve ser a descrição de uma variedade linguística.

Em conformidade com a metodologia da Sociolinguística Variacionista, o principal método para investigação linguística é a observação direta da língua falada, usada em situações naturais de conversação. Nessa interação face a face, utiliza-se o vernáculo – estilo em que o

mínimo de monitoração ou atenção é dispensada à fala (COELHO, 2010, p. 116). Essa variedade utilizada em situações naturais de interação é a que usamos em nossos lares, em botequins, com os amigos, nos encontros de lazer, distantes do local de trabalho. A pesquisa sociolinguística objetiva coletar a variedade vernácula em momentos em que os informantes falam de maneira espontânea em entrevistas gravadas.

Nem sempre é fácil para o pesquisador captar esse momento de interação em que as falas das pessoas emanam de forma natural. Por esse motivo, Labov (1983) sinaliza para o chamado “paradoxo do observador”. Nessa busca pelo vernáculo, o informante depara-se com um grande desafio, que é a presença do gravador ou outro instrumento de registro, bem como a própria presença do inquiridor, que pode inibir a fala natural, levando o informante a prestar mais atenção à sua fala, passando, assim, a monitorá-la.

Alguns impasses da pesquisa variacionista podem ser superados, sobretudo, porque há várias possibilidades de estratégias para serem usadas numa entrevista. Por sua vez, o entrevistador pode utilizar recursos que façam o informante desviar sua atenção do contexto da entrevista.

A coleta dos dados do Projeto ALiB foi realizada por equipes compostas por membros do Projeto, que foram treinadas e instruídas a utilizar os instrumentos adequados para ir a campo. O Comitê Nacional definiu um número máximo de 30 inquiridores para a equipe, garantindo, desta forma, homogeneidade na aplicação dos questionários na área pesquisada. A seleção para documentadores do Projeto aconteceu em 2002, durante o III WorkALiB. Mais tarde, houve a inclusão de novos inquiridores (RIBEIRO, 2012, p. 127).

A pesquisa de campo do ALiB foi feita por um inquiridor juntamente com um ou dois inquiridores auxiliares. Após preparação prévia e levantamento de dados e contatos da área em estudo, a equipe realizava seleção dos informantes, respeitando os critérios pré-estabelecidos pelo Projeto. Depois de selecionar os informantes participantes da pesquisa, os dados eram recolhidos *in loco* por meio de entrevista gravada em áudio digital. Era necessário também o preenchimento de fichas específicas: uma ficha com informações da localidade, para descrever a realidade da área pesquisada, e outra com dados do informante, com informações pessoais e observações comportamentais depreendidas ao longo das entrevistas, a fim de caracterizá-lo e estabelecer possíveis relações com seus usos linguísticos (CARDOSO et al., 2013, p. 41).

Ao findar as gravações, os arquivos eram identificados, digitalizados e armazenados em espaço próprio pelas equipes regionais responsáveis e, posteriormente, pela equipe de bolsistas e pesquisadores do Projeto ALiB, na Universidade Federal da Bahia.

Para a efetivação desta tese, foram utilizados os arquivos em áudio armazenados pelo Projeto ALiB, cuja audição dos questionários selecionados garantiu o levantamento e a transcrição grafemática dos dados.

3.6.2 Constituição do *corpus*

Para o levantamento dos dados, ouvimos todos os questionários (semântico-lexical, morfossintático, questões de prosódia, questões metalinguísticas, questões de pragmática, temas para discurso semidirigido), exceto o questionário fonético-fonológico e o texto para leitura, em um processo de busca das ocorrências de *tu* e *você* como tratamento do interlocutor nos inquéritos. Mesmo o material já tendo sido transcrita, precisávamos não só ter certeza da distinta realização de formas como *cê* e *você* que, por conta da semelhança prosódica na fala espontânea, poderiam ter sido transcritas de forma equivocada pelo transcritor, considerando a quantidade de dados de uma transcrição sociolinguística, como também fazer o acompanhamento das fitas com as transcrições, pois, em determinadas respostas, o documentador só registrava, precisamente, a resposta dada à pergunta.

Cabe salientar que, às vezes, o informante se utilizava de alguma história para responder a uma questão e, dessa forma, interessa-nos registrar não só a resposta como também toda a narrativa. A seguir, exemplificamos com excertos de fala obtidos como resposta à pergunta de nº 3, para identificação do gênero do substantivo, do questionário morfossintático:

INQ. – Como é que se chama aquela folha verde que se come geralmente na salada?

[Para preparar essa folha tem que ser bem lavada. Como pediria a alguém para lavar?]

INF. – Alfa... alface?

INQ. – É. A senhora chega lá na feira, como é que a senhora ia reclamar que tava com o preço muito caro?

INQ. – E, aqui, se a senhora fosse comprar, e a senhora achasse que tá muito caro.

Como é que a senhora ia reclamar?

INF. – “Mas moço, tá caro demais esse alface hoje”. (risos) Que, às vezes, eu digo.

(Risos) Eu digo: “Menino, de quanto tá esse alface?” Aí, ele diz assim: “Esse é de setenta e cinco. Tem de cinquenta, tem um real do (inint)”. “Ave Maria, mas eu queria desse aqui. Tá... faz de cinquenta não?” (risos) Eu só digo a ele. (risos) Aí, ele disse:

“Não, esse aqui...”, que o povo chama o alface manteiga. Ele é uma... como uma flor. Eu gosto muito dele. Não sei se **você** já viu. Aí, só que é mais caro ele do que aquele...

Em síntese, a pergunta feita tinha como objetivo a resposta do substantivo: “o alface” ou “a alface”, porém percebemos que, no evento de fala, há uma ocorrência de um pronome de tratamento, *você*.

Ressaltamos que em alguns momentos em que os dados foram catalogados, houve dificuldade em extraí-los do questionário, devido ao jogo de perguntas e respostas. Por outro lado, procuramos transcrever também as informações obtidas em comentários espontâneos feitos pelos informantes acerca do questionário.

Após o levantamento dos dados, ou seja, de identificar as 1.995 ocorrências pronominais de *tu* e *você* nas falas dos 140 informantes selecionados, distribuídos em 30 localidades do Nordeste, fizemos a codificação do material, considerando as variáveis controladas na pesquisa, como será detalhado a seguir.

3.7 AS VARIÁVEIS CONSIDERADAS

As investigações sociolinguísticas precisam dar conta da descrição detalhada das variantes que fazem parte do espectro da variação, seguindo o pressuposto de que a língua, mesmo heterogênea, revela variação regular e sistemática controlada por variáveis inerentes ao próprio sistema ou externas a ele.

Nesta seção, elencamos, portanto, além da variável dependente, os grupos de fatores de ordem linguística (estrutural e semântico-pragmático-discursiva) e extralinguística (social e/ou geográfica) controlados na pesquisa.

3.7.1 Variável dependente

Para o desenvolvimento desta tese, estabeleceu-se como variável dependente *a referência ao interlocutor*. Como tem sido demonstrado em diversos trabalhos sobre esse fenômeno, os falantes do PB dispõem de mais de uma forma para se dirigir à 2ª pessoa do discurso (cf. seção 1.3), contradizendo o paradigma tradicional (cf. seção 1.2). Vale destacar que esse fenômeno é denominado ‘dependente’ na literatura sociolinguística, tendo em vista que o uso das variantes – formas diferentes para se referir à mesma coisa, do ponto de vista do

sistema (LABOV, 2008 [1972]) – “não é aleatório, mas influenciado por grupos de fatores (ou variáveis independentes) de natureza social ou estrutural” (MOLLICA, 2003, p. 11).

Nessa perspectiva, consideram-se variantes as formas destacadas abaixo, exemplificadas com extratos da amostra.

- (i) Oh fulano, *você* quer um cafezinho com leite ou quer só o café puro? (Sergipe: 80/3)¹⁸
- (ii) J., *tu* toma leite ou café? (Pernambuco: 62/1)¹⁹

Depois de serem levantadas essas alternativas para a referência à 2^a pessoa do discurso (*tu* e *você*), procedeu-se à codificação das mesmas, a fim de averiguar os contextos linguísticos e sociais que favorecem o emprego de cada uma delas. Salientamos que as ocorrências da variante *cê* foram quantificadas juntamente com a forma *você*, uma vez que a primeira é uma redução fonética da segunda e a metodologia estatística adotada para a análise dos dados admite apenas rodadas binárias para a obtenção de pesos relativos.

Controlamos os fatores linguísticos (estruturais e pragmático-discursivos) e extralingüísticos (social e geográfico) descritos nas seções a seguir.

3.7.2 Variáveis independentes

As variáveis independentes consistem em conjuntos de fatores, tanto de natureza linguística como extralingüística, que não dependem de nenhum outro fator, mas podem exercer pressão sobre a variável dependente, aumentando ou diminuindo a frequência de uso das variantes que a constituem (MOLLICA, 2003). A observação atenta das variáveis independentes em uma pesquisa variacionista é de grande importância não somente para examinar quais fatores linguísticos e sociais favorecem ou inibem o uso de uma dada variante como também para garantir que seu uso nunca acontece de maneira aleatória.

18 Em relação ao numeral 80, indica a cidade de Estância. O informante 3 pertence à faixa etária 2, é do sexo masculino e possui o nível de ensino fundamental.

19 Quanto ao numeral 62, é a cidade de Exu. O informante 1 está incluso na faixa etária 1, é do sexo masculino e possui o nível de ensino fundamental.

3.7.2.1 Fatores linguístico-estruturais

Para a definição dos grupos de fatores linguístico-estruturais, partiu-se da ideia de que a opção pelo uso de uma das variantes poderia estar sendo motivada por fatores de ordem interna à estrutura da língua. Foram controlados, então, os seguintes grupos de fatores: referenciação, tipo (parte) do questionário, função sintática do pronome, tempo verbal e paralelismo. A seguir, apresentamos a caracterização de cada uma das variáveis linguístico-estruturais, seguidas das hipóteses aventadas na pesquisa, juntamente com exemplos extraídos da amostra analisada.

3.7.2.1.1 Referenciação

A referenciação *indeterminada* ocorre quando o falante se reporta a um interlocutor imaginário. Trata-se de uma seleção inconsciente, um enquadramento de um interlocutor indeterminado numa situação hipotética. Já na referenciação *específica*, o falante se dirige a um interlocutor como 2^a pessoa do discurso.

No que concerne às estratégias de indeterminação, é sabido que os pronomes pessoais funcionam como o recurso mais utilizado no PB (MENON, 2006). O pronome se caracteriza também pela dicotomia determinação/indeterminação, decorrente da sua especificidade semântica, que possibilita avançar além do nível frasal. Nesse sentido, a variável referenciação vem se destacando em estudos que tratam da alternância pronominal, apesar de muitas deles considerarem essa variável como estatisticamente não significativa (cf. DIAS, 2007; CALMON, 2010).

A fim de testar essa variável, aventamos como hipótese que para a referência genérica os usos mais recorrentes são os do pronome *você*, pois o falante, ao relatar um acontecimento, afasta-se do fato narrado e, consequentemente, do seu envolvimento direto com o discurso, priorizando a forma inovadora. Como exemplos, temos:

a) Referência específica

(Questões de prosódia)

INQ. – Pergunte a uma pessoa se ela vai sair hoje.

INF. – S., *você* vai sair hoje? (Pernambuco: 71/1)²⁰

b) Referência genérica

(Discurso semidirigido)

INF. – A cebola é assim: *você* vai molha, molha a terra todinha, deixa ela toda molhada, vai lá, pega semente, bota tudo dentro de uma caixa, não a semente a sementinha pequena da cebola, ela fica desse tamanho tipo capim, aí *você* pega faz os canteiros, os quadradinhos e suco, corta bem a terra, aí ali vai só com o dedo, pega uma mudinha e só com o dedo.

INQ. – Depois precisa molhar de novo?

INF. – Não, deixa ela pegar e depois pega a bomba de veneno e vai só bombando, venenando. (Pernambuco: 71/1)²¹

3.7.2.1.2 Parte do inquérito

Uma vez que o *corpus* da pesquisa é formado a partir do ALiB, resolvemos verificar como as formas de tratamento com o interlocutor – *tu* e *você* – se realizam nos questionários: QMS; discurso semidirigido; questões de prosódia; QSL; questões metalinguísticas; e questões de pragmática. A seguir, listamos alguns exemplos das ocorrências encontradas no *corpus*.

As expectativas com relação a essa variável eram as seguintes: nos discursos semidirigidos há maior utilização do *você*, especificamente o uso do *você* genérico, uma vez que os relatos pessoais e não pessoais tendem a narrar informações nas quais sempre aparecem um *você* indeterminado (genérico). Já nas questões de prosódia, pelo fato de a pergunta ser direcionada a um amigo, pessoa próxima, acredita-se que haja uma predominância no uso do *tu*.

a) QMS

INQ. – Quando você vê um amigo com uma mala e se quer saber para onde ele vai, como é que se pergunta?

20 O ponto do ALiB de número 71 é a cidade de Floresta. O informante 1 é do sexo masculino, faixa etária I e possui o nível de ensino fundamental.

21 O ponto do ALiB de número 71 é a cidade de Floresta. O informante 1 é do sexo masculino, faixa etária I e possui o nível de ensino fundamental.

INF. – “E aí? *Tu* vai pra onde?” (risos) (Paraíba: 56/2)²²

INQ. – Conhece alguma simpatia para tirar verruga?

INF. – Meu pai, meu avô me ensinava... dizia assim... a lua nova, é... mostrava a verruga pra lua e... contava quantas tinha e dizia que quando ela fosse cheia é... as berrugas teriam acabado. Uma outra é para os dente, nascia os dente, *você* rancava os dente e... aproveitava uma tarde no pô do só e jogava esse dente lá em cima da telha e... dizia algumas coisas assim, aí que dizia: “Morão, Morão, toma meu dente podre e me dá um sôa”. (Alagoas: 74/3)²³

b) Discurso semidirigido

INQ. – Eh... agora, eu quero que a senhora comente com a gente um acontecimento marcante da sua vida. Pode ser o seu casamento...

INF. – Vixe!

INQ. – ... nascimento do primeiro filho...

INF. – Ave Maria!

INQ. – O dia que sua filha disse que ia pra Natal.

INF. – Menina! Olha! (*Tu*) tá... ah, é... é de chorar as coisa, (porque) é muito emocionante, eu acho, que do marido... aff! Do casamento... (Paraíba: 56/4)²⁴

c) Questões de prosódia

INQ. – Se você quer saber se seu amigo toma leite ou café, como pergunta?

INF. – *Você* toma café ou leite? Qual é a sua preferência?? (Alagoas: 74/3)²⁵

d) QSL

INQ. – E.. às vezes tem gente que parece que fala pelo nariz...

INF. – Normalmente chama de fonfom, as pessoas aqui do interior... são linguagens do interior. Eu conheço as duas partes: capital e interior.

INF. – Olá! *Você* conhece aquele que é fonfom?

22 O ponto do ALiB de número 56 é a cidade de Cuité. O informante 2 é do sexo feminino, faixa etária I e do nível de ensino fundamental.

23 O ponto do ALiB de número 74 é a cidade de União dos Palmares. O informante 3 é do sexo masculino, faixa etária II e do nível de ensino fundamental.

24 O ponto do ALiB de número 56 é a cidade de Cuité. O informante 4 é do sexo feminino, faixa etária II e do nível de ensino fundamental.

25 O ponto do ALiB de número 74 é a cidade de União dos Palmares. O informante 3 é do sexo masculino, faixa etária II e do nível de ensino fundamental.

INF. – Você conhece a fonfom? Porque é fanhoso. (Alagoas: 74/3)²⁶

e) Questões metalinguísticas

INQ. – Tem gente que fala diferente aqui?

INF. – Fala, tem cara que fala estrangeiro. Tem gente que fala português, francês, inglês. Teve um cabra um dia desses namorou uma prima minha, disse: “Por que pra *tu* namorar esse cabra (inint), pelo menos falar, não sei nem o que ele tá falando” e ela chama Nanam, “que isso Nanam, *tu* tá doida? (inint) fala, rapaz, o que *tu* quer” e ele não tava entendendo nada. (Paraíba: 58/3)²⁷

f) Questões de pragmática

INQ. – Poderia dar exemplo do modo como falam essas pessoas? “*que falam diferente*”?

INF. – É, como assim... meu sogro chama “meiro irmão”. Eu não sei, assim, porque não chama logo “meu irmão”, né. Aí chama: “Oh, meiro irmão”. Isso aí é uma fala diferente da nossa, né.

INQ. – Hum hum.

INF. – *Tu* tá entendendo como é?

INQ. – Tô, tô entendendo. (Paraíba: 58/1)²⁸

3.7.2.1.3 Função sintática do pronome

Decidimos averiguar se a função sintática também interfere na realização dos pronomes de 2^a pessoa, variável que se mostrou relevante nos estudos de Modesto (2006), Mota (2008) e Andrade (2010).

Na análise desenvolvida por Guimarães (2014), são investigadas seis funções sintáticas dos referidos pronomes: função sujeito, função objeto direto, função objeto indireto, vocativo e complemento nominal. Todavia optou-se, nesta pesquisa, por controlar apenas as funções sujeito e não sujeito.

26 O ponto do ALiB de número 74 é a cidade de União dos Palmares. O informante 3 é do sexo masculino, faixa etária II e do nível de ensino fundamental.

27 O ponto do ALiB de número 58 é a cidade de Itaporanga. O informante 3 é do sexo masculino, faixa etária II e do nível de ensino fundamental.

28 O ponto do ALiB de número 58 é a cidade de Itaporanga. O informante 1 é do sexo masculino, faixa etária I e do nível de ensino fundamental.

Como hipótese, aventamos que, no *corpus* investigado, a função sintática de sujeito favorece a forma *você*. Ao analisar essa variável, Mota (2008) observa que ela se mostra estatisticamente relevante no uso da forma *tu*, sendo esse pronome favorecido pela função de objeto, com peso relativo 0,91.

A seguir, são exemplificados (com dados do *corpus* analisado) os fatores que constituem esse grupo:

a) Sujeito

INF. – R., tô sabendo que *você* é a terrorista da serra, que *você* é brava. (Pernambuco: 64/2)²⁹

INF. – (...) tô sabendo que *tu* tá aí comemorando a facada que *tu* disseste que vai dar em meu filho... (Pernambuco: 64/2)³⁰

b) Não sujeito

INF. – F., saia da minha casa hoje porque eu não aceito mais *você* aqui. (Pernambuco: 64/4)

INF. – Eu disse: olha, o que *tu* tem a ver com a nossa vida? (Pernambuco: 64/4)³¹

3.7.2.1.4 Tempo verbal

Almeida e Assunção (2008) analisam as falas de feirenses analfabetos funcionais e, dentre os grupos de fatores controlados, verificaram que o tempo verbal foi considerado relevante, sendo o modo indicativo o favorecedor do uso do *tu*. Por sua vez, no trabalho de Alves (2015), a realização das rodadas binárias possibilitou constatar que a variável tempo verbal não exerce influência na escolha do pronome *você*, resultado esperado, considerando que esse pronome vem seguido de formas verbais típicas de terceira pessoa.

Acreditamos, nesta pesquisa, que as sentenças com verbos no passado tendem a favorecer o emprego do *você*.

A seguir, encontram-se listados e exemplificados os fatores analisados neste grupo – o tempo passado e o tempo não passado:

29 Limoeiro é a cidade representada pelo ponto 64. O informante 2 é do sexo feminino, da faixa etária I e do nível de ensino fundamental.

30 O ponto do ALiB de número 64 é a cidade de Limoeiro. O informante 2 é do sexo feminino, da faixa etária I e do nível de ensino fundamental.

31 O ponto do ALiB de número 64 é a cidade de Limoeiro. O informante 4 é do sexo feminino, da faixa etária II e do nível de ensino fundamental.

a) Passado

INF. – Começo da noite, como *você* falou assim, né? Ou então, noite mermo. (Sergipe: 80/2)³²

b) Não passado

INF. – Meu filho, saia daí que *você* vai se machucar! (Sergipe: 80/2)³³

3.7.2.1.5 Paralelismo

A variável paralelismo foi apontada em alguns estudos como o grupo de fatores linguísticos de maior relevância nas pesquisas variacionistas de alternância pronominal, dentre os quais estão Santos (2012), Martins (2010), Zilli (2009) e Oliveira (2005; 2007).

O conceito de paralelismo pode ser firmado com base em Oliveira (2009, p. 7), que destaca que “o paralelismo formal indica se o antecedente mais próximo influencia de alguma forma a escolha de uma das variantes”. Essa variável consiste em apontar se os pronomes de segunda pessoa tendem a se repetir durante a fala no mesmo turno.

Para codificar essa variável, consideramos, numa série discursiva, as respostas a uma pergunta até que esta fosse interrompida por uma nova pergunta. Não foi considerado como encerramento de um turno o momento em que o inquiridor se pronuncia apenas como sinal de entendimento (cf. DEUS, 2009).

Segundo Scherre (1998),

A própria repetição das variantes de uma mesma variável dependente no discurso tem se evidenciado como uma restrição importante na análise de fenômenos variáveis de todos os subsistemas linguísticos em diversas línguas. Esta restrição ou variável independente ocorre entre as cláusulas (plano discursivo), no interior da oração (plano oracional), no interior do sintagma (plano sintagmático) e entre palavras e no interior da palavra (plano da palavra). Recebendo denominações diferenciadas dentro da literatura variacionista, ela é hoje bastante conhecida como paralelismo linguístico. (SCHERRE, 1998, p. 30)

A hipótese aventada neste estudo é a de que o princípio do paralelismo atua na escolha das variantes e o uso de uma forma – *tu* ou *você* – ou seus correlatos oblíquos e possessivos, no início da sequência discursiva, desencadeia uma série de repetições da mesma forma.

32 Informante da cidade de Estância, da faixa etária I, do sexo feminino, de nível de escolaridade fundamental.

33 Informante da cidade de Estância, da faixa etária I, do sexo feminino, de nível de escolaridade fundamental.

Como essa variável se apresenta muito relevante em diversos estudos, foi feito o seu controle por meio dos seguintes fatores:

a) *Tu precedido de tu / Você precedido de você*

INF. – ... Então *você* separa... separa ambientes numa ... numa propriedade, num terreno... *você* pode até fazer um tipo galinheiro... Botar o dedo pra lu... pra estrela e mostrar o local, então *você*... (Pernambuco: 70/8)³⁴

b) *Tu precedido de você / Você precedido de tu*

INF. – Então chama: ei cara, dei..., deixasse ô deixaste caí, *você* deixou ou ti ô tu deixar... *Tu* deixou... (Pernambuco: 70/8)³⁵

c) Primeira ocorrência da série (primeira forma que aparece no discurso)

INF. – ... Então *você* separa... separa ambientes numa ... numa propriedade, num terreno... *você* pode até fazer um tipo galinheiro... Botar o dedo pra lu... pra estrela e mostrar o local, então *você* botava a verruga e depois procurava (risos). Pra a verruga *você* iria, é ... botava isso pequeno, na estrela, na estrela cadente. (Pernambuco: 70/8)³⁶

d) Ocorrência isolada / realização única

INF. - Você toma café ou leite? Ou os dois? (Pernambuco: 70/8)³⁷

3.7.2.2 Fatores extralingüísticos

Com base nos princípios da Sociolinguística e da Dialetologia, verificam-se as variantes documentadas, tendo em vista buscar a seleção de formas utilizadas pelos informantes e as características sociais de que elas se revestem. Assim, são considerados os aspectos diatópico, diassexual, diageracional e diastrático, analisando o que as formas selecionadas pelos informantes indicam em relação a estas variáveis. O Projeto ALiB, ao qual se vincula esta

34 O ponto 70 é a cidade de Recife. O informante 8 é do sexo feminino, da faixa etária II e tem nível de ensino universitário.

35 I O ponto 70 é a cidade de Recife. O informante 8 é do sexo feminino, da faixa etária II e tem nível de ensino universitário.

36 O ponto 70 é a cidade de Recife. O informante 8 é do sexo feminino, da faixa etária II e tem nível de ensino universitário.

37 O ponto 70 é a cidade de Recife. O informante 8 é do sexo feminino, da faixa etária II e tem nível de ensino universitário.

pesquisa, contempla 250 localidades brasileiras, com informantes de ambos os sexos e de duas faixas etárias: 1 – 18-30 anos e 2 – 50-65 anos. Do ponto de vista diastrático, nas capitais de Estados, o estudo contempla informantes de nível universitário e de nível fundamental; nas demais cidades do interior, apenas informantes de nível fundamental, como será melhor detalhado nas seções a seguir.

3.7.2.2.1 Localidade

Propaga-se que a região Nordeste é exemplo de unidade em vários aspectos, porém a região é plural e apresenta variação em muitos aspectos linguísticos. Prova disso é a divisão dialetal proposta por Antenor Nascentes (1953), que leva em consideração não as fronteiras geográficas, mas as linguísticas, como pode ser observado na Figura 11. Embora a caracterização do país em zonas linguísticas seja feita a partir de aspectos fonético-fonológicos, trabalhos de outras áreas utilizam a mesma divisão como base para conhecer as características da região.

Figura 11: Divisão dialetal proposta por Nascentes



Fonte: Nascentes (1953, p. 18).

Com o objetivo de identificar a influência da origem geográfica do falante no uso das formas de referência à segunda pessoa, as ocorrências foram analisadas segundo a localidade dos informantes. Assim, procuramos examinar a variação diatópica concernente à aplicação da regra variável em cinco Estados do Nordeste brasileiro: a) Sergipe; b) Alagoas; c) Rio Grande do Norte; d) Paraíba; e e) Pernambuco. Como hipótese, admite-se que o pronome *tu* tem predominância no interior e o *você* é mais utilizado nas capitais.

3.7.2.2.2 Sexo

A divergência no comportamento linguístico dentro de uma mesma comunidade entre os falantes do sexo masculino e os do sexo feminino levou os sociolinguistas a admitir uma possível diferença nos usos desses falantes, uma vez que a linguagem é o resultado da cultura em que estamos inseridos.

Labov (2008 [1972]) aponta o fato de que as mulheres, em situação de variação estável, têm demonstrado preferência pelo uso das formas de prestígio e que, em caso de mudança linguística, seriam inovadoras e responsáveis pela propagação da variante não-padrão. Assim, considerando as informações apresentadas, decidimos analisar o comportamento desta variável no *corpus*, controlando os fatores: a) Masculino e b) Feminino.

A nossa hipótese em relação à interferência da variável sexo na realização do *tu/você* segue o proposto por Labov (2008 [1972]). Assim, assumimos que as mulheres preferem a variante de prestígio *você* em detrimento do *tu*, como forma de afirmação e motivadas pelo discurso com bases feministas que debatem o uso de determinadas formas linguísticas como valorização do papel social das mulheres.

3.7.2.2.3 Faixa etária

O *corpus* desta pesquisa possibilita para análise da variável idade o estudo de duas faixas etárias: a faixa etária I, que agrupa os indivíduos mais jovens, entre 18 e 30 anos, e a faixa etária II, que conta com indivíduos entre 50 e 65 anos.

A faixa etária é compreendida pelos estudos variacionistas como um grupo de fatores importante, uma vez que pode dar indícios de uma mudança linguística em progresso. Em uma pesquisa feita em três lojas de departamento de Nova Iorque, Labov (2008 [1972]) aponta que os jovens preferem as marcas de prestígio. “Enquanto os de menos de 40 anos mostram um padrão bastante confuso em sua reação subjetiva a (r), os que têm entre 18 e 39 anos mostram uma completa unanimidade em sua avaliação positiva deste traço de prestígio” (LABOV, 2008 [1972], p. 164 -165).

Araújo (2007), em seu artigo sobre a variável faixa etária nos trabalhos sociolinguísticos, chegou à conclusão de que, quando a variável é um fenômeno em uma possível mudança em curso, ela é mais recorrente na fala dos jovens do que na fala dos idosos; já quando se trata de uma variável estável, os mais jovens e os mais velhos apresentam frequências mais altas da variante inovadora; e os medianos (devido às pressões de mercado etc.) apresentam uma frequência significativamente mais baixa.

Admitimos a hipótese de que os mais jovens utilizam mais o pronome *você* e os mais velhos o *tu*. O acesso aos meios de comunicação, cuja linguagem enfatiza o uso do *você*, e os recursos disponíveis na Internet, em que este pronome, grafado em geral como “vc”, parece ter

preferência na linguagem dos jovens, induziram-nos a considerar que esta faixa etária apresente um maior uso do *você*.

3.7.2.2.4 Escolaridade

Segundo Mollica (2003, p. 51), “a escola atua como preservadora de formas de prestígio, face às tendências de mudança em curso”. Nesse sentido, faz-se necessário validar ou não se o nível de escolaridade do informante influencia na escolha entre os pronomes *tu* e *você*.

Quanto às pesquisas sobre a alternância dos pronomes de 2^a pessoa, a escolaridade é considerada um grupo de fatores importante, sendo selecionado em muitos trabalhos de orientação quantitativa, destacando-se como mais relevante nos estudos de Modesto (2006), Zilli (2009), Franceschini (2011) e Santos (2012).

Como hipótese, acreditamos que os informantes com maior escolaridade tendem a falar mais o *você* do que informantes com um grau de escolaridade menor, porque este encerra um nível maior de formalidade, além de ser a variante mais frequente nas normas urbanas de maior prestígio da maioria das capitais brasileiras.

3.8 PROCEDIMENTOS PARA A ANÁLISE DOS DADOS

Nesta seção, apresentamos os procedimentos adotados para a análise dos dados, apresentada no capítulo seguinte.

3.8.1 Codificação dos dados

Para proceder à análise estatística no programa GoldVarb X (versão para ambiente Windows do pacote de programas Varbrul³⁸), o nosso primeiro passo foi atribuir códigos para a variável dependente e para as variáveis independentes selecionadas para análise. O pesquisador deve escolher um código que identifica cada fator analisado; esses códigos representam os fatores que caracterizam, identificam cada dado. De acordo com Guy e Zilles (2007, p. 123), “esses códigos podem ser letras, números e outros símbolos de teclado”, com exceção de alguns símbolos que têm significado específico para o programa, como, por

38 Do inglês *Variable Rule Analysis*.

exemplo, o sinal do parêntese aberto, que sinaliza o início de uma codificação, o espaço em branco, entre outros.

A chave de codificação estabelecida para esta tese está exposta no Anexo B. Buscamos lidar com códigos que fossem motivados para facilitar a codificação e não utilizamos códigos como a letra 1 e o numeral 1 ou a letra O e o numeral 0 no mesmo grupo, o que poderia dificultar futuramente o processo de correção e a leitura dos dados.

Os arquivos codificados foram salvos com a extensão .tkn, específica para leitura pela ferramenta computacional, podendo, então, ser submetidos à análise estatística. De acordo com Guy e Zilles (2007),

O Varbrul é um conjunto de programas computacionais de análise multivariada, especificamente estruturado para acomodar dados de variação sociolinguística. A análise se chama ‘multivariada’ porque permite investigar situações em que a variável linguística em estudo é influenciada por vários elementos do contexto, ou seja, múltiplas variáveis independentes. A investigação mede os efeitos, bem como a significância dos efeitos, dessas variáveis independentes sobre a ocorrência das realizações da variável que está sendo tratada como dependente. (GUY; ZILLES, 2007, p. 105)

O programa GoldVarb X, versão mais atualizada do Varbrul, é bastante utilizado pelos pesquisadores variacionistas, uma vez que é através dele que alcançamos resultados mais precisos e confiáveis nos estudos descritivos da língua. O uso dessa ferramenta nos permite fazer uma análise mais adequada da alternância das formas de tratamento pronominais *tu/você*, definindo “a natureza e a extensão de cada um dos efeitos condicionadores, permitindo ‘predizer’ probabilisticamente a taxa aproximada [do uso dos pronomes] por certo indivíduo” (GUY; ZILLES, 2007, p. 49).

Nessa perspectiva, submetidos os dados ao programa, foram geradas as frequências absolutas e relativas de ocorrência da variável dependente para cada grupo de fatores selecionado. Trataremos, na próxima seção, dos procedimentos para a análise estatística dos dados.

3.8.2 Aspectos da análise estatística

Reiteramos que a nossa pesquisa analisa o comportamento de um fenômeno variável e que fatores linguísticos e sociais podem condicioná-lo. Assumimos que essa variação não é aleatória, e, sim, regulada, governada por um conjunto de regras.

Naro (2003, p. 16) argumenta que a questão central para a Teoria da Variação é a avaliação do *quantum* com que cada categoria postulada contribui para a realização de uma ou de outra variante que está em competição. É impossível isolar e medir separadamente o efeito de um fator, mesmo porque quando tal fator aparece nos dados, isto é, no uso real da língua, ou seja, o mesmo nunca se apresenta isoladamente.

A Sociolinguística laboviana, também conhecida como Teoria da Variação, apresenta como uma de suas principais características a quantificação das ocorrências de cada variante em relação às variáveis analisadas. Desse modo, os princípios da Sociolinguística preconizam que as línguas são inherentemente dinâmicas e heterogêneas, isto é, variáveis. O maior objeto de estudo da Sociolinguística é a variação. Entendemos que a variação pressupõe que formas linguísticas variantes estejam coexistindo dinamicamente.

Para que possamos formular esquemas de regras, é necessário que se desenvolva uma metodologia adequada para quantificar os fatores, em número relativamente pequeno, que devem apresentar um peso fixo, independentemente do contexto em que ocorrem.

Segundo Guy e Zilles (2007),

na análise quantitativa, cada fator recebe um ‘peso’ (‘valor do fator’ ou ‘probabilidade de *input*’ (‘Po’), que é, de fato, uma medida global do índice de aplicação da regra. Uma função matemática é usada para combinar os valores dos fatores e a probabilidade de *input*, a fim de produzir os índices esperados de aplicação da regra em cada célula. Diversas funções foram usadas na literatura; a que é favorecida atualmente é a ‘função logística’. (GUY; ZILLES, 2007, p. 51)

A versão inicial do *software* Varbrul data de 1988, desenvolvida para funcionar no antigo sistema MS-DOS³⁹. Mais tarde, o pacote sofreu alterações devido aos avanços da tecnologia computacional. Por vezes, os pesquisadores da área deixaram de utilizar-se de programas como SPSS⁴⁰ e deram preferência a esse novo programa, que foi modelado para verificação de regras variáveis em Linguística, podendo também ser utilizado por estudiosos de outras áreas.

Por meio desse *software*, é possível apontar quais as variáveis independentes que condicionam ou não a escolha das variantes dependentes, favorecendo ou não um determinado processo variável, correspondendo, assim, ao que foi mencionado por Guy e Zilles (2007, p. 106): “o Varbrul permite a construção de um modelo completo e específico dos processos e

39 Do inglês *Microsoft Disk Operating System*.

40 Do inglês *Statistical Package for the Social Sciences* (Pacote Estatístico para as Ciências Sociais).

efeitos, faz os testes de significância etc., que não aparecem num simples cálculo de percentagens”.

Para realizar uma análise estatística, inicialmente, deve-se submeter o arquivo de dados ao *Check Tokens*, que verifica se há erros cometidos na fase de codificação, os quais devem ser corrigidos pelo pesquisador. Depois de conferido o arquivo com os dados codificados e geradas as variáveis específicas para a análise, o pesquisador deve inserir o valor de aplicação segundo a hipótese definida em sua pesquisa. É gerado, então, um arquivo com as porcentagens de cada variável independente com relação à variável dependente.

O programa é uma ferramenta essencial no tratamento quantitativo dos dados, emitindo resultados estatísticos e probabilísticos confiáveis, no entanto a interpretação dos resultados fornecidos é de responsabilidade do pesquisador, que deve fazer uso do programa apenas como um instrumento para chegar às análises definitivas. Por esse ângulo, aludimos ao posicionamento de Scherre e Naro (2010, p. 162), quando afirmam que “os resultados numéricos obtidos pelos programas só têm valor estatístico. O seu valor linguístico é atribuído e interpretado pelo linguista”. Assim, de nada valem os dados se o linguista não souber interpretá-los, até porque a estatística é apenas um instrumento valioso que ajuda a compreender, de forma mais eficaz, o comportamento de alguns fenômenos linguísticos.

O programa revela valores percentuais de ocorrência das variantes em observação, em função de cada uma das variáveis independentes e seus respectivos fatores, controlados na análise, bem como fornece os pesos relativos dos fatores favorecedores da variação, o que corresponde ao seu grau de influência na ocorrência da variável analisada.

3.8.3 Identificação de *nocautes*

No decorrer do processo de análise estatística dos dados levantados nesta pesquisa, o programa GoldVarb X apontou para a ausência de variação, indicada pelo programa com a terminologia *nocaute*, fator que, num dado momento da análise, corresponde a uma frequência de 0% ou de 100% de ocorrência em uma das variantes da variável dependente. Não havendo variação, há um problema na análise, pois o programa não tem como computar pesos e frequências.

Outro problema é o denominado *singleton group*, que consiste na identificação de apenas um fator em uma variável independente. Quando ocorre algum desses problemas –

nocaute ou *singleton group* –, é necessário adotar procedimentos em relação aos fatores controlados, podendo uni-los a outros ou excluí-los da análise, para que o programa avance.

Em tais situações, o pesquisador deve averiguar se eles resultam da pequena quantidade de ocorrências ou se são nocautes verdadeiros. Se esta última possibilidade for a escolhida, os *nocautes* são excluídos das rodadas seguintes, uma vez que “o instrumento quantitativo que produz pesos relativos é apropriado exclusivamente para fenômenos linguísticos variáveis” (SCHERRE; NARO, 2010, p. 152). Por outro lado, apesar de serem retirados, devem ser analisados pelo pesquisador, pois auxiliam na compreensão do fenômeno variável e podem dar indícios de “um início ou fim de uma mudança linguística, ou a especialização de significância referencial, de funções discursivas ou de funções pragmáticas” (SCHERRE; NARO, 2010, p. 153).

Em referência ao *nocaute* aparente, Guy e Zilles (2007, p. 160) revelam duas formas de eliminá-lo e, ao mesmo tempo, reter os dados: amalgamar o *nocaute* a outros fatores ou não o usar. Conforme os autores, quando os dados dos *nocautes* estão isolados num fator com poucos dados, a melhor opção é combinar tal fator com outro que não seja *nocaute*. O grande desafio é decidir com que fator o *nocaute* será amalgamado, para que haja uma lógica em sua combinação.

Os autores também asseguram que, se forem muitos os dados com *nocaute*, o ideal é que o pesquisador pense na possibilidade de não os retirar pois isso poderá afetar os pesos relativos. A segunda forma de eliminar os *nocautes* é recodificá-los como ‘não aplicável’, usando o sinal da barra (/). Segundo Guy e Zilles (2007, p. 161), essa não é a forma preferível de tratar o *nocaute*. Os autores apontam casos em que isso pode ocorrer: “quando não existe nenhuma amalgamação com outro fator do mesmo grupo que seja plausível em termos teóricos; há muito pouco dados; a exclusão de fatores com *nocautes* permite ainda assim que se obtenha um grupo apropriado, talvez com outro título”.

Assim, fica a critério do pesquisador determinar se os *nocautes* devem ser excluídos ou amalgamados, prosseguindo, portanto, a análise para a obtenção da seleção das variáveis relevantes para a regra de aplicação e dos pesos relativos referentes a cada um dos fatores.

3.8.4 Peso relativo

Para uma análise estatística, são fundamentais os valores percentuais, todavia é indicado que façamos a sua leitura junto com a dos pesos relativos, que estabelecem uma significância

estatística para a ocorrência de uma determinada variante. Nessa sequência, o valor do peso relativo oferece ao pesquisador uma informação estatisticamente mais precisa sobre o fenômeno, visto que trabalhar apenas com os percentuais pode acarretar em atribuir uma informação duvidosa ao objeto de estudo (GUY; ZILLES, 2007). A esse respeito, os autores revelam que:

Quando há vários fatores contextuais diferentes afetando uma variável linguística, uma série de tabelas que apresente esses efeitos separadamente [...] pode facilmente dar resultados distorcidos ou até enganadores, se os dados não estiverem uniformemente distribuídos por todas as variáveis independentes. Uma vez que a distribuição dos dados linguísticos geralmente é, de fato, desigual, uma análise multivariada dará resultados mais preciosos, porque, ao mesmo tempo em que computa o efeito de uma variável independente, ela controla explicitamente o efeito de todas as outras variáveis independentes conhecidas. (GUY; ZILLES, 2007, p. 34)

O programa indica qual foi a melhor rodada, em nível *step up* (aquele que considera todos os grupos de fatores que foram selecionados). Depois, verificamos o resultado do processo no *step down* (nível em que são eliminados os fatores não relevantes para a variação analisada).

Em relação à interpretação dos pesos relativos gerados a partir da análise binária, Scherre e Naro (2010), bem como Guy e Zilles (2007), afirmam que, quando os pesos relativos são próximos de 1,0, interpretamos como favorecedores da aplicação da regra em relação ao fenômeno estudado; quando são próximos de 0,5, concluímos que são neutros em relação à aplicação da regra; e, quando são próximos de 0,0, são considerados desfavorecedores em relação à aplicação da regra.

Chamamos a atenção para outro valor também importante para a análise dos resultados apresentados pelo GoldVarb X, o *log likelihood*, que é o logaritmo de verossimilhança que mede a qualidade da aproximação entre o modelo apresentado pelos pesos relativos e os dados observados. Quanto maior o *log likelihood*, ou seja, quanto mais próximo a zero, maior será a robustez dos dados (GUY; ZILLES, 2007).

Como mencionado anteriormente, o programa se baseia em um modelo de regressão logística. Mesmo que os valores de cada variável sejam calculados simultaneamente ao de todas as outras, eles apresentam efeitos independentes. Além disso, os pesos obtidos para um fator, em um grupo, são calculados com relação aos outros fatores desse mesmo grupo, como estão apresentados no Capítulo 4 desta tese.

3.9 CARTOGRAFIA DOS DADOS

Entende-se que a cartografia é arte e ciência. Nesse sentido, o dicionário Houaiss (2001), apresenta o verbete ‘cartografia’ com duas acepções: “(1) conjunto de estudos e operações científicas, técnicas e artísticas que orienta os trabalhos de elaboração de cartas geográficas; (2) descrição ou tratado sobre mapas”. Por outro lado, a ABNT (Associação Brasileira de Normas Técnicas) conceitua cartografia como a arte de levantamento, construção e edição de mapas e cartas de qualquer natureza (TELES; RIBEIRO, 2006, p. 210).

Teles e Ribeiro (2006, p. 210) reiteram que o termo ‘cartografia’ foi criado em 1839 pelo historiador português Manuel Francisco de Barros e Sousa (1791-1856), o Visconde de Santarém, que enviou uma carta a Francisco Adolfo Varnhagem (historiador brasileiro), indicando a ideia de um traçado de mapas e cartas, na qual se lia: “... invento esta palavra já que aí se têm inventado tantas”.

Como descrito por Romano et al. (2014, p. 123), em geral, observa-se, no cenário nacional, que a etapa de cartografia linguística, passo importante para a construção de um atlas linguístico, é realizada, em sua maioria, por profissionais ou da área da Geografia, que possuem conhecimentos específicos de cartografia e de SIG (Sistemas de Informação Geográfica), ou por designers gráficos, pessoas habilitadas a trabalhar com *softwares* de edição de imagens. Na maioria das vezes, é muito difícil o próprio linguista produzir suas cartas devido a dois motivos fundamentais: (i) pelo grande conjunto de dados a serem analisados e estudados, o que requer amplo investimento de tempo; e (ii) pela falta de conhecimento computacional de *softwares* de edição e geração de imagens.

A cartografia linguística exige requisitos que vão além das habilidades técnicas. São necessários conhecimentos intelectuais próprios daqueles que trabalham nesse campo de estudo, uma vez que, além de saber como representar, é necessário saber o que representar, quais variantes são válidas, qual extensão de legenda, a natureza da carta (léxica, fonética, morfossintática, sintática, semântica, neológica, isoglóssica), o tipo de representação, entre outros fatores, que o “não linguista”, em geral, não comprehende e, às vezes, ignora por priorizar aspectos estéticos ou mesmo conceituais de outras áreas do saber. (ROMANO et al., 2014, p. 123)

Enfim, buscamos, através da demonstração das cartas, com indicação exclusiva da variação diatópica (ver itens 4.2.1.4 e 4.2.2.3), fornecer uma imagem seletiva e coerente da variação pronominal nas localidades analisadas. O processo cartográfico resultou da junção dos dados e informações das áreas (rede de pontos) que perfizeram a amostra.

Algumas convenções são de extrema importância para que a expressividade, a legibilidade e a eficácia sejam conquistadas. Na legenda há explicações das convenções utilizadas, normalmente dentro de uma moldura, na parte inferior da carta. Contemplam-se também valores percentuais nas cartas.

Outras informações importantes são acrescidas nas cartas, na tentativa de elucidar, agregar e esclarecer: são as informações marginais. As mais utilizadas são: o título, a projeção usada e os dados a ela associados, fontes das informações usadas na execução, data e escalas (gráfica e numérica) (TELES; RIBEIRO, 2006, p. 223).

A Comissão de cartografia, revestida de sua característica principal, que é a interdisciplinaridade, vem tratando da informatização e cartografia dos dados do ALiB desde 2006. Agora, em 2020, sofreu algumas alterações, inserindo novos membros, devido à morte da professora Ana Teles.

Quadro 11: Comissão de informatização e cartografia

| 2006 | 2020 |
|---------------------------------------|----------------------------|
| Silvana Ribeiro (UFBA) | Silvana Ribeiro (UFBA) |
| Abdelhak Razky (UFPA) | Abdelhak Razky (UFPA) |
| Cléo Altenhofen (UFRGS, saiu em 2013) | Daniela Claro (UFBA, 2007) |
| Ana Regina Teles (UFBA, †2020) | Valter Romano (UFSC, 2014) |

Fonte: Mota (2020, apresentação de *slides* em *live*: 1:15).

As cartas linguísticas 1 e 2, apresentadas na seção de análise dos dados, propiciam a representação espacial das formas de tratamento *tu* e *você* em alguns Estados do Nordeste; as mesmas expressam graficamente informações quantitativas e qualitativas. Permitem também que “sejam feitas análises sobre o tema (ou assunto) cartografado e que estas possam ser relacionadas a outras informações obtidas através de outros instrumentos” (TELES; RIBEIRO, 2006, p. 212), no nosso caso, as tabelas que estão exibidas na mesma seção.

Para finalizar, reportamo-nos à memória desse recurso: a cartografia. Nos anos 1960, o desenho era feito com a ajuda do computador; mais adiante, nos anos 1970, as informações nos mapas e cartas eram registradas através de números. Já nos anos 1980,

[...] todos os domínios da Cartografia abriram-se irremediavelmente à Informática: surgiram métodos matemáticos e estatísticos para o tratamento das informações geográficas contidas nos mapas em termos de análise espacial e não-espacial. A cartografia automatizada tornou-se um verdadeiro sistema de informação geográfica, compreendendo seus próprios bancos de dados, suportes lógicos de aplicação e equipamentos. Os SIG (Sistemas de

Informações Geográficas) revolucionaram os conceitos e as técnicas de pesquisa, de análise e de manipulação dos dados. (TELES; RIBEIRO, 2006, p. 213)

Inferimos, portanto, que o processo cartográfico vem avançando bastante nas pesquisas do ALiB. O Comitê Nacional do Projeto tem dedicado esforços, como rapidez das respostas, novos recursos utilizados, capacidade de armazenamento dos dados e novas formas de apresentação, dentre outros, para melhor servir à pesquisa.

Salientamos ainda que o banco de dados do ALiB tem alimentado o ALiBWeb, sistema (em desenvolvimento) que permite gerenciar os dados armazenados, o qual possibilita consulta aos dados e elaboração de cartas em formato dinâmico e a fala do entrevistado poderá ser ouvida, com a geração de mapas sonoros. O sistema também admite que algumas cartas passem a ser públicas e outras não bem como permite a criação de novas tabelas referentes à análise linguística e às cartas linguísticas.

3.10 SÍNTESE DO CAPÍTULO

Explicamos, neste capítulo, os procedimentos metodológicos que nortearam, direta e indiretamente, a análise dos dados, cuja amostra de fala pertence ao Projeto ALiB, o qual estuda a realidade linguística do português brasileiro em seus mais diversos níveis: fonético-fonológico, morfossintático e léxico-semântico, enfocando fatores sociais e a perspectiva diatópica.

O referido Projeto utiliza como rede de pontos um total de 250 localidades, recobrindo uma extensão nacional do Oiapoque ao Chuí, selecionadas de acordo com critérios demográficos, históricos, culturais e de natureza do povoamento. Nessa perspectiva, o ALiB recobre toda a área do território brasileiro e, para este trabalho, elegemos a região Nordeste do país como *locus*, especificamente cinco Estados, perfazendo um total de trinta localidades.

Apresentamos os critérios seguidos na pesquisa de campo bem como na seleção dos informantes que constituíram a amostra. Relatamos a respeito dos questionários utilizados, os quais foram organizados pelos membros do Comitê Nacional do ALiB.

A partir de orientação da Sociolinguística Quantitativa, revelamos a importância do pacote computacional GoldVarb X para a quantificação dos dados deste estudo, uma vez que o mesmo aponta valores globais e percentuais de ocorrência das variantes em observação, além dos pesos relativos dos fatores das variáveis consideradas. Por outro lado, as orientações da

Dialectologia serviram como premissa para a organização de cartas linguísticas, as quais propiciam a representação geográfica de fenômenos em variação.

O capítulo seguinte é destinado à análise dos resultados, no qual descreveremos como os fatores selecionados atuam no comportamento dos falantes com relação ao uso das variantes em análise bem como se as hipóteses consideradas para a pesquisa foram comprovadas ou refutadas. A análise de cada um desses fatores levou em consideração os resultados estatísticos e os percentuais de ocorrência gerados pelo GoldVarb X e é acompanhada de um olhar qualitativo dos dados.

4 ANÁLISE DOS DADOS

Neste capítulo, são interpretados os resultados obtidos em análises estatísticas e probabilísticas gerados a partir do pacote GoldVarb X. Foram feitas duas análises: uma geral, incluindo todos os Estados estudados e outra de cada Estado, separadamente. Em cada análise, foram feitas duas rodadas no GoldVarb X: a primeira, que forneceu a frequência das variantes na amostra e a distribuição das mesmas em função das localidades; e a segunda, a rodada binária⁴¹, após serem solucionados os casos de nocautes⁴², que gerou os pesos relativos e a ordem de relevância de cada variável (ou grupo de fatores) na ocorrência das variantes do fenômeno. Esses resultados serão aqui interpretados, seguidos de uma análise qualitativa, na qual se faz uma abordagem crítica dos valores obtidos mediante a submissão dos dados ao programa estatístico. Vale destacar que no Capítulo 3, apresentamos a metodologia utilizada na pesquisa e foram feitos comentários acerca do manuseio deste programa computacional.

Também nesta seção, fazemos algumas considerações sobre os fatores que não foram selecionados pelo programa, mas que, mesmo descartados, são importantes para a compreensão geral tanto da análise quanto da variação em estudo. É importante assinalar que a variante escolhida com o valor de aplicação para análise estatística dos dados foi o pronome *tu*. Em relação à rodada com os cinco Estados, tivemos *input* de 0,058, *log likelihood* de -509.482 e *significance* de 0,001.

4.1 DISTRIBUIÇÃO GERAL DAS VARIANTES NA AMOSTRA

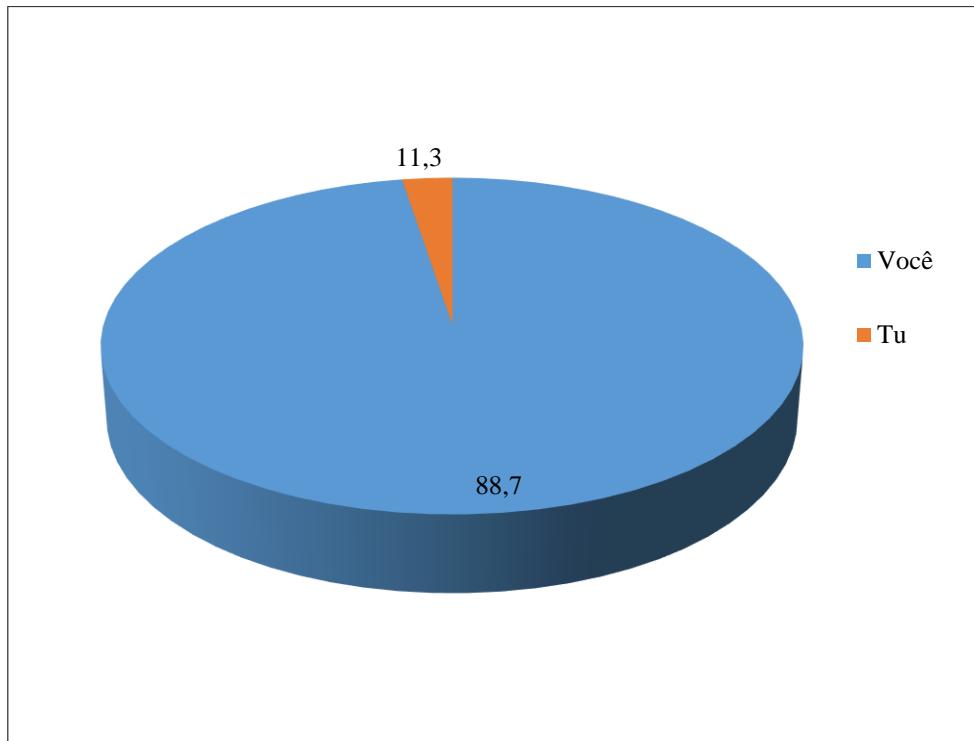
Inicialmente, como foi mencionado anteriormente, exibimos um panorama das formas de referência documentadas, mediante as variáveis pré-estabelecidas nos cinco Estados do Nordeste brasileiro. Após a apresentação dos resultados iniciais, passamos a comentar sobre a importância ou não de cada variável controlada para a seleção das formas de tratamento com referência à 2^a pessoa nos Estados investigados.

A Figura 12 apresenta o resultado geral das ocorrências registradas nesta primeira análise.

41 A rodada binária não foi feita em todos os Estados, devido ao grande número de nocautes, não sendo possível obter o peso relativo de algumas variáveis (ou fatores).

42 Ocorre nocaute (*knockout*) sempre que uma variante é usada em 100% dos dados, o que significa que a outra variante não ocorre, portanto não há variação. Trata-se, pois, de regra categórica.

Figura 12: Distribuição geral do *tu* e *você* nos Estados do Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e Sergipe



Fonte: Elaborada pela autora.

A Figura 12 mostra que, de um total de 1.995 ocorrências de pronomes de segunda pessoa analisadas, o *você*⁴³ foi empregado em 1.769 ocorrências, correspondendo a 88,7%, enquanto o *tu* foi empregado em 226 ocorrências, totalizando 11,3% dos dados. Em termos gerais, a distribuição do fenômeno estudado revela que nos cinco Estados selecionados para a análise há uma significativa preferência pela forma *você*, que tem aplicação semicategórica.

Esse resultado indica a comprovação de uma das hipóteses iniciais, a predominância do *você* em relação ao *tu*. Cardoso (2008) apresenta os resultados dos levantamentos de ocorrências do *tu* e *você* em seis capitais do Nordeste e duas do Sudeste, utilizando o *corpus* do ALiB. Nesta análise, a autora constatou uma variação diatópica desses pronomes, como mostram os dados da Tabela 14:

43 Convém salientarmos que, em relação à variante *cê* na nossa amostra, a mesma foi contabilizada juntamente com o pronomes *você*.

Tabela 14: Distribuição diatópica dos pronomes *tu* e *você* em capitais brasileiras

| PRONOMES | TU | | VOCÊ | | TOTAL GERAL |
|----------------|-------------------|----|-------------------|-----|-------------------|
| Localidades | Nº de ocorrências | % | Nº de ocorrências | % | Nº de ocorrências |
| Aracaju | 3 | 6 | 46 | 94 | 49 |
| Maceió | 4 | 6 | 60 | 94 | 64 |
| Recife | 6 | 16 | 31 | 84 | 37 |
| Salvador | - | 0 | 85 | 100 | 85 |
| João Pessoa | 9 | 13 | 60 | 87 | 69 |
| Teresina | 5 | 13 | 32 | 87 | 37 |
| São Paulo | - | 0 | 189 | 100 | 189 |
| Rio de Janeiro | - | 0 | 192 | 100 | 182 |
| Totais | 27 | | 865 | | 712 |

Fonte: Cardoso (2008, p. 9).

Conforme exposto na Tabela 14, a variante inovadora predomina em todas as capitais analisadas. No Nordeste, observa-se o uso categórico do *você* na amostra de Salvador, resultado que se assemelha às duas capitais do Sudeste (São Paulo e Rio de Janeiro). Em Teresina, Aracaju e Recife, destaca-se a baixa frequência do *tu*.

Os resultados obtidos na amostra de dados analisada nesta tese parecem então se equipararem aos resultados apontados em Cardoso (2008) para os dados do ALiB (Tabela 14), em que há uma suplantação do *tu* pelo *você* nos Estados nordestinos.

Em análise mais detalhada da distribuição das formas de tratamento no *corpus* estudado, verifica-se, na Tabela 15, o total de ocorrências de cada forma sob análise:

Tabela 15: Frequência absoluta das formas de tratamento

| VARIANTES | TU | | VOCÊ | |
|-----------|-----------|------|-------------|-------------|
| | Oc./Total | % | Oc./Total | % |
| Total | 226/1.995 | 11,3 | 1.769/1.995 | 88,7 |

Fonte: Elaborada pela autora.

A Tabela 15 confirma os resultados já mostrados na Figura 10: o uso expressivo do pronomes *você* com frequência de 88,7%, em detrimento do *tu*, com 11,3% de frequência nos cinco Estados nordestinos analisados.

É interessante ressaltar que, diferente do que ocorre em outras pesquisas nas quais a forma *tu* é usada com maior frequência, como em Bezerra (1994), Andrade (2010), Salles (2001) (ver item 1.4.4.1), nos Estados de Rio Grande do Norte, Paraíba, Pernambuco, Alagoas

e Sergipe, os falantes têm preferência pela forma *você*.

Merece destaque o número inibido de ocorrências do *tu* nos Estados em análise, o que se justifica ao modelo das entrevistas que parece não favorecer a ocorrência do fenômeno. Pode-se cogitar que tenha ocorrido nos dados desta pesquisa o mesmo que em Paredes Silva (2003) ao pesquisar as formas de tratamento no Projeto Censo/PEUL, o gênero do discurso – entrevista nos moldes sociolinguísticos – influenciou diretamente os números obtidos. A aplicação de entrevistas (método) sociolinguística consiste, segundo Tarallo (1985, p. 21), em “minimizar o efeito negativo causado pela presença do pesquisador”, de maneira que o informante se expresse o mais natural possível. Acreditamos, portanto, que o fato de o entrevistador empregar o *você* no momento das entrevistas pode ter levado o informante a monitorar a sua fala, repetindo a mesma forma usada pelo entrevistador.

Paredes Silva (2003, p. 162) afirma que a fala monitorada não facilita a interação entre os interlocutores, e é justamente nos contextos de maior interação, intimidade e informalidade que ocorre o pronome *tu*.

Mesmo com um número reduzido do pronome *tu* na amostra investigada, focaremos a nossa análise em elucidar quais os contextos em que esse pronome ainda resiste.

Nessa perspectiva, nas seções a seguir, exibiremos os resultados da rodada binária, em que foram selecionadas pelo GoldVarb X quatro variáveis (referenciação, parte do inquérito, localidade e escolaridade) como condicionantes linguísticas para o emprego do *tu* pelos informantes.

4.1.1 Variáveis selecionadas

Os resultados da rodada do GoldVarb X apontaram o uso categórico do pronome *você* com relação a algumas variantes/variáveis controladas, sendo essas estatisticamente irrelevantes e, consequentemente, foram descartadas para a rodada binária, uma vez que o método e o suporte quantitativo utilizados buscam identificar os fatores que regulam a variação.

Considerando a exclusão de alguns fatores da análise por não apresentar ocorrências da forma *tu*, recorre-se ao que sugerem Guy e Zilles (2007, p. 155) para o procedimento da análise quando não se constata variação em um dos fatores de determinado grupo. Os autores, com base na pesquisa sobre *nós* e *a gente*, afirmam que, “[...] se aceitarmos a hipótese de que esse é um contexto realmente categórico que exige o pronome *nós*, é preferível tirar os dados do *corpus*”.

Assim, os grupos de fatores ausentes nesta análise geral da amostra foram: função sintática do pronome, tempo verbal, paralelismo, sexo e faixa etária, uma vez que os mesmos não revelaram resultados significativos e, por isso, foram excluídos da análise.

Extraídas as variáveis/variantes que apresentaram nocaute, a rodada binária do GoldVarb X apontou como variáveis relevantes no emprego do *tu*, em ordem de seleção, (1) Referenciação, (2) Parte do inquérito, (3) Localidade e (4) Escolaridade. Foram selecionadas, portanto, variáveis linguísticas e sociais. Assim, confirmou-se uma das hipóteses específicas deste trabalho, segundo a qual o uso das formas de referência à 2ª pessoa estaria sujeito a fatores de ordem interna e externa à língua.

Sobre a atuação desses fatores, vale salientar que os resultados, expostos ao longo desta seção, são oriundos da seleção do GoldVarb X, que teve o nível de significância 0,001, o *log likelihood* -509.482 e o *input* de 0,058. A apresentação dos resultados seguirá a ordem de seleção do programa.

4.1.1.1 Referenciação

Optamos por controlar esta variável com base nas pesquisas sobre indeterminação, ou seja, referência genérica, a saber: Martins (2010), Deus (2009) e Divino (2008). Esses estudos observam as estratégias de referenciação dos pronomes *tu* e *você*. Ainda sobre a variável em questão, partimos da hipótese de que para a referência genérica os usos mais recorrentes são os do pronome *você*. A hipótese aqui defendida foi confirmada, uma vez que das 1.588 ocorrências, 1.363 foram do *você* com referenciação genérica.

A Tabela 16 exibe a distribuição dos dados desta variante pelos tipos de referenciação considerados.

Tabela 16: A variante *tu* em função da Referenciação nos cinco Estados nordestinos analisados

| REFERÊNCIA | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|------------|------------------|------------|--------------|
| Específica | 225/1588 | 14,2% | 0,711 |
| Genérica | 1/407 | 0,2% | 0,029 |

Input: 0,058; Log likelihood: -509.482; Significance = 0,001

Fonte: Elaborada pela autora.

Notamos, a partir dos resultados expostos na Tabela 16, que, de 1.588 ocorrências de pronome com referenciação específica, obtivemos 225 realizações do pronome *tu*,

respectivamente 0,711 de peso relativo e frequência de 14,2%, o que comprova que a referenciação específica foi a mais utilizada. Por outro lado, percebemos um desfavorecimento para a utilização do pronome *tu* em contextos de referenciação genérica, respectivamente, com peso relativo de 0,029.

Seguem exemplos dos tipos de referenciação controlados:

a) Referenciação específica:

INF. – (inint) Ô fulano, *tu* vai pra onde? O que é isso aí? (risos) (75-2)⁴⁴

INQ. – Se você quer saber se alguém vai sair hoje, como é que você pergunta?

INF. – *Você* vai sair hoje? (79-5)⁴⁵

b) Referenciação genérica:

INF. – Acho muito interessante, porque trata do cotidiano familiar, que, assim, *você* vê muita coisa que acontece na sua família, assim, coisas engraçadas... (79-2)⁴⁶

Contrariamente às conclusões a que chegamos através dos dados, Martins (2010), em análise da cidade de Tefé (AM), encontrou uma predominância do *tu*, favorecido pela variável referência específica, com peso relativo de 0,55, como mostra a Tabela 17:

Tabela 17: Uso do *tu* pelo tipo de referência

| TIPO DE REFERÊNCIA | FREQUÊNCIA DO <i>TU</i> | | PESO RELATIVO |
|--------------------|-------------------------|------|---------------|
| | N | % | |
| Específica | 334/447 | 74,7 | 0,55 |
| Genérica | 186/359 | 51,8 | 0,43 |
| TOTAL | 520/806 | 64,5 | |

Fonte: Martins (2010, p. 79).

O resultado da pesquisa de Deus (2009), que analisa as formas pronominais de 2^a pessoa em três capitais do Sul e três capitais do Nordeste, comprovou que o fator referência específica favoreceu o uso do *você* com peso relativo de 0,59. Isso implica dizer que o tipo de referência geral ou indeterminada tem favorecido a ocorrência do *tu*, visto que apresenta peso de apenas 0,30 para o uso do *você*.

⁴⁴ Informante de Santana do Ipanema-AL. Sexo feminino. Faixa etária I.

⁴⁵ Informante de Aracaju-SE. Sexo masculino. Faixa etária II. Nível universitário.

⁴⁶ Informante de Aracaju-SE. Sexo masculino. Faixa etária I. Nível fundamental.

Outros estudos dão notícia do uso categórico do *você* como referência genérica, como a pesquisa de Nogueira (2013) e a de Modesto (2006). Neste último, o autor notou que a referência indeterminada favorece o uso do *você*, ao registrar peso relativo de 0.67.

Nesse sentido, podemos concluir que o uso do *tu*, de modo geral, é de referência específica, como hipotetizávamos. Isso não ocorre somente entre os informantes deste estudo, uma vez que a maioria das pesquisas consultadas registra o uso genérico feito com o pronome *você*, como afirma Dias (2007), que confirma a referência genérica com 3,7% para *tu*, enquanto a aplicação do *você* foi de 96,3%.

4.1.1.2 Parte do inquérito

A utilização dos questionários é de grande importância para a pesquisa dialetológica, visto que a metodologia do ALiB, preocupada com o fornecimento de dados dialetais, foi desenvolvida para tornar a entrevista o mais próximo possível da conversa informal. Sabemos que esse instrumento permite a homogeneização do procedimento da coleta de dados necessária à pesquisa, atendendo a parâmetros científicos. Dessa maneira, controlamos o tipo/parte do questionário e apresentamos, na Tabela 18, os resultados para a utilização do pronome *tu*.

Como mencionado anteriormente, essa variável foi controlada para termos uma ideia de como estão atualmente distribuídas as formas de referência à segunda pessoa nos Estados em análise. Soma-se a isso a possibilidade de observar em qual tipo de questionário essas formas são mais recorrentes.

Tabela 18: A variante *tu* em função do Tipo do questionário nos cinco Estados nordestinos analisados

| | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|---------------------------|------------------|------------|--------------|
| Questões de prosódia | 38/530 | 7,2% | 0,312 |
| QSL | 27/338 | 8,0% | 0,489 |
| QMS | 113/742 | 15,2% | 0,615 |
| Discurso semidirigido | 29/264 | 11,0% | 0,567 |
| Questões metalinguísticas | 16/59 | 27,1% | 0,777 |
| Questões pragmáticas | 3/62 | 4,8% | 0,278 |

Input: 0,058; Log likelihood: -509,482; Significance = 0,001

Fonte: Elaborada pela autora.

A análise estatística das ocorrências permite visualizar que, na variável tipo de inquérito, as questões metalinguísticas apresentaram um favorecimento considerável para o uso do *tu*, obtiveram um peso relativo de **0,777** e frequência de 27,1%; já as questões morfossintáticas apresentaram o peso relativo de **0,615**, com percentual de 15,2%, enquanto o discurso semidirigido mostrou-se um pouco favorável com peso relativo de **0,567** e frequência de 11,0%. Sobre as questões metalinguísticas, sabemos que o falante dá esclarecimento sobre sua própria língua, respondendo as 6 questões que fazem parte deste questionário. Podemos inferir que o informante sentiu-se mais à vontade e confiante no segundo momento de respostas aos questionários citados, prova disso foi uma crescente utilização do pronome *tu*, concorrendo com o *você*. A relação entre informante e entrevistado, a partir da metade da entrevista, talvez, tornou-se solidária e não monitorada, facilitando, portanto, a utilização do pronome mais informal e íntimo *tu*, conforme excertos de fala:

INQ. – Era amiga mesmo.

INF. – Era duas chave ela levava uma, e outra ficava aí nesse torno aí comigo. Todo dia eu ia na casa dela olhá. E casa dela, era uma casa, não como a minha. A casa dela tinha muita coisa. Não fazia vergonha ninguém chegá na casa dela, era bem arrumada. Uma vez ela chegou de Recife, ela chegava qualquer hora, chegava com a chave dela e abria a porta. Aí de manhã ligava pra mim. *Tu* tá onde, ela dizia *tô* em casa vem tomar café mais eu.

INQ. – Olha.

INF. – E uma dia ela veio com uma visita. E quando a mulher chegou, eu arrumei a casa dela porque ela tava pra chegar. E quando a mulher chegou perguntou assim... L. *tu* passa esse tempo todinho fora e quando *tu* chega tua casa tá limpa, tá cheirando. Ela dizia não eu tenho uma vizinha eu deixo com ela. Ela dizer L. é de muita confiança essa pessoa, porque tua casa é muito bem arrumada, só televisão ela tinha quatro. (63-2)⁴⁷

Seguem outros exemplos de ocorrências com a variável tipo do questionário do ALiB:

a) Questões de prosódia

INQ. – Se você quer saber se seu amigo toma leite ou café, como pergunta?

47 Informante de Salgueiro- PE. Sexo feminino. Faixa etária I

INF. – *Você* toma café ou leite? Qual é a sua preferência? (74-3)⁴⁸

b) QSL

INF. – Ah, *você* tá perguntano quando vem a chuva forte. Acho que o nome é inverno, né? Que *você* queria? (74-1)⁴⁹

e) QMS

INQ. – Conhece alguma simpatia para tirar verruga?

INF. – Meu pai, meu avô me ensinava... dizia assim... a lua nova, é... mostrava a verruga pra lua e... contava quantas tinha e dizia que quando ela fosse cheia é... as berrugas teriam acabado. Uma outra é para os dente, nascia os dente, *você* rancava os dente e... aproveitava uma tarde no pô do só e jogava esse dente lá em cima da telha e... dizia algumas coisas assim, aí que dizia: “Morão, Morão, toma meu dente podre e me dá um sâo”. (74-3)⁵⁰

f) Discurso semidirigido

INF. – Ele tava sentadinho num banco assim, e ele chegô, o colega dele, né? Chegô e disse: Ei *você* um cabra safado e xingô, xingô, xingô. Ele... por que *você* foi me intregá? Ele respondeu pra ele: Eu, intregá *você*? O que tá aconteceno? Foi dizê que eu tava durmino em casa, só pode sê *você*! Pergunte quem foi que disse. (inint) Mas ele cheio de droga, que ele usa droga, pegô uma pau, botô na mão e disse: Eu sei onde é que eu tô que não lhe mato n'estante. (77-4)⁵¹

g) Questões metalínguísticas

INQ. – Tem gente que fala diferente aqui em Maceió?

INF. – Fala. Fala. As meninas já acha que eu falo diferente... sendo que eu não falo diferente. Porque ela diz: “ah, porque *você* foi... pa Bahia... *você* veio falano diferen...” Eu falei: “não, mulé... é impressão sua”. Ela disse: “Foi”. Que ela disse... eu passei... eu passei pouco tempo. (77-2)⁵²

48 Informante de União dos Palmares- AL. Sexo feminino. Faixa etária II.

49 Informante de União dos Palmares- AL. Sexo masculino. Faixa etária I.

50 Informante de União dos Palmares- AL. Sexo masculino. Faixa etária II.

51 Informante de Maceió. Sexo feminino. Faixa etária II

52 Informante de Maceió. Sexo feminino. Faixa etária I

h) Questões de pragmática

INQ. – Um objeto caiu do bolso de um rapaz jovem, e ele não viu. Como um outro rapaz jovem chama a atenção desse rapaz?

INF. – Fulano, *você* deixou cair uma chave! (77-2)⁵³

Lembramos que o programa GoldVarb X selecionou o grupo de fatores acima explicado como segundo de maior importância para o fenômeno sob análise.

Sobre a nossa hipótese, a mesma foi parcialmente confirmada, pois acreditávamos que no discurso semidirigido a forma *você* predominava e nas questões de prosódia, o *tu* era favorável, essa última, foi refutada.

É importante assinalar que, nas relações simétricas, os interlocutores são pessoas da mesma faixa etária, mesmo grupo social, familiares, amigos etc. Nesse sentido, acreditávamos na predominância do *tu*, nas questões de prosódia, pelo fato da pergunta ser direcionada a um amigo, pessoa próxima. Essa simetria, relação de reciprocidade entre os falantes, determinará as formas de tratamento escolhidas na situação comunicativa.

4.1.1.3 Localidade

Em relação à variação diatópica, algumas localidades apresentaram nocautes, pois não tiveram nenhuma ocorrência do pronome *tu*, a saber: Rio Grande do Norte (Mossoró, Pau dos Ferros, Caicó, Natal); Alagoas (União dos Palmares, Arapiraca); e Sergipe (Aracajú e Propriá). Em Pernambuco, excluímos Arco Verde por não apresentar também variação das formas de referência à segunda pessoa. Dessa forma, podemos concluir que no interior do Rio Grande do Norte, exceto em Angicos, não houve utilização do *tu*, ou seja, há suplantação do *você* nesses interiores.

Com base nos resultados encontrados, conferidos na Tabela 19, a cidade de Cabrobó, em Pernambuco, foi a localidade mais favorecida pelo uso do *tu*, com peso relativo de 0,78 e frequência de 31,7%. Depois, temos as cidades de Itaporanga, na Paraíba e Afrânio, em Pernambuco, com pesos relativos bem significantes, respectivamente, 0,77 (frequência de 35,5%), 0,717 (frequência de 22,9%). Ainda no Estado de Pernambuco, encontramos a cidade de Salgueiro com o peso relativo de aplicação do *tu* de 0,696 e frequência de 21,3%.

É relevante destacar que nas cidades de Propriá (SE), Aracajú (SE), União dos Palmares

53 Informante de Maceió. Sexo feminino. Faixa etária I

(AL), Arapiraca (AL), Arco Verde (PE), Natal (RN), Caicó(RN) Mossoró (RN) e Pau dos Ferros (RN) não houve nenhuma ocorrência do pronome *tu*.

Tabela 19: Resultado da variante *tu* em função da Localidade nos cinco Estados analisados

| | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|-------------------------|------------------|------------|--------------|
| Cabrobó (PE) | 26/82 | 31,7% | 0,782 |
| Itaporanga (PB) | 61/172 | 35,5% | 0,774 |
| Afrânio (PE) | 11/48 | 22,9% | 0,717 |
| Salgueiro (PE) | 19/89 | 21,3% | 0,696 |
| Cajazeiras (PB) | 8/58 | 13,8% | 0,630 |
| Recife (PE) | 18/122 | 14,8% | 0,629 |
| Garanhuns (PE) | 3/27 | 11,1% | 0,585 |
| Limoeiro (PE) | 17/84 | 20,2% | 0,584 |
| Caruaru (PE) | 6/31 | 19,4% | 0,567 |
| João Pessoa (PB) | 6/49 | 12,2% | 0,527 |
| Exu (PE) | 6/76 | 7,9% | 0,526 |
| Olinda (PE) | 8/52 | 15,4% | 0,524 |
| Cuité (PB) | 6/52 | 11,5% | 0,461 |
| Patos (PB) | 8/94 | 8,5% | 0,453 |
| Santana do Ipanema (AL) | 5/44 | 11,4% | 0,421 |
| Maceió (AL) | 12/208 | 5,8% | 0,294 |
| Petrolina (PE) | 2/42 | 4,8% | 0,245 |
| Angicos (RN) | 1/28 | 3,6% | 0,179 |
| Estância (SE) | 2/79 | 2,5% | 0,130 |
| Floresta (PE) | 1/64 | 1,6% | 0,113 |
| Aracajú (SE) | 0/121 | 0% | - |
| Caicó (RN) | 0/79 | 0% | - |
| Propriá (SE) | 0/49 | 0% | - |
| União dos Palmares (AL) | 0/83 | 0% | - |
| Arapiraca (AL) | 0/38 | 0% | - |
| Arco Verde (PE) | 0/11 | 0% | - |
| Natal (RN) | 0/43 | 0% | - |
| Mossoró (RN) | 0/30 | 0% | - |
| Pau dos Ferros (RN) | 0/35 | 0% | - |

Input: 0,058; Log likelihood: -509.482; Significance = 0,001

Fonte: Elaborada pela autora.

Incluímos, na Tabela 19, as cidades que foram excluídas para análise estatística, por acreditarmos na necessidade de uma visão mais ampla de como o fenômeno se revelou em todas as cidades escolhidas para amostra desta tese.

Levando em consideração a divisão dialetal (ver 4.2.2.1) proposta por Antenor Nascentes (1953, p. 20) que subdivide o falar nortista em dois subfalares (o amazônico e o nordestino) e o do sul em quatro subfalares: (i) o baiano, intermediário entre os dois grupos, abrangendo Sergipe, Bahia, o norte, nordeste e noroeste de Minas e o norte de Goiás (atualmente território do Estado do Tocantins); (ii) o fluminense, abrangendo Espírito Santo, Rio de Janeiro, a Mata e o leste de Minas Gerais; (iii) o mineiro, compreendendo o centro, o oeste e parte do leste de Minas Gerais; e (iv) o sulista, compreendendo São Paulo, Paraná, Santa Catarina, Rio Grande do Sul, Minas (sul e Triângulo mineiro), sul de Goiás e Mato Grosso, que até então compreendia os atuais Estados de Mato Grosso e Mato Grosso do Sul; a área dialetal intitulada de ‘falar nordestino’, deu a essa região, segundo o autor, uma certa peculiaridade, com traços que podem constatar sua unidade ou revelar sua diversidade. Assim, reiterando essa divisão feita por Nascentes, podemos reconhecer que nos Estados de Pernambuco e Paraíba, pertencentes a área linguística do ‘falar nordestino’, a variante *tu* ainda resiste, numa proporção maior, conforme explanado na Tabela 19.

Nossos resultados têm um ponto de congruência com a pesquisa de Deus (2009), cuja amostra foi constituída por seis capitais (três do Nordeste e três do Sul). Em relação ao Nordeste, os falantes soteropolitanos demonstram clara preferência pelo pronome *você*, não tendo sido observada, inclusive, nenhuma ocorrência do *tu* na amostra. Os falantes de Recife e Teresina também empregaram majoritariamente o *você*. Os índices de utilização da forma *tu* foi de 5% e 2%, respectivamente, o que não colocam o *tu* numa concorrência direta com o *você* nessas capitais. A autora salienta, no entanto, que é preciso levar em conta que o tipo de inquérito utilizado para captar os dados pode não ter favorecido a utilização de um pronome mais informal.

Diferentemente do que foi encontrado nos nossos dados, Lucca (2005) analisa a variação *tu* ~ *você* na fala de jovens brasilienses e conclui que o *tu* é empregado nas três regiões administrativas de Brasília, em maior grau naquela em que os moradores são de procedência nordestina. Os dados colhidos para a pesquisa de Lucca (2005) partiram de três regiões administrativas do Distrito Federal: Taguatinga, Ceilândia e Brasília.

Defendemos a hipótese de que os falantes do interior utilizariam mais o pronome *tu*, enquanto os falantes da capital prefeririam o *você*. Nossa hipótese não se sustenta pelos

resultados obtidos, pois tanto nas capitais como no interior a regra de aplicação do *você* é semicategórica, com percentuais estatisticamente semelhantes.

4.1.1.4 Escolaridade

A variável escolaridade foi a quarta selecionada pelo GoldVarbX. A escolaridade tem sido uma variável social pesquisada em vários trabalhos variacionistas, uma vez que as pessoas escolarizadas refletem, em geral, sinais da norma padrão, valorizada e ensinada na escola, quer dizer, a escola gera mudança na fala e na escrita das pessoas que a frequentam, no sentido de incorporarem no seu acervo linguístico uma maior produção de formas de expressão socialmente prestigiadas.

Como esperado, a escolaridade se mostrou um dos fatores mais relevantes neste estudo, tendo sido selecionada pelo GoldVarb X, confirmando que o fenômeno pesquisado é, em parte, explicado sob o aspecto extralinguístico.

Apresentamos na Tabela 20 os resultados para esta variável social na amostra de fala dos cinco Estados nordestinos analisados:

Tabela 20: A variante *tu* em função da Escolaridade do informante nos cinco Estados nordestinos analisados

| | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|---------------|------------------|------------|--------------|
| Fundamental | 216/1720 | 12,6% | 0,523 |
| Universitário | 10/275 | 3,6% | 0,361 |

Input: 0,058; Log likelihood: -509.482; Significance = 0,001

Fonte: Elaborada pela autora.

Os dados comprovam que das 1.720 ocorrências de referência à segunda pessoa utilizada por falantes de ensino fundamental, 216 realizações foram da variante *tu*, obtendo 0,523 de peso relativo, representando 12,6%. Por outro lado, os falantes de maior escolaridade desfavoreceram o uso do *tu*, com peso relativo de 0,361 e frequência, respectivamente, de 3,6%.

Salientamos que a nossa rodada estatística foi em função do *tu*, objetivando resgatar a utilização deste pronome nas falas nordestinas. Enfim, o pronome *você* foi, majoritariamente, escolhido pelos falantes de nível fundamental e universitário.

Nessa direção, a partir dos dados apresentados na Tabela 20, verifica-se que todos os falantes, independentemente do nível de escolaridade, utilizam mais o pronome *você* em detrimento do *tu*. Sendo assim, a nossa hipótese de que as pessoas com um maior grau de

escolarização tendem a utilizar mais *você* e as pessoas com grau menor tendem a usar mais o *tu*, foi descartada.

Observemos os excertos de fala:

- 1) INQ. – Você quer dizer a algumas pessoas que estão presentes que você está muito aborrecido com o que acontece. Como é que você diz?

INF. – Rapaz, eu tô é arretado! *Tu* não sabe o que é que aconteceu! (70-5)⁵⁴

- 2) INQ. – Se você quer saber se seu amigo toma leite ou café, como é que você pergunta?

INF: O que é que *tu* toma, leite ou café? (70-1)⁵⁵

As pesquisas de Santos (2012), Franceschini (2011), Zilli (2009), Modesto (2006) e Loregian-Penkal (2004) mostram um resultado diferente do encontrado nesta pesquisa: quanto menor a escolaridade, maior a frequência de uso do *tu*; quanto maior a escolaridade, maior o uso do *você*. Por outro lado, os dados da pesquisa de Guimarães (2014) apontam que, diferentemente dos estudos elencados, os mais escolarizados usam mais o pronome *tu*.

4.2 DESCRIÇÃO DAS RODADAS POR ESTADO

Nesta seção, analisaremos cada Estado separadamente, mais especificamente, serão explorados os Estados de Pernambuco, Alagoas e Paraíba, dos quais obtivemos além da frequência de uso, o peso relativo de cada variante do fenômeno dentro dos grupos selecionados. A análise de cada um desses fatores levou em consideração os resultados estatísticos e os percentuais de ocorrência gerados pelo GoldVarb X.

Cabe lembrar que foram feitas as rodadas individuais com as dez variáveis escolhidas por nós, testando assim, todos os grupos inicialmente propostos, tanto linguísticos como os extralingüísticos. Porém, nos Estados mencionados acima, ocorreram exclusões de alguns desses fatores, os quais serão explicados mais adiante.

54 Falante morador de Recife. Faixa etária II. Sexo masculino Nível universitário.

55 Falante morador de Recife. Faixa etária I. Sexo masculino. Nível fundamental.

4.2.1 Pernambuco

Para a análise do fenômeno no Estado de Pernambuco, foram escolhidas 12 cidades para fazerem parte dos pontos do ALiB, a saber: Exu, Salgueiro, Limoeiro, Olinda, Afrânio, Cabrobó, Arcoverde, Caruaru, Recife, Floresta, Garanhuns e Petrolina.

No primeiro momento da análise, fez-se uma visão geral da alternância pronominal, observando como as variáveis dependente e independente estão partilhadas. O mesmo procedimento será repetido nesse segundo momento de análise do fenômeno em variação.

Apresentamos, a seguir, as variáveis que foram selecionadas como significativas pelo programa estatístico, na rodada referente ao Estado de Pernambuco: (1) Referenciação; (2) Parte do inquérito; (3) Faixa etária; e (4) Localidade.

4.2.1.1 Referenciação

A referência diz respeito à forma como são usados os pronomes: se de forma específica, referindo-se ao interlocutor ou a alguém já referenciado antes, ou de forma genérica, quando se usa a segunda pessoa para se reportar a qualquer um. Na Tabela 21, apresentamos os resultados percentuais dessa variável.

Tabela 21: A variante *tu* em função da Referenciação no Estado de Pernambuco

| REFERÊNCIA | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|------------|------------------|------------|--------------|
| Específica | 116/582 | 19,9% | 0,690 |
| Genérica | 1/146 | 0,7% | 0,040 |
| | 117/728 | 16,1% | |

Input: 0,084 Log likelihood = -265.409 Significance = 0,032

Fonte: Elaborada pela autora.

Como vemos, a distribuição do fenômeno observado mostra que o *tu* é mais favorável em conversas nas quais a referência é específica, apresentando peso relativo de 0,690 e frequência de 19,9% de uso. Já em relação a referenciação genérica, obtivemos apenas uma ocorrência do *tu*, cujo peso relativo foi de 0,40 e frequência de 0,7%. Observando o total de ocorrências de referência específica, 582, obtivemos 466 de uso do *você*, o que implica que este pronome é altamente favorecido, nesta modalidade.

Como ilustração do que estamos chamando do *tu* genérico, observemos um trecho da transcrição do único inquérito nesse perfil:

INF. – As aulas são boas, tendeu?... só que, assim, depois que saiu de greve não vão todos os professores.

INQ. – É assim?

INF. – É eh, lá é assim. A hente tá largano, geralmente, mais cedo. Lá tem uma sala de computação, eles diz curso de computação, só que não num é um curso. Quando a gente... vai pra sala de computação num tem nem mais graça... te Se *tu* vai uma vez perdida e quando vai... assim, se tem o professô, assim, o professô que tem que dá aula, o... (inint) eh, a hente tem que fazê o que o professô manda. ó... Não, só que eles uns entra em... em site, Paparazo, essas coisa, ôtros já... eh, bota em jogos... Num é uma coisa organizada. Às veze... a gente tá numa, numa turma, mas tá levano falta em outra ôtra turma que a gente num não sabe... às veze a gente... não passa, aí bota como a gente passô, entendeu? Sim, algumas coisa são muito errada... (70-2)⁵⁶

Outro estudo realizado no Estado do Ceará é a pesquisa de Guimarães (2014), que analisa os dados de Fortaleza, comprova que o pronome *tu* é, discretamente, favorecido quando é usado de forma específica (frequência 53,3% e peso relativo 0,529), mas é amplamente desfavorecido se empregado de modo genérico (frequência 8,4% e peso relativo 0,110). A pesquisadora teve um total de 1.472 ocorrências, sendo 785 de referência específica com o *tu*. Percebe-se, portanto, que o pronome *tu* como forma genérica não é muito usual entre os informantes, e sim, o *você*.

O trabalho de Alves (2015) comunga com nossa hipótese: o peso de 0,32 e 0,58 indicam, respectivamente, que as variantes *tu* sem concordância e *tu* com concordância tendem a ser mais favorecidas em referências específicas, enquanto a variante *você* apresentou peso relativo de 0,08 e mostrou-se bastante favorecida com peso relativo de 0,77, em referências genéricas.

4.2.1.2 Parte do inquérito

Verificamos, nesta seção, como as formas de tratamento com o interlocutor – *tu* e *você* – se realizam a depender do tipo/parte do questionário. Esta variável foi a segunda selecionada em ordem de seleção pelo GoldVarb X e os resultados encontrados estão arrolados na Tabela 22.

56 Informante de Recife. Sexo feminino. Faixa etária I. Nível fundamental.

Tabela 22: A variante *tu* em função do Tipo do questionário no Estado de Pernambuco

| TIPO DO QUESTIONÁRIO | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|---------------------------|------------------|------------|--------------|
| Questões de prosódia | 28/239 | 11,7% | 0,331 |
| QSL | 13/102 | 12,7% | 0,603 |
| QMS | 52/247 | 21,1% | 0,589 |
| Discurso semidirigido | 18/99 | 18,2% | 0,616 |
| Questões pragmáticas | 2/21 | 9,5% | 0,351 |
| Questões metalinguísticas | 4/20 | 20,0% | 0,533 |
| | 117/728 | 16,1% | |

Input: 0,084 Log likelihood = -265.409 Significance = 0,032

Fonte: Elaborada pela autora.

A partir dos dados apresentados na Tabela 22, podemos depreender que a forma *tu* foi favorecida nas questões do discurso semidirigido, com peso relativo de 0,616 e 18,2% de uso; enquanto no QSL apresentou, respectivamente, peso relativo de 0,603 e percentual de 12,7%. Já no QMS, o pronome *tu* obteve 0,589 de peso relativo e 21,1% de frequência. Sobre as questões metalinguísticas, as mesmas exibiram efeito levemente favorecido, com peso relativo de 0,533 e 20% de percentual. Infere-se portanto, que nos relatos pessoais, muitas vezes emitido com muita emoção, os falantes deram preferência pelo *tu*.

No que concerne às questões metalinguísticas, formuladas com objetivo de registrar a variação diarreferencial⁵⁷, a presença do pronome *você* revelou-se bastante expressiva, com 16 ocorrências e, respectivamente, 4 ocorrências da variante *tu*.

Os questionários são instrumentos imprescindíveis para a realização do trabalho de campo. Como já salientamos anteriormente, as nossas expectativas eram as seguintes: nos discursos semidirigidos haveria maior utilização do *você* e nas questões de prosódia, maior uso do *tu*. Por conseguinte, nossas hipóteses foram rejeitadas.

⁵⁷ As seis questões metalinguísticas procuram apurar o conhecimento linguístico do falante. Em resposta a tais questões, buscam-se as formas lexicais utilizadas por falantes de outras áreas, por indivíduos mais idosos ou mais jovens, assim como aos inúmeros “sotaques”.

4.2.1.3 Faixa etária

A investigação desta variável social é indispensável para a análise de qualquer fenômeno a partir do método da sociolinguística laboviana, tanto em pesquisas em *tempo aparente* como em *tempo real*, pois vai indicar se o estágio de determinada variação linguística corresponde à variação estável ou mudança em progresso.

Essa variável foi selecionada pelo GoldVarb X como a terceira mais relevante para o estudo da variação dos pronomes no Estado de Pernambuco.

Tabela 23: A variante *tu* em função da Faixa etária do informante em Pernambuco

| | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|-------------------|------------------|------------|--------------|
| Faixa etária 1 | 56/326 | 17,2% | 0,569 |
| Faixa etária 2 | 61/402 | 15,2% | 0,444 |
| | 117/728 | 16,1% | |

Input: 0,084 Log likelihood = -265.409 Significance = 0,032

Fonte: Elaborada pela autora.

De posse dos dados expressos na Tabela 23, notamos que a variável faixa etária aponta para o favorecimento do uso do *tu* na faixa etária I, com peso relativo de 0,569 e percentual de 17,2%, ao passo que a faixa II revelou um peso relativo de 0,444 e percentual de 15,2% de uso.

Em contrapartida, nas duas faixas etárias prevaleceram a variante *você* como forma majoritária.

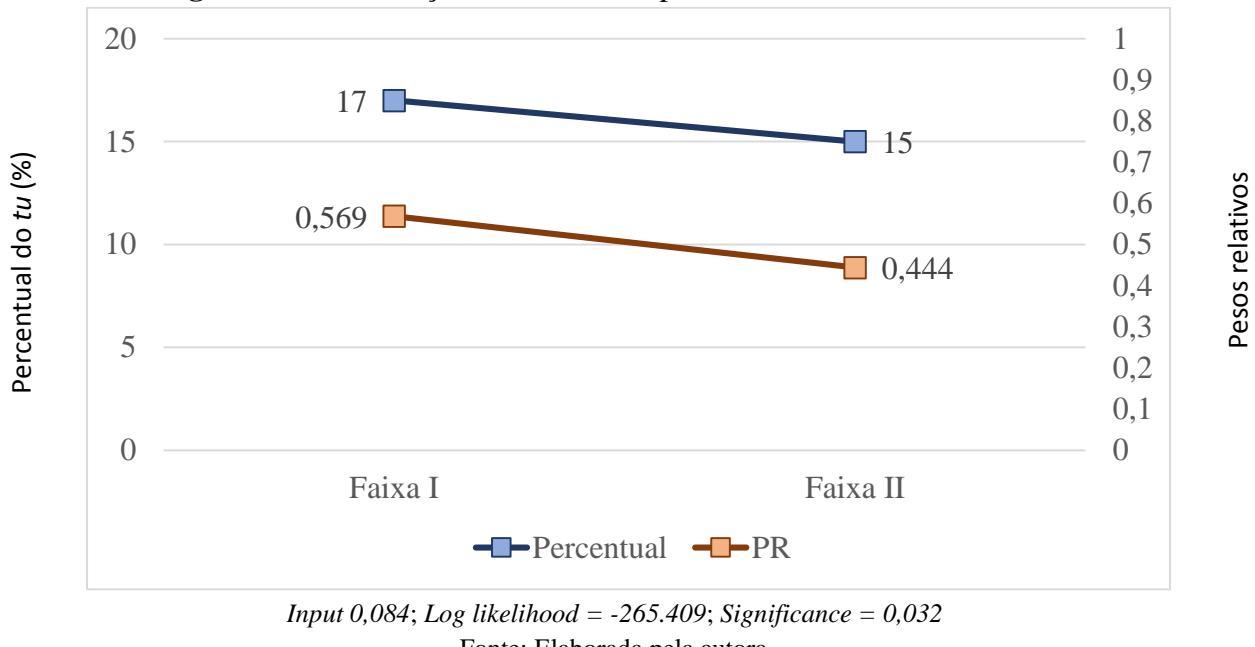
Para a análise dessa variável, adotamos como hipótese a crença de que a faixa etária II (50 a 60 anos) utiliza mais o pronome *tu* e, por outro lado, os falantes da faixa etária I (18 a 30 anos) utilizam mais a forma *você*, pelo fato dessa variante ser mais prestigiada socialmente e, estes falantes (faixa etária I) encontrarem-se, normalmente, inseridos no mercado de trabalho, fato que influencia seu comportamento linguístico. Desse modo, a nossa hipótese foi parcialmente confirmada, pois das 326 ocorrências na faixa etária I, 270 foram realizações do pronome *você* e 56 do *tu*.

Na pesquisa de Martins (2010), os resultados mostram que a faixa etária mais jovem tem efeito fortemente favorecedor sobre o uso do *tu* na amostra, com peso relativo de 0,95, contrariando, portanto, nossos resultados. Já nas pesquisas de Franceschini (2011), Divino (2008) e Oliveira (2007) o uso do *você* é favorecido na faixa etária mais jovem.

Encontram-se, na Figura 13, os percentuais referentes à distribuição das faixas etárias I

e II no Estado de Pernambuco.

Figura 13: Distribuição da variante *tu* por faixa etária em Pernambuco



Esse gráfico permite uma melhor visualização dos resultados, evidenciando que a faixa etária I apresenta peso relativo (0,569) bem próximo à faixa etária II (0,444).

É interessante salientar também que o *tu* no Nordeste aparece associado ao verbo na terceira pessoa, acarretando assim, o estigma da variante. Nesse sentido, o traço de mais intimidade, vinculado a essa variante, traz uma avaliação social negativa e, portanto, é a forma não usada pelos falantes que pretendem ser mais formais.

4.2.1.4 Localidade

A fim de observar a interferência da origem geográfica do falante no uso das formas de referência à segunda pessoa, as ocorrências foram identificadas conforme a localidade dos informantes. Com isso, buscou-se investigar a variação diatópica concernente à aplicação da regra variável nas cidades de Pernambuco.

A cidade de Arcoverde, como já mencionamos anteriormente, foi excluída da análise do programa GoldVarb X, visto que, no cômputo do fenômeno, identificou-se o uso categórico da variante *você*.

Tabela 24: A variante *tu* em função da Localidade em Pernambuco

| CIDADE | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|----------------|------------------|--------------|--------------|
| Exu | 6/76 | 7,9% | 0,479 |
| Salgueiro | 19/89 | 21,3% | 0,647 |
| Limoeiro | 17/84 | 20,2% | 0,533 |
| Olinda | 8/52 | 15,4% | 0,510 |
| Afrânio | 11/48 | 22,9% | 0,679 |
| Cabrobó | 26/82 | 31,7% | 0,764 |
| Caruaru | 6/31 | 19,4% | 0,528 |
| Floresta | 1/64 | 1,6% | 0,086 |
| Garanhuns | 3/27 | 11,1% | 0,508 |
| Petrolina | 2/42 | 4,8% | 0,192 |
| Recife | 18/122 | 14,8% | 0,530 |
| Total | 117/717 | 16,3% | |

Input: 0,084 Log likelihood = -265.409 Significance = 0,032

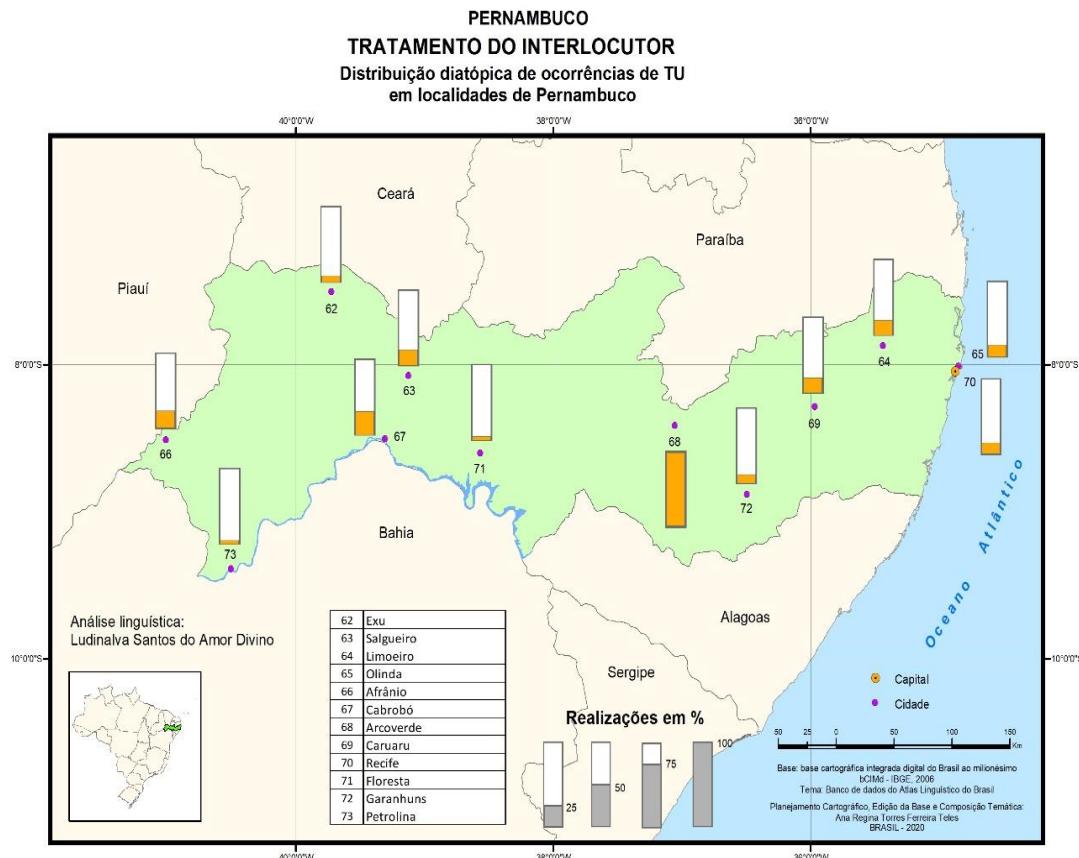
Fonte: Elaborada pela autora.

Observamos, a partir dos dados da Tabela 24, que das 717 ocorrências de pronomes de referência ao interlocutor neste Estado, 117 referem-se à variante *tu*. É possível visualizar, nessa distribuição, a atuação do fator geográfico na variação das formas de referência ao interlocutor.

Verificamos, assim, que a cidade de Cabrobó, lidera como contexto favorável ao uso do *tu*, com peso relativo de 0,764. Logo em seguida, aparecem as cidades de Afrânio, com peso relativo de 0,679, Salgueiro, com 0,647 e Limoeiro, com 0,533. Destaca-se também a capital Recife com peso relativo quase neutro de 0,530. A respeito desta capital, os dados coletados por Deus (2009) indicam peso relativo de 0,32 em relação à variante *você* e 88% de utilização da mesma, ou seja, esse resultado corrobora com estudos já divulgados sobre o fenômeno no referido Estado. A cidade de Floresta desfavoreceu fortemente a variante *tu*, visto que das 64 ocorrências, contou-se apenas com uma realização do *tu*, perfazendo 1,6% de percentual.

A Figura 14 exibe a representação dessas ocorrências, na área geográfica de Pernambuco.

Figura 14: Carta linguística 1: Distribuição diatópica de ocorrência de *tu* em Pernambuco



Fonte: Elaborada pela autora. Colaboração cartográfica: Ana Regina Teles.

A partir da leitura da Figura 14, confirmam-se as respostas documentadas para o tratamento do interlocutor no Estado de Pernambuco bem como se demonstra que as cidades de Cabrobó, Afrânio e Salgueiro, as quais apresentaram um maior uso da variante *tu*, estão localizadas em um mesmo eixo geográfico.

As cidades de Cabrobó e Salgueiro, coincidentemente, possuem um número significativo de indígenas e remanescentes de quilombo, com histórias que estão atreladas às lutas pela garantia da terra e pela preservação das suas culturas. Nesse sentido, acreditando que língua e cultura são indissociáveis, podemos pensar que o fenômeno linguístico em análise esteja atrelado à realidade vivida por essas duas localidades, isto é, a utilização do pronome *tu* justifica-se nessa direção.

A pesquisa de Sette (1980), realizada em Recife, coincide com os nossos resultados. A autora salienta que, em relação às formas *tu* e *você*, a maioria dos informantes diz usar mais a forma *você* (65%), em segundo lugar aparecem as pessoas que dizem usar somente *você* (20%)

e, em último lugar, vêm as que usam mais a forma *tu* (2,5%). A maioria dos informantes conceituou o *tu* como forma muito íntima, reservada apenas para familiares ou pessoas com quem se tem grande aproximação.

Por fim, a variação das formas analisadas *tu* e *você* justifica-se também pelo fato de cada região apresentar características distintas em seus respectivos vernáculos, em decorrência de fatores diversos, sobretudo da questão sócio-histórica. Assim, mesmo os dados apresentando um baixo percentual de uso da variante *tu*, cumpre destacar que a construção da identidade linguística e social do falar nordestino sempre esteve arraigada às dificuldades de sobrevivência, problemas histórico-sociais, precariedade ocupacional, analfabetismo etc.

4.2.2 Paraíba

Para compor os dados do Estado da Paraíba, os pontos do ALiB foram: Cuité, Cajazeiras, Itaporanga, Patos, Campina Grande e João Pessoa⁵⁸.

As rodadas binárias, geradas pelo GoldVarb X, nos forneceram, em ordem de seleção, as variáveis (ou grupo de fatores) significativas na análise: (1) Parte do inquérito; (2) Faixa etária; e (3) Localidade.

Passaremos a apresentar e discutir, na ordem de seleção, os resultados estatísticos gerados para cada variável e suas implicações para nossa análise.

4.2.2.1 Parte do inquérito

Na Tabela 25, é apresentada a distribuição do *tu* com relação aos questionários que compõem a metodologia do ALiB.

⁵⁸ Salientamos que a cidade de Campina Grande não compôs a nossa amostra, devido aos ruídos nos áudios, que impediram a audição das gravações.

Tabela 25: A variante *tu* em função do Tipo do questionário no Estado da Paraíba

| TIPO DO QUESTIONÁRIO | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|---------------------------|------------------|--------------|--------------|
| Questões de prosódia | 9/109 | 8,3% | 0,333 |
| QSL | 12/74 | 16,2% | 0,371 |
| QMS | 49/174 | 20,2% | 0,627 |
| Discurso semidirigido | 10/48 | 20,8% | 0,537 |
| Questões metalinguísticas | 8/10 | 80,0% | 0,921 |
| Questões pragmáticas | 1/10 | 10,0% | 0,318 |
| Total | 89/425 | 20,9% | |

Input: 0,166; Log likelihood = -182.557; Significance = 0,016

Fonte: Elaborada pela autora.

Tendo em vista o elevado número de ocorrências com a variante *tu* nas questões metalinguísticas, percebe-se a preferência do falante paraibano por esta variável, confirmando o peso relativo de 0,921 e 80% de frequência no uso. O contexto do QMS favorece levemente o emprego do *tu*, com peso relativo 0,627 e, por último, o discurso semidirigido, cujo peso é 0,537 e percentual uso de 20,8%, contexto que se mostra muito próximo ao ponto neutro.

Sobre as questões metalinguísticas, podemos reiterar que a percepção do falante sobre outras variantes e outros falares, verificando até que ponto fatos identificados como característicos de outro estrato social, de outra geração ou de outras áreas estão presentes em sua fala, nem sempre coincide inteiramente com a sua própria realização linguística.

As seis questões metalinguísticas, que levam às respostas sobre crenças e atitudes linguísticas, dão um passo à frente nessa área do conhecimento e foi objeto de análise realizada por Vanderci Aguilera, na Universidad de Alcalá de Henares, Espanha, em 2006, em nível de pós-doutorado. Aqui, no nosso estudo, percebemos que, devido a ser o penúltimo questionário do ALiB, o informante paraibano sentiu-se mais próximo do entrevistador e, por isso, utilizou uma linguagem mais espontânea. Prova disso foi o percentual de 80% de realização do pronome *tu*.

Analizando os resultados por outra perspectiva, a variante *você* tem seu uso majoritário nas questões de pragmática, visto que se obteve apenas uma ocorrência da variante *tu*.

4.2.2.2 Faixa etária

Nesta pesquisa, de caráter sincrônico, utilizamos o recurso do tempo aparente, segundo o modelo laboviano. Como já mencionado, para a análise deste fator, consideramos duas faixas etárias, faixa I (18 - 30 anos) e faixa II (50 a 65 anos).

Estudos sobre a referência à segunda pessoa, tais como os de Paredes Silva (2003), Lorean-Penkal (2004), Dias (2007), Divino (2008) e Martins (2010), indicam que são os mais jovens que utilizam a variante *tu* com maior frequência.

Vejamos os resultados obtidos através da análise da amostra na Tabela 26.

Tabela 26: A variante *tu* em função da Faixa etária do informante na Paraíba

| | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|-------------------|------------------|------------|--------------|
| Faixa etária 1 | 21/161 | 13% | 0,391 |
| Faixa etária 2 | 68/264 | 25,8% | 0,567 |
| | 89/425 | 20,9% | |

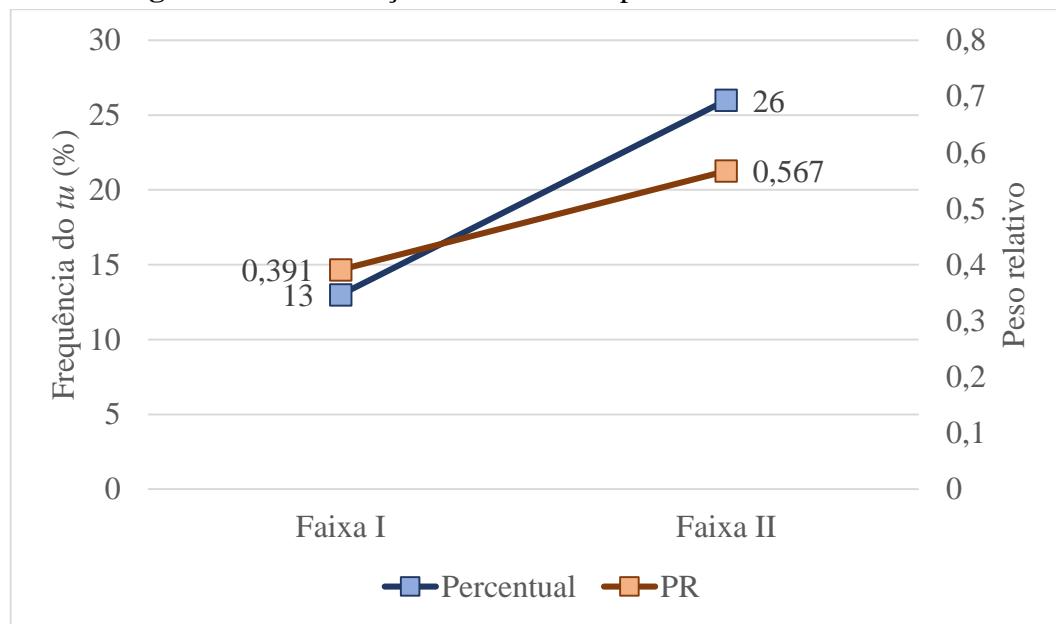
Input 0,166; Log likelihood = -182.557; Significance = 0,016

Fonte: Elaborada pela autora.

Conforme os resultados apontados na Tabela 26, fica evidente que a hipótese com que se trabalhou, inicialmente, não foi confirmada, ou seja, o grupo mais jovem usou menos o pronome *tu* do que o grupo mais velho (faixa II). O peso relativo de 0,567, alcançado na faixa II, sinaliza que as pessoas mais velhas ainda utilizam a variante *tu*, nas suas interlocuções. Por outro lado, ressalta-se que, na Paraíba, os falantes da faixa II favorecem a utilização da variante *você*.

O gráfico a seguir apresenta uma melhor visualização dos resultados:

Figura 15: Distribuição da variante *tu* por faixa etária na Paraíba



Input: 0,166; Log likelihood = -182.557; Significance = 0,016

Fonte: Elaborada pela autora.

No que concerne ao peso relativo apresentado na Figura 15, observamos que na faixa I há desfavorecimento do uso (0,391), enquanto na faixa II, o favorecimento é quase neutro (0,567), como mencionado anteriormente. A faixa II demonstra que a probabilidade de uso da forma inovadora *você* é maior entre os falantes desse grupo, já que das 264 ocorrências, em 68 o *tu* é empregado (25,8%) e em 196, o *você*.

4.2.2.3 Localidade

Assim como fizemos com os outros Estados, para observar a variação geográfica no Estado da Paraíba, separamos cada localidade no intuito de verificar uma possível dialetação interna. Os percentuais de uso dos pronomes de 2^a pessoa aqui analisados estão expostos por localidade na Tabela 27.

Tabela 27: A variante *tu* em função da Localidade na Paraíba

| CIDADES | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|--------------|------------------|--------------|--------------|
| Cuité | 6/52 | 11,5% | 0,413 |
| Cajazeiras | 8/58 | 13,8% | 0,493 |
| Itaporanga | 61/172 | 35,5% | 0,685 |
| Patos | 8/94 | 8,5% | 0,295 |
| João Pessoa | 6/49 | 12,2% | 0,342 |
| Total | 89/425 | 20,9% | |

Input: 0,166; Log likelihood = -182.557; Significance = 0,016

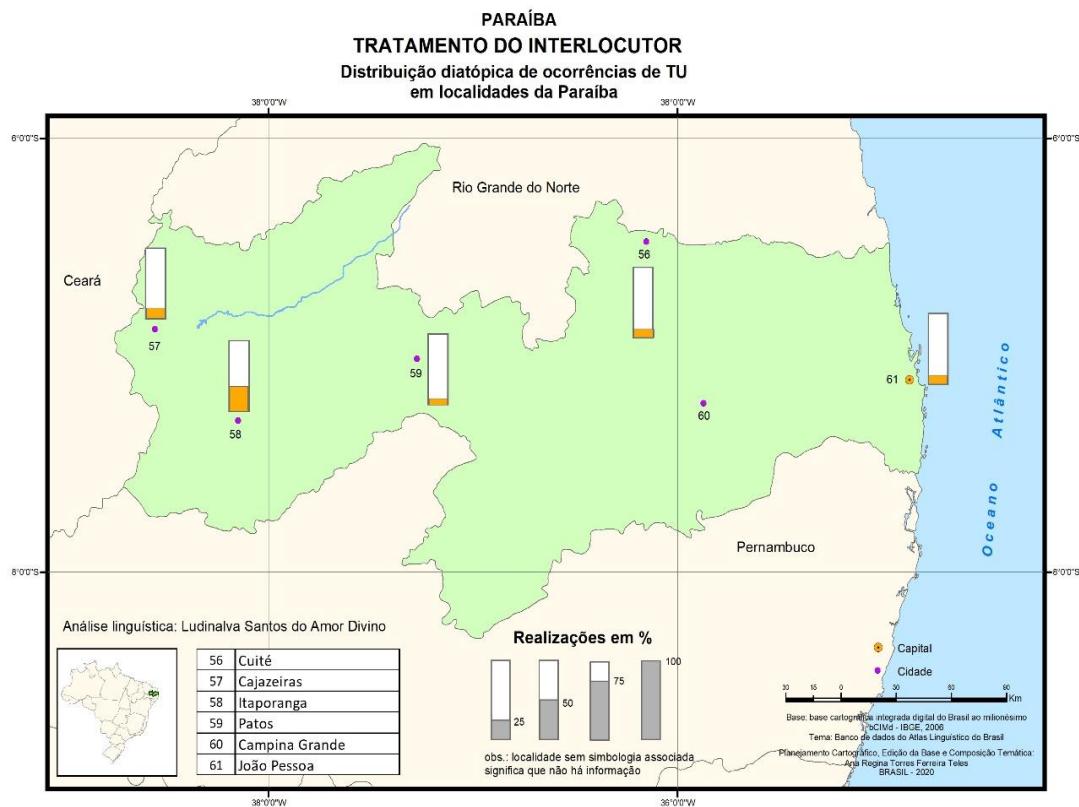
Fonte: Elaborada pela autora.

De posse dos dados sobre a variação com relação à origem geográfica do informante, comprova-se que das cinco localidades arroladas, Itaporanga é a única que resiste ao fenômeno da variação do *tu* e *você*, favorecendo a variante *tu*, respectivamente, com peso relativo de 0,685 e percentual de 35,5%. Por conseguinte, é na cidade de Patos, onde encontramos o menor percentual de uso (8,5%) do *tu*.

Diante disso, temos a variante inovadora sendo usada de forma majoritária no Estado da Paraíba. Não confirmando a hipótese inicial, nem a distribuição equitativa das formas em análise, acredita-se que isto revele o contexto de resistência do *tu* como forma de tratamento íntimo. Por outro lado, o *você* confirma-se como forma prestigiada nas mais diversas diádias sociais.

Assim sendo, na junção geral dos dados no Estado em pauta, obtivemos 425 ocorrências de pronomes de segunda pessoa, com 336 utilizações do *você* (79,1%) e apenas 89 utilizações do *tu* (20,9%), conforme Tabela 27. A distribuição do fenômeno ora observado mostra que os falantes paraibanos dão preferência a variante *você*. A Figura 16 destaca essa realidade.

Figura 16: Carta linguística 2: Distribuição diatópica de ocorrência de *tu* na Paraíba



Fonte: Elaborada pela autora. Colaboração cartográfica: Ana Regina Teles.

A Figura 16 revela todas as ocorrências de segunda pessoa, por localidade, no Estado da Paraíba. Essa carta linguística foi desenhada com a intenção de ressaltar a verticalidade das respostas obtidas para o tratamento do interlocutor. Buscamos, assim, demonstrar como a variante *tu* está distribuída nesse espaço geográfico.

Como observado, nossos resultados se distanciam de Bezerra (1994), que através dos seus dados coletados na Paraíba, fez uma análise sociolinguística interacional e não quantitativa, ou seja, não deu prioridade aos dados estatísticos. A pesquisa comprovou a predominância do *tu* sobre o *você*, com 69% das ocorrências do *tu*, contra 31% de ocorrências do *você*. O uso de uma ou outra variante parece estar relacionado ao tipo de ato comunicativo envolvido na interação – pedido, ordem, pergunta, desculpa, insulto, consentimento, entre outros.

Por ora, permite-nos observar que, mesmo o *você* apresentando-se como uso preferencial nas localidades elencadas na Tabela 26 e demonstradas na Figura 14, encontramos a variante *tu* em algumas falas paraibanas, alterando com o *você*, como veremos nos excertos de fala:

1-

INF. – Tem um primo dele que é promotor, no Piauí, paraibano é corajoso. Ele trabalha de delegado em Campina Grande. Aí disse: “Pai, me dê um dinheiro aqui pra eu fazer uma inscrição pro Piauí.” Eu disse: “Oxen! *Tu* ganha quase mil reais e tá pedindo dinheiro.” Foi ano passado, tem quase dois ano que ele trabalha lá. “Ah, foi que eu paguei uns negócio que tinha, uma prestação do carro.” “Quanto que é a inscrição?” “Quinhentos reais.” O pai dele deu. “Pro Piauí?!” “Vou fazer pra promotor.” “*Você* é doido, *tu* num vai passar nada. Quantas vaga tem?” “Cinco. O paraibano sou eu, vou ser o vitorioso.” Tá bom... ele fez, passou pro Piauí. (58/3)⁵⁹

2- A)

INQ. – Poderia dar exemplo do modo como falam essas pessoas? “*que falam diferente*”
INF. – É, como assim... meu sogro chama “meiro irmão”. Eu num sei, assim, porque num chama logo “meu irmão”, né. Aí chama: “Oh, meiro irmão”. Isso aí é uma fala diferente da nossa, né.

INQ. – Humrum.

INF. – *Tu* tá intendeno como é?

INQ. – Tô, tô entendendo. (58/1)⁶⁰

B)

INQ. – Quando você toma conhecimento de que um amigo casou, como comenta, com esse amigo essa novidade?

INF. – Eu já sabia, mai...mai...eu já sabia, mai o cara poderia tê vindo pelo menos falá: “Ói, num vai dá pra *você* í não. Ó, desculpa aí”. (58/1)⁶¹

Chamou nossa atenção, o fato de todos os informantes de Itaporanga utilizarem, nas suas interlocuções o *tu* e o *você*, simultaneamente.

Essa composição nos leva a inferir que a seleção por uma ou outra variante não depende apenas do grau de atenção prestado à fala (LABOV, 2001), mas de outros elementos que regem o momento da interação que, juntos, levam o indivíduo a mudanças estilísticas. Dessa maneira, acreditamos que por se tratar de entrevistas sociolinguísticas em que os informantes estão

59 Falante morador de Itaporanga-PB. Sexo masculino. Faixa etária II. Nível fundamental.

60 Falante morador de Itaporanga-PB. Sexo feminino. Faixa etária I. Nível fundamental.

61 Falante morador de Itaporanga-PB. Sexo feminino. Faixa etária I. Nível fundamental.

interagindo com um desconhecido, numa relação assimétrica⁶², utilizam um estilo (ver 2.2.2) semi-monitorado, por isso a variante *você*, bastante utilizada em relações assimétricas, concorre com a variante *tu*, utilizada em relações simétricas⁶³.

Bortoni-Ricardo (2004, p. 62) ressalta o contínuo de monitoração estilística, pois há interações nas quais o falante situa-se de forma espontânea e outras que são previamente planejadas e requerem atenção do falante à forma da língua, ou seja, há situações menos monitoradas e mais monitoradas.

4.2.3 Alagoas

O Estado de Alagoas é composto, na sua área geográfica, por 102 localidades. Dentre elas, foram escolhidas quatro cidades para integrarem a rede de pontos do ALiB: União dos Palmares, Santana do Ipanema, Arapiraca e Maceió.

Mesmo depois do mapeamento realizado por Scherre et al. (2015) mostrar todas as regiões do país, percebemos que alguns Estados não apresentam dados, ou seja, neles não existem muitos estudos relacionados à variação do *tu* e *você*, como é o caso, por exemplo, do Estado de Alagoas. Na pesquisa de Silva (2019), que analisa a variação dos pronomes *tu/você/cê* na posição de sujeito no sertão alagoano (Projeto Lusa⁶⁴), foram registradas apenas 3 realizações do *tu* e, por esse motivo, a autora optou por analisar apenas a variação entre o pronome *você* e sua variante *cê*.

Os resultados obtidos nesta tese, com relação ao Estado de Alagoas, incluindo capital e interior, indicaram que das 374 ocorrências de pronomes de 2^a pessoa, 356 foram do *você* (95,4%) e 17 foram do *tu* (4,6%). Na rodada binária, o GoldVarb X selecionou como estatisticamente significativas as variáveis Sexo e Tipo (parte) do questionário, nesta ordem.

Cumpre mencionar que as cidades de Arapiraca e União dos Palmares foram excluídas da análise por não apresentarem variação do fenômeno em questão.

27 Relações assimétricas: descendente (superior-inferior) e ascendente (inferior-superior) – aquelas mantidas entre informantes de diferentes faixas etárias, e em relações de hierarquia como pais/filhos, filhos/pais, professor/aluno, aluno/professor, entre outras.

63 As relações simétricas ou recíprocas são aquelas que os interlocutores são próximos; são relações igualitárias.

64 O projeto Lusa conta com um banco de dados composto por 96 entrevistas de informantes residentes no sertão alagoano.

4.2.3.1 Sexo

As diferenças mais nítidas entre a fala de homens e mulheres se situam no plano lexical. Vários estudos em referência à correlação linguística entre variação linguística e sexo/gênero apontam para uma diferença de valorização social, isto é, as mulheres demonstram maior tendência ao uso de variantes linguísticas mais prestigiadas socialmente.

Como salientado, diferentes pesquisas indicam para o que poderíamos denominar uma maior consciência feminina do *status* social das formas linguísticas. Assim, os padrões de correlação entre sexo/gênero e a variação linguística permitem às mulheres uma maior aproximação e receptividade à atuação normativa da escola. Nesse sentido, acreditamos que as mulheres estão mais dispostas à utilização do *você*.

Essa variável foi a primeira selecionada como estatisticamente relevante pelo GoldVarb X. A Tabela 28 exibe os resultados encontrados.

Tabela 28: A variante *tu* em função do Sexo do informante em Alagoas

| SEXO | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|--------|------------------|------------|--------------|
| Homem | 3/175 | 1,7% | 0,250 |
| Mulher | 14/198 | 7,1% | 0,725 |

Input: 0,033; Log likelihood = -55.437; Significance = 0,000

Fonte: Elaborada pela autora.

Conforme expresso na Tabela 28, as mulheres empregaram mais a variante *tu* com peso relativo de 0,725, ao passo que a utilização do *tu* foi inexpressível na fala dos homens (0,250). No seu turno, das 198 ocorrências, 14 foram da variante *tu* e 184 da variante *você*, pre eminentemente a favorita entre as mulheres; confirmado, portanto a nossa hipótese.

Diferentemente dos nossos resultados, as pesquisas realizadas por Andrade (2010), Dias (2007) e Lucca (2005) nas regiões Centro-Oeste e por Santos (2012), no Sudeste, mostraram que os homens favorecem o uso do *tu*.

Cardoso (2008), analisando a variação do *tu* e *você*, com dados do Projeto ALiB, encontrou um total de 64 realizações desses pronomes na cidade de Maceió, que estão distribuídas em 4 realizações do pronome *tu* e 60 realizações do pronome *você*, representando 6% de uso do *tu* contra 94% de uso do *você*. A autora seguiu a visão da Geolinguística Pluridimensional e, sobre as variáveis sociais consideradas, o pronome *você* é mais frequente entre os homens da faixa etária 1 do ensino fundamental.

Outra pesquisa para contrastarmos os resultados aqui obtidos é a de Nogueira (2013), cujo

corpus foi formado por conversas espontâneas, intitulado como “dados complementares”. Nesta, as mulheres favoreceram o uso do pronome *tu*, com 50,5% dos dados e os homens, respectivamente, 33,3%.

Modesto (2006), ao analisar esse fator, observa que, quando se trata de fala menos monitorada, as mulheres utilizam a variante *tu* em maior frequência que os homens, correspondendo a 54% dos dados.

Dessa maneira, podemos concluir que as mulheres utilizam mais a variante *tu* quando se trata de falas menos monitoradas; acreditamos também no papel feminino atrelado ao valor social, atuando assim na escolha pela variante *você* como forma inovadora, considerando que essa variante atua em níveis de consciência social.

Homens e mulheres, mesmo vivendo na mesma comunidade de fala, usam meios diferentes para se comunicar. A escolha pela norma padrão, pelas mulheres, deve-se a vários fatores; inclusive estas tendem a se corrigir mais que os homens nos contextos mais formais. Consideradas como “sexo frágil”, é necessário quebrar preconceitos para serem aceitas e isso se reflete nas escolhas linguísticas.

4.2.3.2 Parte do inquérito

Para a variável tipo de inquérito, nossa hipótese era de que o falante tenderia a usar mais o pronome *você* no discurso semidirigido e, nas questões de prosódia, a tendência maior seria para o uso da variante *tu*. Vejamos na Tabela 29 como as variantes revelaram-se nas falas dos alagoanos.

Tabela 29: A variante *tu* em função do Tipo do questionário no Estado de Alagoas

| PARTE DO QUESTIONÁRIO | OCORRÊNCIA/TOTAL | PERCENTUAL | PR |
|---------------------------|------------------|-------------|--------------|
| QSL | 2/95 | 2,1% | 0,376 |
| QMS | 9/128 | 7 % | 0,578 |
| Discurso semidirigido | 2/50 | 4 % | 0,334 |
| Questões metalinguísticas | 4/16 | 25% | 0,933 |
| Total | 17/209 | 5,9% | |

Input: 0.033; Log likelihood = -55.437; Significance = 0,000

Fonte: Elaborada pela autora.

Os resultados obtidos para esta variável, contrariamente à hipótese aventada, demonstram que a variante *tu* não faz parte da fala alagoana.

Conforme a disposição dos dados na Tabela 29, a variante *tu* revelou-se bastante resistente, apenas nas questões metalinguísticas, com peso relativo de 0,933 (25%). No QMS, das 128 ocorrências, 119 foram realizações do *você* e 9 da variante *tu*, apresentando peso relativo de 0,578 e frequência de 7%.

As questões metalinguísticas foram as mais favorecidas pelos falantes de Alagoas, fato ocorrido também na Paraíba, o que nos leva a acreditar na mesma justificativa para os dois Estados (cf. 4.2.2.1).

Nos discursos semidirigidos, o *tu* surge de forma bem retraída, com valor relativo de 0,334 e percentual baixíssimo, 4,0%.

Uma vez que esse enfoque é inovador e não tenha sido localizado nenhum trabalho com análise semelhante até então desenvolvido nesta localidade, que pudéssemos comparar com nossos resultados, buscamos alguns trabalhos realizados em Alagoas para comentá-los.

Silvia (2019), analisando a variação entre o *tu* e *você*, na cidade de Coité do Nônia (interior de Alagoas), com diálogos entre dois informantes, coletou 18 diálogos com faixa etária entre 18 a 35 anos e faixa etária entre 40 a 55 anos, totalizando 36 informantes. Obteve um total de 520 realizações que apresentaram 57 realizações do pronome *tu* e 397 realizações do pronome *você/cê*, respectivamente, 11% do *tu* contra 89% do *você/cê*.

A pesquisa realizada com falantes maceioenses, desenvolvida por Vitório (2018), teve como objetivo principal estudar os pronomes *tu* e *você* na função de sujeito. O estudo obteve como resultado um total de 404 realizações do *tu* e *você*, distribuídas da seguinte forma: 6 realizações do pronome *tu* e 398 realizações do pronome *você*. Esses dados representam percentuais de 2% do *tu* versus 98% do *você*.

É oportuno assinalar que os Estados do Rio Grande do Norte e Sergipe não apresentaram variação do fenômeno em análise, sendo quase categórico o uso do *você* nessas localidades, porém achamos pertinente fazer algumas considerações de cunho qualitativo.

4.2.4 Rio Grande do Norte

Para este Estado, os dados foram coletados nas localidades: Natal, Mossoró, Angicos, Pau dos Ferros e Caicó. Os resultados gerais para o Rio Grande do Norte estão dispostos na Tabela 30.

Tabela 30: Distribuição dos pronomes *você* e *tu* no Rio Grande do Norte

| PRONOMES | OCORRÊNCIA/TOTAL | % |
|-------------|------------------|-------------|
| <i>Você</i> | 219/220 | 99,4 |
| <i>Tu</i> | 1/220 | 0,6 |

Fonte: Elaborada pela autora.

Os dados apresentados revelam que, no Estado do Rio Grande do Norte, os falantes preferem empregar o pronome *você* nas suas interlocuções (99,4%), conforme números da Tabela 30. Das 220 realizações de pronomes, obteve-se apenas 1 (uma) utilização do *tu*, conforme o exemplo seguinte:

INF. – *Tu* quer tomar leite ou café? (52/3)⁶⁵

Trata-se de um caso em que o pronome *tu*, empregado com referênciação específica, apareceu numa questão de prosódia, usado com a função sintática de sujeito, no presente do indicativo. O dado foi documentado no interior do Estado, na cidade de Angicos. O *tu* foi produzido por um homem da faixa etária 2 e com nível fundamental de escolaridade.

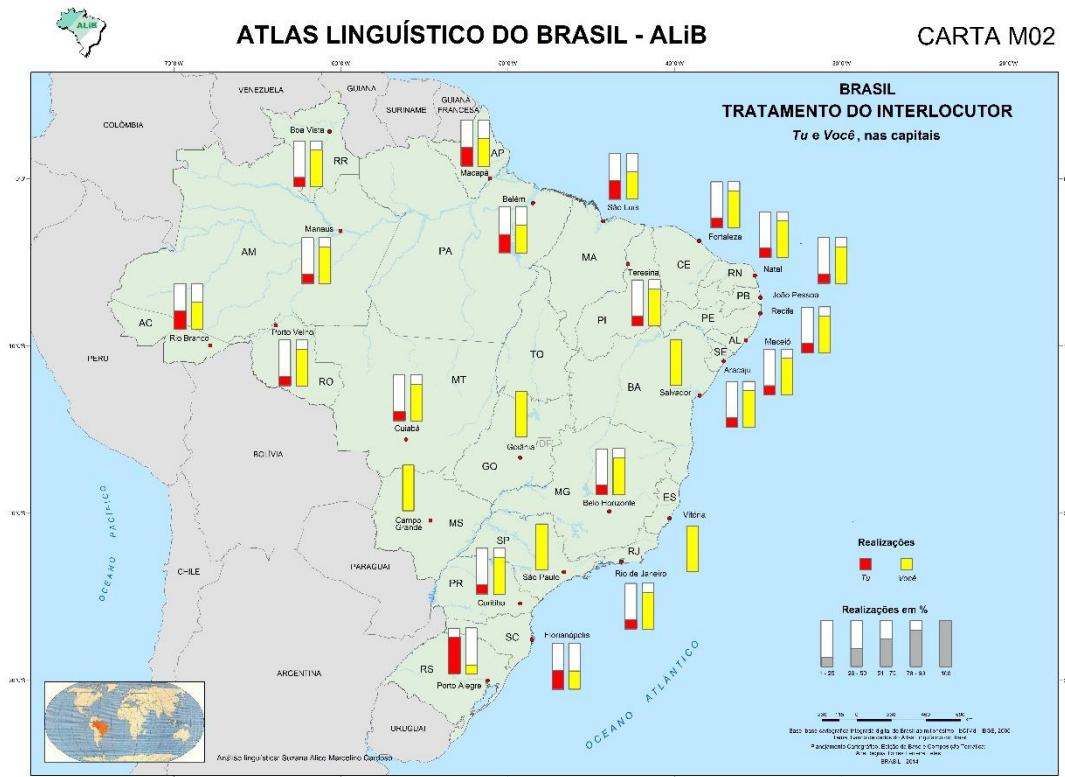
O informante em questão, na sua entrevista, utilizou 3 ocorrências do *você* e uma do *tu*, podendo sugerir que essa variação na fala do mesmo está atrelada ao “tipo de relação” estabelecida com seu interlocutor, o que pode estar atuando para a mudança estilística do falante, ou seja, o tipo de relação existente entre os interlocutores condiciona o uso da forma linguística escolhida pelos mesmos. Nesse caso, a pergunta foi direcionada a alguém (receptor) conhecido, de relações simétricas.

Para Bell (1994), quando o falante escolhe um estilo é porque ele acomodou sua fala a determinadas características do seu interlocutor. Isto posto, conscientes da mudança e motivados, sobretudo, pela relação com seus interlocutores, o falante presta mais atenção a sua produção verbal. “Pode-se resumir este processo, argumentando que o interlocutor é um dos fatores que determina o grau de pressão comunicativa que incide sobre o falante” (BORTONI-RICARDO, 2002, p. 335).

65 Informante de Angicos-RN. Sexo masculino. Faixa etária II. Nível fundamental

A fim de ilustrarmos melhor a baixa frequência da variante *tu* nas capitais do Nordeste, vejamos a Figura 17, que apresenta o tratamento do interlocutor (com *tu* ou *você*) nas capitais:

Figura 17: Carta M02: Tratamento do interlocutor (com *tu* ou *você*) nas capitais



Fonte: Cardoso et al. (2014, p. 349).

Visando a uma melhor compreensão do fenômeno nas capitais analisadas nesta tese, observemos na Carta M02 que, nas cidades de Natal, João Pessoa, Maceió, Aracajú e Recife, a realização da variante *tu* se encaixa entre 1-25%. Somente em São Luís que a realização desta variante está acima de 25%.

4.2.4.1 Referenciação

No que concerne ao uso do *você/tu* e o tipo de referência expressa, obteve-se o seguinte resultado:

Tabela 31: Distribuição dos pronomes *você* e *tu* por Referenciação no Rio Grande do Norte

| REFERENCIAÇÃO | VOCÊ | | TU | |
|---------------|-----------|-------------|-----------|-----|
| | Nº./total | % | Nº./total | % |
| Específica | 153/154 | 99,3 | 1/154 | 0,7 |
| Genérica | 66/66 | 100 | 0/66 | 0 |

Fonte: Elaborada pela autora.

Conforme disposição da Tabela 30, os resultados indicam que os falantes do Rio Grande do Norte preferem utilizar o *você*, uma vez que das 154 ocorrências de referência específica, obtivemos apenas 1 utilização do *tu*. Por outro lado, evidenciamos uma utilização de 100% da referenciação genérica com o *você*, ou seja, os falantes potiguares não utilizam o *tu* de forma genérica.

Nossos resultados encontram também um ponto de congruência com outro estudo realizado na cidade de Natal, a pesquisa de Silva (2015), que constituiu um *corpus* de conversações gravadas em diferentes ambientes sociais. A pesquisadora selecionou 8 conversas, das quais obteve 62 dados do *tu* (16%) e 316 do *você* (84%) na função de sujeito.

Em relação à variável tipo de inquérito, apenas as questões de prosódia apresentam um uso expressivo do pronome *você* com um percentual do 97,7% e 4,3% do *tu*. Os demais questionários evidenciaram, categoricamente, o uso do *você*.

A variável sexo demonstra que o fator feminino empregou, categoricamente, o pronome *você*, enquanto o fator masculino obteve um percentual de 99,2% de uso do *você*. Comparando esses dados com os dados do Estado de Alagoas, encontramos uma maior utilização do *você* no sexo masculino.

No que concerne à amostra do Rio Grande do Norte em relação à faixa etária, verifica-se que a faixa 1, categoricamente, utilizou o pronome *você*, já os falantes com idade mais avançada utilizaram 0,8% do *tu*, isto é, nas 118 ocorrências, há apenas uma utilização do *tu*.

Em relação à escolaridade, registramos o uso do *você*, categoricamente, pelos falantes de nível universitário (100%), enquanto os falantes de escolaridade fundamental utilizaram 0,5% do *tu*. Comparando os dados dessa variável com dados de escolarização de Sergipe, coincidentemente, obteve-se 100% de uso do *você* por falantes de nível universitário.

Diferentemente dos resultados aqui apresentados, a pesquisa de Guimarães (2014), realizada no Ceará, com a norma popular, apontou a escolaridade como a décima primeira

variável que mais influencia a regra. Os resultados alcançados para este grupo de fatores revelam que o *tu* só é favorecido, embora discretamente, pelos falantes de maior escolaridade (54,2% e 0,546), já que os de menor escolaridade (0-4 anos: 46,9% e 0,451; 5-8 anos 49,7% e 0,478) inibem o seu uso.

4.2.5 Sergipe

A variação espacial, também conhecida como regional ou dialetal, vem corroborar o fornecimento de dados sobre o português brasileiro. Desse modo, almejando mapear linguisticamente o Estado de Sergipe, no que concerne à variação dos pronomes *tu* e *você*, elencamos na Tabela 32 os resultados gerais para as três localidades examinadas: Aracaju, Estância e Propriá.

Tabela 32: Distribuição dos pronomes *tu* e *você* em Sergipe

| PRONOMES | OCORRÊNCIA/TOTAL | % |
|-------------|------------------|-------------|
| <i>Tu</i> | 2/249 | 0,8 |
| <i>Você</i> | 247/249 | 99,2 |

Fonte: Elaborada pela autora.

Explorando os dados do Estado de Sergipe, confirmamos 249 ocorrências de pronomes de 2^a pessoa, das quais 247 foram do *você* (99,2%) e apenas 2 do *tu* (0,8%). Nota-se que no referido Estado, a aplicação da regra do pronome *você* é quase categórica. Os únicos dados do *tu*, listados a seguir, ocorreram na cidade de Estância:

- (i) INF. – ... eu vou falá com meu pai, *tu* vai ficar na minha... na casa do meu pai comigo, meu pai não tem mulher mora sozinho lá eu e ele (80/1)⁶⁶
- (ii) INF. – *Tu* vai... Fulana, *tu* vai na rua agora? (80/4)⁶⁷

Devido à má distribuição dos dados, uma vez que 99,2% dos dados foram do pronome *você*, o GoldVarb X não selecionou nenhuma variável como estatisticamente significativa.

⁶⁶ Informante do sexo masculino. Faixa etária I. Morador de Estância- SE. Nível fundamental.

⁶⁷ Informante do sexo feminino. Faixa etária II. Moradora de Estância- SE. Nível fundamental.

Assim, alguns resultados serão apresentados a seguir com seus percentuais e outros, de maneira qualitativa.

Diferentemente dos resultados encontrados no Rio Grande do Norte e em Sergipe, temos a pesquisa de Franceschini (2011), na cidade de Concórdia (SC), que apresentou 55% para uso do *tu* e 45% para *você*, revelando, portanto, o predomínio do pronome *tu*.

4.2.5.1 Referenciação

Encontram-se, na Tabela 33, os percentuais referentes à distribuição das variantes quanto à referenciação específica e genérica no Estado de Sergipe.

Tabela 33: Distribuição dos pronomes *tu* e *você* com relação à Referenciação em Sergipe

| REFERENCIAÇÃO | TU | | VOCÊ | |
|---------------|-----------|-----|-----------|-------------|
| | Oc./Total | % | Oc./Total | % |
| Específica | 2/227 | 1,8 | 225/227 | 98,2 |
| Genérica | 0/22 | 0 | 22/22 | 100 |

Fonte: Elaborada pela autora.

Como se pode depreender dos dados, no Estado de Sergipe, encontramos 227 ocorrências com referenciação específica, das quais 225 foram do pronome *você* (98,2%) e 2 aplicações do *tu* (1,8%). Por outro lado, com referenciação genérica ocorreu apenas o emprego do pronome *você* (22), assim como aconteceu no Estado de Alagoas. Logo, a preferência dos falantes sergipanos em relação à referência específica, quanto à genérica é pela variante *você*.

Sobre as outras variantes, como salientamos anteriormente, as nossas expectativas eram as seguintes: no tipo de questionário, acreditávamos que nos discursos semidirigidos haveria maior utilização do *você* e nas questões de prosódia, maior uso do *tu*. As duas ocorrências da variante *tu* aconteceram no QMS.

Em relação à variável função sintática, encontramos 2 ocorrências da variante *tu* no tempo verbal não passado (presente, futuro e imperativo), 196 ocorrências do *você* no tempo não passado e 51 ocorrências do *você* no passado.

No que tange à variável faixa etária, das 247 ocorrências, obteve-se, na faixa II, 128 falantes utilizando a variante *você* e 119 na faixa I. Por outro lado, obteve-se 2 ocorrências da

variante *tu*, na faixa I.

Observou-se, na análise da variável escolaridade, que, de maneira categórica, todos os falantes de nível universitário empregaram o *você*.

4.2.5.2 Sexo

Encontram-se, na Tabela 34, os percentuais referentes à distribuição da variável sexo no Estado de Sergipe.

Tabela 34: Distribuição dos pronomes *tu* e *você* por Sexo do informante em Sergipe

| SEXO | TU | | VOCÊ | |
|-----------|-----------|-----|-----------|-------------|
| | Oc./Total | % | Oc./Total | % |
| Masculino | 2/116 | 1,7 | 114/116 | 98,3 |
| Feminino | 0/133 | 0,0 | 133 | 98,3 |

Fonte: Elaborada pela autora.

De acordo com a Tabela 34, a variável sexo mostra-se bem equilibrada, ou seja, homens e mulheres utilizam o *você* nas suas interlocuções. Temos, respectivamente, um favorecimento do *você* de 98,4% na fala masculina e 98,3% na fala feminina.

Os nossos dados opõem-se aos resultados obtidos por Alves (2015) em seu estudo feito em São Luiz (Maranhão). O mesmo objetivou analisar o comportamento linguístico do falante ludovicense escolarizado em situações naturais de interação social, buscando entender o significado social que as formas *tu* (com e sem concordância), *você*, *cê*, *senhor/a* têm na comunidade.

Seus dados totalizaram 1.110 ocorrências, dos quais 741 *tu* sem concordância e 130 *tu* com concordância e apenas 175 do *você/cê*. Percebemos, assim, que para representar a segunda pessoa, os falantes ludovicense conservam a presença viva e marcante da forma *tu*. A autora afirma que o *tu* é um pronome típico do falar ludovicense, e por que não dizer, marca identitária da comunidade, visto sua alta ocorrência mesmo em entrevistas formais estruturadas entre um documentador e um informante.

4.3 SÍNTESE DO CAPÍTULO

Neste capítulo, foram apresentados e discutidos os resultados do *corpus*, composto por 140 informantes dos dois sexos, divididos em duas faixas etárias. O estudo contempla cinco fatores linguísticos estruturais – referênciação, tipo (parte) do questionário, função sintática do pronome, tempo verbal e paralelismo; e quatro fatores extralingüísticos – localidade, sexo, faixa etária e escolaridade.

O Quadro 12 mostra, de maneira mais condensada, a quantificação dos dados em referência à 2^a pessoa, nos estados analisados.

Quadro 12: Total de ocorrências das variantes: *tu* e *você* por Estado

| ESTADO | TU | VOCÊ | TOTAL |
|---------------------|-----|-------|-------|
| Rio Grande do Norte | 1 | 219 | 220 |
| Paraíba | 89 | 336 | 425 |
| Pernambuco | 117 | 611 | 728 |
| Alagoas | 17 | 356 | 373 |
| Sergipe | 2 | 247 | 249 |
| | 226 | 1.769 | 1.995 |

Fonte: Elaborado pela autora.

Tivemos um total de 1.995 dados de ocorrências de pronomes de segunda pessoa analisadas, das quais 1.769 foram do pronome *você* e 226 realizações do pronome *tu*. Foram feitas duas análises: uma geral, incluindo todos os Estados considerados, e outra, para cada Estado separadamente. Devido ao uso categórico do pronome *você*, os grupos de fatores excluídos da análise geral da amostra foram: função sintática do pronome, tempo verbal, paralelismo, sexo e faixa etária.

Na análise geral, o programa GoldVarb X, a partir da rodada binária, apontou como variáveis relevantes no emprego do *tu*: referênciação, parte do inquérito, localidade e escolaridade. Em relação às rodadas individuais (por Estado), o programa selecionou os Estados de Pernambuco, Alagoas e Paraíba.

No Estado do Rio Grande do Norte, apenas no interior de Angicos encontramos uma ocorrência do pronome *tu*. Já no Estado de Sergipe, registramos duas ocorrências do pronome *tu*, na cidade de Estância. Isto posto, justifica-se a exclusão destes Estados na rodada do programa GoldVarbX.

Na seção seguinte, sintetizamos, como conclusões, os resultados a que chegamos a partir das análises dos dados coletados dos cinco estados nordestinos.

CONSIDERAÇÕES FINAIS

Conforme propusemos no Capítulo de introdução desta tese, objetivamos investigar o uso das formas de tratamento *tu* e *você* em cinco Estados da região Nordeste do país, a partir da análise de um extrato do *corpus* do ALiB, e depreender quais os fatores linguísticos e sociais que estão em jogo no processo de escolha do falante.

As opiniões divergem muito sobre o emprego das formas de tratamento. Muitos pesquisadores postulam que as formas de tratamento no Brasil estão reduzidas a duas: *você* como forma de intimidade, usada no tratamento familiar; e *o senhor*, como forma de respeito, usada no tratamento ceremonioso (DIVINO, 2008). A forma *tu* é vista como um uso relativamente restrito, mas frequente em algumas regiões do Norte e no extremo Sul do país.

Recorremos aos pressupostos teóricos da Dialetologia Pluridimensional Contemporânea e da Sociolinguística Variacionista, que priorizam os princípios gerais para o estudo da mudança linguística descritos por Weinreich, Labov e Herzog (2006 [1968]), ou seja, a variação é inerente à língua, sistemática e condicionada por fatores intra e extralinguísticos; apresenta funções sociais; e é fundamental para que a mudança linguística ocorra de maneira ordenada.

No Capítulo 1, apresentamos um histórico das formas de tratamento, discorremos sobre os pronomes de tratamento utilizados pelos falantes brasileiros e elencamos alguns trabalhos sobre referência de 2^a pessoa no PB – baseados tanto em língua oral quanto em língua escrita – selecionados e caracterizados por região. No Capítulo 2, tratamos da parte teórica que serviu de base para a análise dos dados, situando nossa pesquisa na linha de estudos da Geolinguística e da Sociolinguística. Enfocamos as áreas dialetais no Brasil, bem como enumeramos os atlas linguísticos concluídos e em andamento. No Capítulo 3, expusemos a metodologia empregada na pesquisa e mostramos o suporte estatístico utilizado para análise dos dados. No Capítulo 4, apresentamos a análise dos dados e os resultados de cálculos estatísticos obtidos com o auxílio do programa GoldVarb X, que confirmaram algumas hipóteses e refutaram outras. Apontamos os fatores sociais e linguísticos que têm efeito na escolha dos falantes quanto aos pronomes *tu* e *você*.

No tocante aos objetivos específicos – observar como as formas *tu* e *você* estão distribuídas em cada localidade, verificar como essas formas se realizam nos diferentes questionários do Projeto ALiB, bem como verificar o efeito da escolaridade, sexo e faixa etária, quanto à utilização dos pronomes –, seus resultados foram alcançados no decorrer da análise e são expostos juntamente com a testagem das hipóteses, a seguir.

Retomando os objetivos específicos levantados para esta pesquisa, convém dizer que confirmamos que o *você* é a forma preferida pelos falantes nordestinos dos cinco Estados analisados. Porém buscamos analisar os contextos nos quais a variante *tu* ainda resiste. Essa variante se fez presente em interações mais informais, regidas por relações simétricas.

É interessante salientar que a forma *tu*, no Nordeste, aparece associada ao verbo na terceira pessoa, acarretando, assim, o estigma da variante. Nesse sentido, o traço de mais intimidade, vinculado a essa variante, traz uma avaliação social negativa e, portanto, é a forma não usada pelos falantes que pretendem ser mais formais.

Vale destacar que os resultados do estudo aqui apresentado derivam de uma investigação em *corpus* de língua oral (oriundo do Projeto ALiB, como detalhado no Capítulo 3), cumprindo, assim, o objetivo primordial das pesquisas de cunho dialetológico e sociolinguístico: contribuir para o conhecimento da realidade linguística do PB, do ponto de vista da sua distribuição espacial (variação diatópica) e social (variação diastrática, diageracional e diagenérica).

No tocante à interpretação dos dados, foram feitas duas análises: uma geral, incluindo todos os Estados estudados, e outra de cada Estado, separadamente. Em cada análise, realizamos duas rodadas no GoldVarb X: a primeira, que forneceu a frequência das variantes na amostra e a distribuição das mesmas em função das localidades; e a segunda, a rodada binária para a geração dos pesos relativos.

Constatamos que, nos Estados de Sergipe, Alagoas, Rio Grande do Norte, Paraíba e Pernambuco, obtivemos um total de 1.995 ocorrências de pronomes de segunda pessoa, sendo 1.769 contextos com o emprego de *você* (88,7%) e 226 usos de *tu* (11,3%). Em termos gerais, a distribuição do fenômeno estudado revelou que, nos cinco Estados selecionados para a análise, há uma significativa preferência pela forma *você*, que teve aplicação bastante alta. Nesse sentido, comprovou-se uma das hipóteses aventadas inicialmente: a predominância de *você* em relação a *tu*.

Mesmo com um número reduzido do pronome *tu* na amostra investigada, focamos a nossa análise em elucidar quais os contextos em que esse pronome ainda perdura.

Os grupos de fatores ausentes na análise geral da amostra foram: função sintática do pronome, tempo verbal, paralelismo, sexo e faixa etária. Estes não revelaram resultados significativos e, por isso, foram excluídos da análise. Assim, extraídas as variáveis/variantes que apresentaram nocaute, a rodada binária do GoldVarb X apontou como variáveis relevantes

no emprego de *tu*: referencião, parte do inquérito, localidade e escolaridade. Foram selecionadas, portanto, variáveis linguísticas e sociais.

Em relação às hipóteses inicialmente levantadas para o desenvolvimento desta pesquisa, chegamos às seguintes conclusões:

- Em nossa análise geral, com os cinco Estados, no que concerne à variável referencião, os dados evidenciaram que das 1.588 ocorrências de pronomes com referencião específica, 225 foram do *tu* (0,711) e 1.363, do *você*. Em relação à referencião genérica, obtivemos 407 ocorrências, das quais apenas uma aparece com o pronome *tu*. Para testar essa variável, partimos da hipótese de que em referência genérica os usos mais recorrentes são do pronome *você*, por acreditar que o falante, ao relatar um acontecimento, afasta-se do fato narrado e, consequentemente, do seu envolvimento direto com o discurso, priorizando a forma inovadora. Diante disso, a nossa hipótese foi confirmada.
- Sobre a variável tipo de questionário, reunindo os Estados selecionados, encontramos uma maior utilização do *tu* nas questões metalinguísticas, com peso relativo de 0,777; das 59 ocorrências de pronomes de 2^a pessoa, houve 16 realizações do *tu*. Por sua vez, no QMS, conforme revelaram os resultados, obtivemos peso relativo de 0,615, isto é, das 742 ocorrências, 113 realizações foram do pronome *tu*. No discurso semidirigido, ocorreu um favorecimento de 0,567 (peso relativo) para realização do *tu*. A nossa expectativa foi a de que nos discursos semidirigidos acontecesse uma maior utilização de *você*, especificamente *você* genérico, uma vez que os relatos pessoais e não pessoais tendem a narrar informações nas quais sempre aparece *você* indeterminado (genérico). Já nas questões de prosódia, pelo fato de a pergunta ser direcionada a um amigo, pessoa próxima, acreditamos na predominância do *tu*. Isto posto, chegamos à conclusão de que a nossa hipótese em relação ao pronome *você* foi confirmada. Já a hipótese em relação ao *tu* foi rejeitada.
- Conforme revelaram os dados analisados, a respeito da amostra de fala dos cinco Estados do Nordeste, sobre a variável localidade, em algumas cidades, não foi localizada nenhuma ocorrência do pronome *tu*, a saber: Rio Grande do Norte (Mossoró, Pau dos Ferros, Caicó, Natal); Alagoas (União dos Palmares, Arapiraca); Sergipe (Aracaju e Propriá) e Pernambuco (Arco Verde). Por outro lado, a cidade de Cabrobó, em Pernambuco, foi a localidade em que o uso do *tu* foi mais favorecido, com peso relativo de 0,78 e frequência de 31,7%. Em seguida, as cidades de Itaporanga, na Paraíba, e

Afrânio, em Pernambuco, apresentaram pesos relativos bem significativos, respectivamente, 0,77 (frequência de 35,5%), 0,717 (frequência de 22,9%). Defendemos a hipótese de que os falantes do interior utilizariam mais o pronome *tu*, enquanto os falantes da capital prefeririam o *você*; portanto a hipótese não se sustenta pelos resultados obtidos, pois tanto nas capitais como no interior a regra de aplicação do *você* foi bastante alta.

- Em relação à escolaridade, nos cinco Estados, houve 0,523 de peso relativo de utilização do *tu* pelos falantes do Ensino Fundamental, ou seja, das 1.720 ocorrências de referência à segunda pessoa, 216 realizações foram do *tu*. Dessa forma, a nossa crença de que os informantes com maior escolaridade tendem a falar mais o *você* do que o *tu* foi comprovada.

Num segundo momento, fizemos as rodadas individuais, através do programa GoldVarb X, nos Estados de Pernambuco, Paraíba e Alagoas, controlando as dez variáveis escolhidas, a saber: função sintática do pronome, tempo verbal, paralelismo, sexo, faixa etária, referênciação, parte do inquérito, faixa etária e localidade, testando, assim, todos os grupos inicialmente propostos, tanto linguísticos como os extralingüísticos.

Em Pernambuco, foram apontadas como mais relevantes as variáveis: referênciação, parte do inquérito, faixa etária e localidade. A partir dos resultados, evidenciamos que:

- A variável referênciação foi favorecida pela referência específica, a qual apresentou peso relativo de 0,690 e frequência de 19,9% de uso do *tu*.
- A variável parte do inquérito teve maior realização do *tu* nas questões do discurso semidirigido, com peso relativo de 0,616 e 18,2% de uso; em seguida, no QSL; e depois, no QMS.
- A variante faixa etária I destacou-se com peso relativo 0,569, embora bem próximo à faixa etária II (0,444).
- A variável localidade revelou favorecimento da utilização do pronome *tu* nas cidades pernambucanas de Cabrobó, Afrânio, Salgueiro e Limoeiro. No total de 717 ocorrências de pronomes de segunda pessoa, 117 foram da variante *tu* nas elocuções, obtendo 16,3%. A liderança foi de Cabrobó com, 0,764 de peso relativo.

Sobre a variação diatópica nas cidades de Cabrobó e Salgueiro, destacamos o fato de essas cidades estarem localizadas bem próximas, na mesorregião do sertão pernambucano, uma área remanescente de quilombos e descendentes indígenas, fato que, provavelmente, deixou resquícios na linguagem.

Podemos reconhecer que, nos relatos pessoais, emitidos nas falas pernambucanas, os falantes participantes de contextos de mais envolvimento e menos monitoramento utilizaram o pronome *tu*. São situações carregadas de expressividade e emoção; relatos provindos de momentos descontraídos, com fortes traços de informalidade. Nessa direção, retomamos as palavras de Bortoni-Ricardo (2004, p. 63), a qual afirma que “com um mesmo interlocutor, o estilo poderá tornar-se mais ou menos monitorado em função do alinhamento que assumimos em relação ao tópico e ao próprio interlocutor”.

No que concerne ao Estado da Paraíba, das 425 ocorrências de pronomes de segunda pessoa, 336 apresentaram o uso do *você* (79,1%) e apenas 89 de emprego do *tu* (20,9%). Os grupos de fatores mais significativos foram parte do inquérito, faixa etária e localidade, com os seguintes resultados:

- A variante *tu* foi mais usada nas questões metalinguísticas (0,921);
- A faixa etária II deu preferência pela variante *tu* (0,567); e
- A localidade de Itaporanga apresentou um maior uso do *tu* (0,685).

Em relação ao Estado de Alagoas, das 374 ocorrências de pronomes de 2^a pessoa, 356 foram do *você* (95,4%) e 17 foram do *tu* (4,6%). Mostraram-se como estatisticamente significativas as variáveis:

- Sexo, com 0,725 de peso relativo por parte das mulheres; e
- Tipo de questionário, com 0,933 de peso relativo nas questões metalinguísticas.

A nossa hipótese em relação à interferência da variável sexo segue o proposto por Labov (2008 [1972]), reconhecendo que as mulheres preferem a variante de prestígio *você*. Nessa direção, a hipótese foi confirmada no Estado de Alagoas, uma vez que, das 198 ocorrências, 14 foram da variante *tu* e 184 da variante *você*.

Os Estados do Rio Grande do Norte e Sergipe não apresentaram variação. No Rio Grande do Norte obtivemos 220 ocorrências de *você* (99,4%) e apenas 1 ocorrência do pronome *tu* (0,6%), empregado na cidade de Angicos, por um falante do sexo masculino, faixa etária II, de nível fundamental de escolaridade. Já em Sergipe, das 249 ocorrências de pronomes, 247 são da variante *você* (99,2%) e 2 da variante *tu* (0,8%), na cidade de Estância.

Isto posto, fazemos aqui algumas reflexões de caráter conclusivo. Os falantes dos Estados em questão, que utilizaram a variante *tu*, também fizeram uso do *você*, de maneira majoritária, o que fica evidente que a variação estilística, nesses casos, está atrelada aos fatores contextuais. Os contextos discursivos nos quais houve realização da variante *tu* se configuram em interlocuções em que as relações são assimétricas (informante e entrevistador) e, por outro

lado, simétricas (informante e pessoa próxima/intima). Assim, além da atenção prestada à fala (monitoramento) e da projeção que o falante faz de seu ouvinte, há outros fatores que entram no processo de escolha estilística feita pelo indivíduo.

Por não termos objetivado controlar, criteriosamente, a variação *tu* e *você* no indivíduo, a atenção para esse fenômeno aconteceu no momento de análise dos dados, quando estávamos revendo alguns excertos de fala. Pensamos, então, que esse tipo de análise pode trazer resultados de grande valia, sobretudo no que se refere ao efeito da mudança de estilo linguístico, uma vez que a variação estilística de um falante é caracterizada pela variação de alternância de estilos e da quantidade de atenção que o falante dá à sua fala.

Lembremos, assim, das palavras de Labov (2010, p. 189), quando afirma que “há muito a ser aprendido com o estudo da variação individual, em ver como os indivíduos fazem uso da estrutura complexa de variação da comunidade ao evocarem diferentes identidades sociais”. Segundo o autor, estudos como esses têm de ir além da descrição de atos individuais e observar como uma pessoa muda seu estilo de uma situação para outra. Por ora, lembramos Bell (1984), quando defende que a investigação de estilo exige seus próprios *designs* e metodologia: para progredir além do nosso nível atual de conhecimento. A pesquisa sobre estilo precisa ser adaptada para esse fim, não apenas enxertada em um estudo com outros objetivos principais. Percebemos, portanto, que os caminhos se abrem para novas perspectivas, novos estudos.

Sobre as contribuições da Sociolinguística e da Dialetologia, lembramos que o estudante de Letras deve realizar pesquisas sobre variação linguística ao longo de sua formação, tanto na área da Sociolinguística quanto da Geolinguística. Os professores que conhecem a diversidade linguística estão habilitados a investigar e a descrever as variedades do português em uso nas comunidades onde seus alunos residem, podendo, portanto, posicionar-se de maneira crítica em relação ao ensino da língua. Dessa maneira, poderão contribuir para o desenvolvimento de uma pedagogia sensível às diferenças dialetais, sociolinguísticas e culturais dos alunos.

Acreditamos que a nossa pesquisa apresenta dados e resultados que podem contribuir para uma descrição mais acurada do fenômeno da variação *tu/você* no PB, mais especificamente no Nordeste. Entretanto novas pesquisas poderão ampliar as explicações para a variação aqui investigada, usando, por exemplo, uma amostra com novos critérios de coleta de dados. Desse modo, poderemos obter mais evidências capazes de testar nossas conclusões, como, por exemplo, o tratamento por *você* no interior nordestino.

REFERÊNCIAS

- AGUILERA, V. A. **Atlas Linguístico do Paraná** (ALPR). Curitiba: Imprensa Oficial do Estado, 1994.
- ALMEIDA, N. L. F. Urbanização, escolarização e variação linguística em Feira de Santana-Bahia (Século XX). **Tabuleiro de Letras**. UNEB/ DCH I, n. 4, junho 2012. Disponível em: <http://www.revistas.uneb.br/index.php/tabuleirodeletras/article/view/170/120>. Acesso em: 17 dez. 2018.
- ALVES. C. C. B. **Pronomes de segunda pessoa no espaço maranhense**. 2015. 153f. Tese (Doutorado em Linguística) – Universidade de Brasília, Brasília, 2015.
- AMARAL, A. **O dialeto caipira**. 3. ed. São Paulo: Hucitec, 1976 [1920].
- ANDRADE, C. Q. **Tu e mais quantos?** A segunda pessoa na fala brasiliense. 2010. 139f. Dissertação (Mestrado em Letras e Linguística) – Programa de Pós-Graduação em Linguística, Instituto de Letras, Universidade de Brasília, Brasília, 2010.
- ARAÚJO, L. E. S. A variável faixa etária em estudos sociolinguísticos. **Estudos Linguísticos de São Paulo**, v. XXXVI, n. 2, p. 389-398, maio-agosto 2007, p. 389-398. Disponível em: <http://www.gel.org.br/estudoslinguisticos/edicoesanteriores/4publica-estudos-2007/sistema06/71.PDF>. Acesso em: 20 jan. 2018.
- ASSUNÇÃO, J. S.; ALMEIDA, N. L. F. **A realização do “tu” e “você” na variedade linguística de falantes feirenses**. In: Anais do XII Seminário de Iniciação Científica da UEFS, Feira de Santana: UEFS, p. 1-3, 2008.
- AZEREDO, J. C. **Gramática Houaiss da língua portuguesa**. 3. ed. São Paulo: Publifolha, 2011 [2008].
- BAGNO, M. A inevitável travessia: da prescrição gramatical à educação linguística. In: BAGNO. M.; STUBBS, M.; GAGNÉ, G. **Língua materna, letramento, variação e ensino**. São Paulo: Parábola, 2002. p. 13-84.
- BECHARA, E. **Moderna gramática portuguesa**. 37. ed. Rio de Janeiro: Lucerna, 2004.
- BELL, A. Language style as audience design. **Language in Society**, v. 13, n. 2, p. 145-204, 1984.
- BENVENISTE, E. **Problemas de linguística geral**. São Paulo: EDUSP, 1976.
- BEZERRA, M. A. Uso de tu/você em interações infantis. **Revista Letras**. PUCCAMP, Campinas, v. 1, n. 13, p. 96-118, dezembro 1994.

- BIDERMAN, M. T. C. Formas de Tratamento e Estruturas Sociais. **Alfa: Revista de Linguística**. São Paulo, n. 18, p. 339-381, 1972-1973.
- BORTONI-RICARDO. S. M. **Nós chegou na escola, e agora?**: sociolinguística e educação. São Paulo: Parábola, 2005.
- _____. **Educação em língua materna**: a sociolinguística na sala de aula. São Paulo: Parábola, 2004.
- _____. Um modelo para a análise sociolinguística do português do Brasil. In: BAGNO, M. **Linguística da norma**. São Paulo: Edições Loyola, 2002. p. 333-350.
- BRANDÃO, S. F. **A geografia linguística no Brasil**. São Paulo: Ática, 1991.
- BRASIL. Ministério da Educação e do Desporto. Secretaria de Educação Fundamental. **Parâmetros curriculares nacionais** – terceiro e quarto ciclos do ensino fundamental: Introdução aos parâmetros curriculares nacionais. Brasília, DF: MEC/SEF, 1998. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/introducao.pdf>. Acesso em: 5 jun. 2019.
- BRIGTH, W. As dimensões da sociolinguística. In: FONSECA, M. S. V.; NEVES M. F. (org.). **Sociolinguística**. Rio de Janeiro: Livraria Eldorado Tijuca, 1974. p. 17-22.
- BROWN, R.; GILMAN, A. The pronouns of power and solidarity. In: SEBEOK, T. A. (Ed.). **Style in Language**. New York/Cambridge: MIT Press, 1960. p. 253-276.
- BUSSE, S. **Um estudo Geossociolinguístico da fala do Oeste do Paraná**. 2010. 288f. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2010.
- CALLOU, D. Quando Dialectologia e Sociolinguística se encontram. **Estudos linguísticos e literários**. Salvador, n. 41, p. 29-48, janeiro-junho 2010.
- _____; LEITE, I. **Como falam os brasileiros**. R. de Janeiro: Jorge Zahar Editor, 2002.
- CALMON, E. N. **Ponte da passagem**: você e cê transitando na fala de Vitória (ES). 2010. 139f. Dissertação (Mestrado em Linguística) – Universidade Federal do Espírito Santo, Vitória, 2010.
- CARDOSO, S. A. M. Projeto ALiB: o sentido desta caminhada. In: MOTA, J. A. et al. (org.). **Documentos 3**: Projeto Atlas Linguístico do Brasil. Salvador: Vento Leste, 2012. p. 13-32.
- _____. **Geolinguística**: tradição e modernidade. São Paulo: Parábola, 2010.
- _____. Caminhos dos pronomes pessoais no português brasileiro: considerações a partir de dados do projeto ALiB. **Anais do Simpósio Mundial de Estudos de Língua Portuguesa**. São Paulo: USP, p. 321-345, 2008.
- _____. Dialectologia atual: tendências e perspectivas. **Revista do Gelne**. v. 5, n. 1 e 2, p. 185-192, 2003.

- _____. Dialectologia: trilhas seguidas, caminhos a perseguir. **DELTA**. São Paulo, v. 17, n. esp., p. 25-44, 2001a.
- _____. A variação diagenérica no português do Brasil: uma visão a partir de inquéritos experimentais do Projeto ALiB. **Boletim da ABRALIN**. Fortaleza, v. 26, n. esp., p. 120-122, 2001b.
- _____. A geolinguística no Brasil: meio século de contribuição à ciência da linguagem e ao ensino da língua materna. **Boletim da ABRALIN**. Florianópolis, v. 23, p. 18-34, 1999.
- _____. O atlas linguístico do Brasil: um projeto nacional. In: AGUILELA, V. A. (org.). **A geolinguística no Brasil: caminhos e perspectivas**. Londrina: UEL, 1998. p. 165-176.
- _____. (org.). **Diversidade linguística e ensino**. Salvador: EDUFBA, 1996.
- CARDOSO, S. A. M.; MOTA, J. Dialectologia e ensino da língua materna. In: GOESKI, E.; COELHO, I. L. **Sociolinguística e ensino: contribuições para a formação do professor de português**. Florianópolis: UFSC, 2006. p. 83-107.
- _____; MOTA, J. A Geolinguística no Brasil e o Projeto ALiB. **Bollettino dell'Atlante Linguistico Italiano**. Torino, n. 27, p. 225-267, 2005.
- _____; FERREIRA, C. **A dialetologia no Brasil**. São Paulo: Contexto, 1994.
- CARDOSO, S. A. M. et al. (org.). **Documentos 6: Projeto Atlas Linguístico do Brasil: histórias e memórias**. Salvador: Vento Leste, 2016.
- _____. et al. **Atlas linguístico do Brasil**. v. 1. Introdução. Londrina: EDUEL, 2014a.
- _____. et al. **Atlas Linguístico do Brasil**. v. 2. Cartas linguísticas. Londrina: EDUEL, 2014b.
- _____. et al. (org.). **Documentos 4: Projeto Atlas Linguístico do Brasil**. Salvador: Vento Leste, 2013.
- CARVALHO, N. **A Língua do Nordeste**. 2000. Disponível em: <http://intervox.nce.ufrj.br/~edpaes/linguane.htm>. Acesso em: 20 ago. 2019.
- CHAMBERS, J. K.; TRUDGILL, P. **La Dialetología**. Traducción Carmen Morán González. Madrid: Visor Libros, 1994.
- CINTRA, L. F. L. **Sobre formas de tratamento na língua portuguesa**. Lisboa: Livros Horizonte, 1972.
- COELHO, I. **Sociolinguística**. Florianópolis: LLV/CCE/UFSC, 2010.
- COELHO, M. S. V. **Uma abordagem variacionista do uso da forma você no norte de Minas**. 1999. 85f. Dissertação (Mestrado em Linguística) – Universidade Federal de Minas Gerais, Belo Horizonte, 1999.

- COMITÊ NACIONAL. **Projeto Atlas Linguístico do Brasil**. Salvador: Instituto de Letras, 1998. Disponível em: www.alib.ufba.br. Acesso em: 12 jul. 2018.
- COMITÊ NACIONAL DO PROJETO ALiB. **Questionários**. Londrina: UEL, 2001.
- CRISTIANINI, A. C. **Atlas Semântico-Lexical da região do Grande ABC**. 2007. 635f. Tese (Doutorado em Semióticas e Linguística Geral) – Faculdade de Filosofia e Ciências Humanas da Universidade de São Paulo, São Paulo, 2007.
- CUNHA, C.; CINTRA, L. **Nova gramática do português contemporâneo**. 4. ed. Rio de Janeiro: Lexikon, 2007.
- DEUS, V. G. **Você ou tu? Nordeste versus Sul**: o tratamento do interlocutor no português do Brasil a partir de dados do Projeto ALiB. 2009. 166f. Dissertação (Mestrado em Letras) – Programa de Pós-graduação em Língua e Cultura, Instituto de Letras, Universidade Federal da Bahia, Salvador, 2009.
- DIAS, E. P. **O uso do tu no português brasiliense falado**. 2007. 114f. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2007.
- DIVINO, L. S. A. **Como trato o meu receptor?** (a propósito do uso de tu/você em Santo Antônio de Jesus-BA). 2008. 170f. Dissertação (Mestrado em Letras) – Universidade Federal da Bahia, Salvador, 2008.
- FARACO, C. A. O tratamento do “você” em português: uma abordagem histórica. **Fragmenta**. Editora da UFPR, Curitiba, n. 13, p. 51-82, 1996.
- FEMERICK, R. T. **O ciclo do algodão em Caicó**. 2010. Disponível em: <http://www.tribunadonorte.com.br/noticia/o-ciclo-do-algod%C3%A3o-no-serido-i/150373>. Acesso em: 20 ago. 2019.
- FRANCESCHINI, L. T. **Variação pronominal nós/a gente e tu/você em Concórdia-SC**, 2011. 252f. Tese (Doutorado em Letras) – Universidade Federal do Paraná, Curitiba, 2011.
- _____. A influência dos fatores sociais no uso do “tu”/“você”. **Anais do I Congresso Internacional de Dialetologia e Sociolinguística**. São Luís, 2010.
- GUIMARÃES, T. A. A. S. **Tu é doido, macho!** A variação das formas de tratamento no falar de Fortaleza. 2014. 237f. Dissertação (Mestrado em Linguística Aplicada) – Programa de Pós-Graduação em Linguística Aplicada, Centro de Humanidades, Universidade Estadual do Ceará, Fortaleza, 2014.
- GUY, G. R.; ZILLES, A. **Sociolinguística quantitativa**: instrumental de análise. São Paulo: Parábola, 2007.

- HERÊNIO, K. K. P. “Tu” e “você” em uma perspectiva intra-linguística. 2006. 120f. Dissertação (Mestrado em Linguística) – Programa de Pós-graduação em Linguística, Instituto de Letras e Linguística, Universidade Federal de Uberlândia, Uberlândia, 2006.
- ILARI, R. **Gramática do Português Falado**. v. II: Níveis de Análise Linguística. São Paulo: Editora da UNICAMP, 1996.
- ILARI, R. et al. Os pronomes pessoais do português falado: roteiro para análise. In: CASTILHO, A. T.; BASILIO, M. **Gramática do Português Falado**. v. IV: Estudos descritivos. São Paulo: Editora da UNICAMP, 2002. p. 73-159.
- ISQUERDO, A. N; TELES, A. R. A rede de pontos. In: CARDOSO, S. A. M. et al. **Atlas Linguístico do Brasil**: v. I. Introdução. Londrina: EDUEL, 2014. p. 37-78.
- KARLBERG, L. G. L. **Atlas etnolinguístico do Acre – ALAC**: fronteiras léxicas. Rio Branco: Edufac, 2018.
- LABOV, W. **Padrões sociolinguísticos**. Tradução de Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].
- _____. **Principios del cambio lingüístico** – fatores sociales. Tradução de Pedro Martín Butragueño. Madrid: Gredos, 2006 [1994].
- _____. The anatomy of style-shifting. In: RICKFORD, J. R.; ECKERT, P. **Style and sociolinguistic variation**. Cambridge: Cambridge University Press, 2001. p. 85-108.
- LEITE DE VASCONCELOS, J. **Esquisse d'une dialectologie portugaise**. 2. ed. com aditamentos e correções do autor, preparada com base no exemplar conservado no Museu Etnológico Dr. Leite de Vasconcelos / Maria Adelaide Valle Cintra. Lisboa: Centro de Estudos Filológicos, 1970.
- LEMLE, M.; NARO, A. J. **Competências básicas do português**. Relatório final de pesquisa apresentado às instituições patrocinadoras Fundação Movimento Brasileiro (MOBRAL) e Fundação Ford. Rio de Janeiro, 1977.
- LOREGIAN-PENKAL, L. **(Re)análise da referência de segunda pessoa na fala da região sul**. 2004. 261f. Tese (Doutorado em Letras) – Programa de Pós-Graduação em Letras, Universidade Federal do Paraná, Curitiba, 2004.
- LUCCA, N. N. G. **A variação “tu/você” na fala brasiliense**. 2005. 126f. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Instituto de Letras, Universidade de Brasília, Brasília, 2005.

- LUCENA, R. O. P. **Pronomes possessivos de segunda pessoa:** a variação *teu/seu* em uma perspectiva histórica. 2016. 220f. Tese (Doutorado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.
- LUCCHESI, D. O conceito de transmissão linguística irregular. In: RONCARATI, C.; ABRAÇADO, J. (org.). **Português brasileiro:** contato lingüístico, heterogeneidade e história. Rio de Janeiro: 7 Letras, 2003. p.272-284.
- MARTINS, G. F. **A alternância Tu/Você/Senhor no município de Tefé - Estado do Amazonas.** Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2010.
- MATTOS E SILVA, R. V. Variação, mudança, norma e a questão do ensino no português no Brasil. In: CARDOSO, S. (org.). **Diversidade linguística e ensino.** Salvador: EDUFBA, 1996. p. 19-43.
- MENON, O. P. S. A indeterminação do sujeito no português do Brasil: NURC-SP e VARSUL. In: VANDRESEN, P. **Variação, mudança e contato linguístico no Português da Região Sul.** Pelotas: Educat, 2006a. p. 125-167.
- _____. A história de “você”. In: GUEDES, M.; BERLINCK, R. A.; MURAKAWA, C. A. A. (org.). **Teoria e análises lingüísticas:** novas trilhas. 1. ed. Araraquara, São Paulo: UNESP, 2006b. p. 99-160.
- _____. O sistema pronominal na região sul. Mesa redonda: Aspectos do sistema pronominal relevantes para o ensino básico. **Anais do XI Encontro Nacional das ANPOLL**, João Pessoa, p. 510-512, 1997.
- _____. O sistema pronominal do português do Brasil. **Revista Letras**, Curitiba: Editora da UFPR, n. 44, p. 91-106, 1995.
- MENON, O. P. S.; LOREGIAN-PENKAL, L. Variação no indivíduo: *tu/você* no Sul do Brasil. In: VANDRESSEN, P. (org.). **Variação e Mudança no português na região Sul.** Pelotas: EDUCAT, 2002. p. 147-188.
- MODESTO, A. T. T. **Formas de tratamento no português brasileiro:** a alternância “*tu/você*” na cidade de Santos-SP. 2006. 152f. Dissertação (Mestrado em Letras) – Universidade de São Paulo, São Paulo, 2006.
- MONTEIRO, J. L. **Para compreender Labov.** Rio de Janeiro: Petrópolis, Vozes, 2000.
- _____. **Pronomes pessoais:** subsídios para uma gramática do português do Brasil. Fortaleza: Edições Universidade Federal do Ceará, 1994.

- MOTA, M. A. **A variação dos pronomes ‘tu’ e ‘você’ no português oral de São João da Ponte (MG)**. Dissertação (Mestrado em Linguística). Universidade Federal de Minas Gerais, 2008.
- MUNIZ, C. Dostoiévski: biografia e resumo das principais obras. **Toda Matéria**, 2019. Disponível em: <https://www.todamateria.com.br/dostoievski/>. Acesso em: agosto de 2019.
- NARO, A. J. Modelos quantitativos e tratamento estatístico. In: MOLLICA, M. C.; BRAGA, M. L. (org.). **Introdução à sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2003. p. 15-25.
- NASCENTES, A. **Bases para elaboração do atlas linguístico do Brasil**. Rio de Janeiro: Casa de Rui Barbosa, 1958.
- _____. O tratamento de “você” no Brasil. **Letras**. Curitiba, n. 5, p. 114-122, 1956.
- _____. **O linguajar carioca**. 2. ed. Rio de Janeiro: Organização Simões, 1953.
- NOGUEIRA, F. M. S. B. **Como os falantes de Feira de Santana e Salvador tratam o seu interlocutor?** 2013. 138f. Dissertação (Mestrado em Letras) – Programa de Pós-graduação em Língua e Cultura, Instituto de Letras, Universidade Federal da Bahia, Salvador, 2013.
- OLIVEIRA, J. M. **O futuro da língua portuguesa ontem e hoje: variação e mudança**. 2006. 254f. Tese (Doutorado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.
- OLIVEIRA, J. M.; PAIM, M. M. T.; RIBEIRO, S. A importância do Atlas linguístico do Brasil para o ensino de português. **Tabuleiro de Letras** (PPGEL, Salvador, online), vol.: 12; n. 03, dezembro 2018. ISSN: 2176-5782.
- OLIVEIRA, L. A. F. O. **Tu e Você no português popular do estado da Bahia**. Comunicação ao VIII Seminário de Pesquisa e Pós-Graduação da UFBA. Salvador, 2007.
- _____. **Tu e você no português afro-brasileiro**. Comunicação ao VI Seminário de Pesquisa e Pós-Graduação da UFBA. Salvador, 2005.
- PAIM, M. M. T. **Tudo é diverso no universo**. Salvador: Quarteto, 2019.
- _____. Variação e ensino no Atlas Linguístico do Brasil. **Revista línguas e letras**. v. 17, n. 35, p. 71-84, 2016.
- PAREDES SILVA, V. L. O retorno do pronome tu à fala carioca. In: RONCARATI, C.; ABRAÇADO, J. (org.). **Português brasileiro: contato lingüístico, heterogeneidade e história**. Rio de Janeiro: 7 Letras, 2003. p.160-169.

- PEREIRA, M. N. **Atlas geolinguístico do litoral potiguar**. v. I. 2007. 123f. Tese (Doutorado em Letras Vernáculas) – Programa de Pós-graduação em Letras Vernáculas da Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2007.
- PEREIRA, R. O. **Pronomes possessivos de segunda pessoa**: a variação *teu/seu* em uma perspectiva histórica. Tese (Doutorado em Letras). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.
- PRETI, D. Formas de tratamento e papéis sociais em A ilustre casa de Ramires de Eça de Queiroz. In: BERRINI, B. (org.). **Eça de Queiroz**. A ilustre casa de Ramires. Cem Anos. São Paulo: Educ/FAPESB, 2000. p. 85-110.
- RAMOS, M. P. B. **Formas de tratamento no falar de Florianópolis**. 1989. 106f. Dissertação (Mestrado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal de Santa Catarina, Florianópolis, 1989.
- RAZKY, A. Atlas Linguístico Sonoro do Pará: uma nova perspectiva para a organização de *corpus* geolinguísticos. In: AGUILERA, V. A. (org.). **A geolinguística no Brasil**: trilhas seguidas, caminhos a percorrer. Londrina: EDUEL, 2013. p. 209-227.
- RIBEIRO, S. Projeto ALiB: 20 anos pesquisando a fala dos brasileiros. Congresso virtual UFBA 2020: Universidade em movimento. 1 vídeo (1h 29min10seg). **TV UFBA**. Salvador, 26 de maio de 2020. Disponível em: https://www.youtube.com/watch?v=QnsmXFRPbMI&fbclid=IwAR1ybMS-Sq03z1T1tFxiGxjm3gy4Nah_C1f3iJ0KU-GMVN3R-vW8NN22XTs. Acesso em: 20 maio 2020.
- _____. **Brinquedos e brincadeiras infantis na área do Falar Baiano**. 2012. 752f. Tese (Doutorado em Letras) – Programa de Pós-graduação em Língua e Cultura, Instituto de Letras, Universidade Federal da Bahia, Salvador, 2012.
- ROMANO, V. P. et al. [SGVCLin] – Software para geração e visualização de cartas linguísticas. **Estudos Linguísticos**, Belo Horizonte, v. 22, n. 1, p. 119-151, janeiro-junho 2014.
- SAID ALI, M. **Gramática Histórica da Língua Portuguesa**. 8. ed. revista e atualizada por Mário Eduardo Viaro, São Paulo: Companhia Melhoramentos, 2001 [1921].
- SALLES, M. **Pronomes de tratamento do interlocutor no português brasileiro**: um estudo de Pragmática Histórica. 2001. 246f. Tese (Doutorado em Linguística) – Universidade de São Paulo, São Paulo, 2001.
- SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. **GoldVarb X** – A multivariate analysis application. Toronto: Department of Linguistics/ Ottawa: Department of P. Mathematics, 2005.

Disponível em: http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm#ref. Acesso em: 18 jan. 2018.

SANTOS, V. M. “**Tu vai para onde?... Você vai para onde?**”: manifestações da segunda pessoa na fala carioca. 2012. 137f. Dissertação (Mestrado em Letras Vernáculas) – Programa de Pós-graduação em Letras Vernáculas da Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

SCHERRE, M. M. P. Paralelismo linguístico. **Revista de estudos da linguagem**, Belo Horizonte, v. 7, n. 2, p. 9-28, julho-dezembro 1998.

_____. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: SILVA, G. M. O.; SCHERRE, M. M. P. (org.). **Padrões sociolinguísticos** – análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Editora da UFRJ, p. 41-62, 1996.

SCHERRE, M. M. P.; NARO, A. J. Análise quantitativa e tópicos de interpretação do Varbrul. In: MOLLICA, M. C.; BRAGA, M. L. (org.). **Introdução à sociolinguística**: o tratamento da variação. São Paulo: Contexto, 2010. p. 147-177.

SCHERRE, M. M. P. et al. Variação dos pronomes tu e você. In: ABRAÇADO, J.; MARTINS, M. A. (org.) **Mapeamento Sociolinguístico do português brasileiro**. São Paulo: Contexto, 2015. p. 132-172.

SETTE, N. D. **Formas de tratamento no português coloquial**. 1980. 168f. Dissertação (Mestrado em Linguística) – Universidade Federal Pernambuco, Recife, 1980.

SILVA, F. C. **Variação entre os pronomes tu e você na função de sujeito na fala de Natal (RN)**: uma abordagem sociofuncionalista. 2015. 105f. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Federal do Rio Grande do Norte, Natal, 2015.

SILVA, L. A. Tratamentos familiares e referênciação dos papéis sociais. In: PRETI, D. (org.). **Léxico na língua oral e na escrita**. 1. ed. São Paulo: Humanitas, 2003. p. 169-194.

SILVA, S. O. P. **A variação pronominal de segunda pessoa do singular em Coité do Nôia, AL**. 2019. 126f. Dissertação (Mestrado em Letras e Linguística) – Universidade Federal de Alagoas, Maceió, 2019.

SILVA-CORVALÁN, C. **Sociolinguística**: teoría y análisis. Madrid: Alhambra, 1988.

SILVA NETO, S. **História da língua portuguesa**. 4. ed. Rio de Janeiro: Presença, 1986.

_____. **Guia para estudos dialectológicos**. Belém: Conselho Nacional de Pesquisas; Instituto Nacional de Pesquisas da Amazônia, 1957.

SILVA, S. C. **A variação dos pronomes tu e você na fala mineira de Ressaquinha (MG)**.

2017. 101f. Dissertação (Mestrado em Letras) – Universidade Federal de Ouro Preto, Mariana, 2017.

SOARES, M. E. **As formas de tratamento nas interações comunicativas**: uma pesquisa sobre o português falado em Fortaleza. 1980. 157f. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1980.

SOUZA, P. D. S.; SANTOS, A. K. B. Urbanização e monitoração estilística: a variação linguística e as representações da fala caipira nas histórias e m quadrinhos. **Fórum linguístico**. Florianópolis, v.15, n.1, p. 2 843 - 2859, jan./mar. 2018. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=6566272>. Acesso em: 12 mar. 2020.

SOUZA, V. V. **Os (des)caminhos do você**: uma análise sobre a variação e mudança na forma, na função e na referência do pronome você. 2008. 220f. Tese (Doutorado em Letras) – Universidade Federal da Paraíba, João Pessoa, 2008.

TARALLO, F. **A pesquisa sociolinguística**. São Paulo: Ática, 1985.

TELES, A. R. **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. 483f. Tese (Doutorado em Língua e Cultura) – Universidade Federal da Bahia, Salvador, 2018.

TELES, A. R.; RIBEIRO, S. Apresentando a cartografia aos linguistas: o Projeto ALiB. In: MOTA, J. A.; CARDOSO, S. A. M. (org.). **Documentos 2**: Projeto Atlas Linguístico do Brasil. Salvador: Quarteto, 2006. p. 207-228.

VALLE, C. R. M. **Sabe? ~ não tem? ~ entende?**: itens de origem verbal em variação como requisitos de apoio discursivo. 2001. 183f. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2001.

VITÓRIO, E. G. S. L. A. A variação tu e você em Maceió, Alagoas. **Todas as letras**, São Paulo, v. 20, n. 2, p. 85-99, 2018.

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

WENKER, G. **Atlas da Língua da Alemanha do Norte e Central**. 1881. Disponível em: <https://regionalsprache.de/nmd.aspx>. Acesso em: 12 abr. 2018.

ZILLI, G. N. **Por que “tu” e não “você”?** 2009. 50f. Monografia/Trabalho de Conclusão de Curso (Especialização em Língua e Literatura) – Universidade do Extremo Sul Catarinense, Criciúma, 2009.

APÊNDICE – CHAVE DE CODIFICAÇÃO

G1 Variáveis dependentes

v – você

t – tu

G2 Referenciação

g – genérico

e – específico

G3 Parte do inquérito

I – QSL (L MINÚSCULO)

m – QMS

d – discurso semidirigido

q – questões de prosódia

p – pragmática

n – metalinguística

G4 Sexo

H – homem

M – mulher

G5 Faixa etária

1 – 18-30 anos

2 – 50-65 anos

G6 Localidade

SERGIPE:

J – Aracaju

P – Propriá

E – Estâncio

ALAGOAS:

6 – Maceió

U – União dos Palmares

S – Santana do Ipanema

A – Arapiraca

RIO GRANDE DO NORTE:

N – Natal

W – Mossoró

Y – Angicos

b – Pau dos Ferros

1 – Caicó

PARAÍBA:

T – João Pessoa

4 – Cuité

Z – Cajazeiras

7 – Itaporanga

8 – Patos
Q – Campina Grande

PERNAMBUCO:

R – Recife
X – Exu
G – Salgueiro
L – Limoeiro
O – Olinda
F – Afrânio
C – Cabrobó
V – Arco Verde
K – Caruaru
B – Floresta
5 – Garanhuns
D – Petrolina

G7 Escolaridade

f – fundamental
u – superior

G8 Função sintática

a – Sujeito
h – Não-sujeito

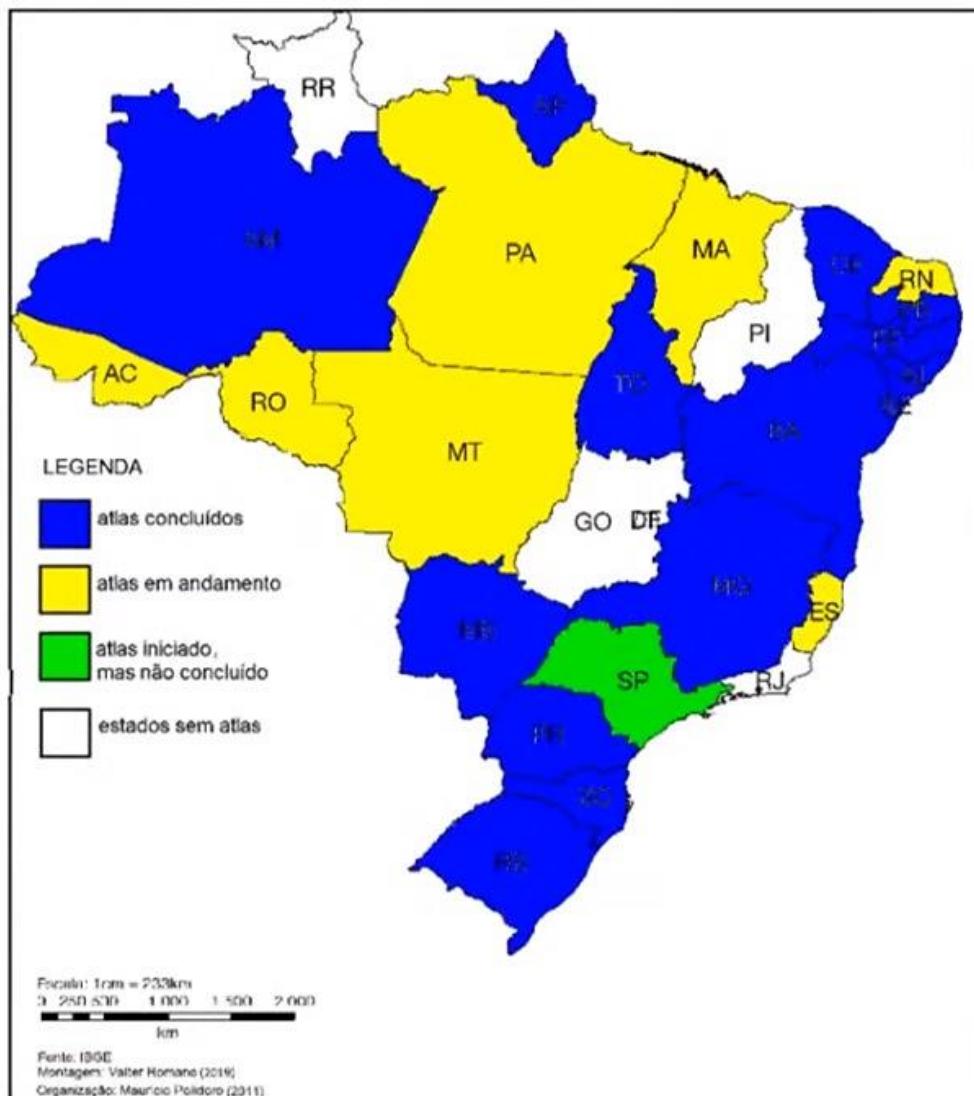
G9 Tempo Verbal

i – Passado
j – Não-passado

G10 Paralelismo

x – tu precedido de tu | você precedido de você
z – tu precedido de você | você precedido de tu
+ – 1^a ocorrência da série
* – Ocorrência isolada

ANEXO A – PANORAMA DA COMPOSIÇÃO DOS ATLAS LINGUÍSTICOS NO BRASIL



Fonte: ROMANO, Valter. Apresentado em 2018. No prelo.

ANEXO B – DOCUMENTOS PROJETO ALiB

DOCUMENTOS PROJETO ATLAS LINGÜÍSTICO DO BRASIL



Fonte: www.alib.ufba.br.

ANEXO C – CDS DESCREVENDO A LÍNGUA, FORMANDO JOVENS PESQUISADORES



Fonte: www.alib.ufba.br.

ANEXO D – AUTORIZAÇÃO DO COMITÊ NACIONAL PARA USO DO *CORPUS* DO PROJETO ALiB



DECLARAÇÃO

Ao utilizar como referencial empírico do trabalho de Pós-Graduação, intitulado ***Tu e você em cinco Estados do Nordeste a partir dos dados do Projeto Atlas Linguístico do Brasil***, que desenvolvo sob a orientação da **Profa. Dra. Josane Moreira Oliveira**, Diretor Científico do Projeto Atlas Linguístico do Brasil/ Membro Credenciado da Equipe Regional (**Projeto ALiB**), dados do *corpus* desse Projeto, **declaro**:

1. Estar ciente de que os materiais do Banco de Dados do **Projeto ALiB** a mim facultados não podem ser repassados, enquanto conjunto de dados, a outro(s) pesquisador(es) e/ou interessado(s) na matéria.
2. Ter pleno conhecimento de que a divulgação parcial ou final do trabalho deve ser sempre acompanhada da indicação da fonte (Banco de Dados do Projeto ALiB) e da citação do nome do orientador.
3. Autorizar que os resultados da análise por mim efetuada sejam utilizados nas publicações do Atlas Linguístico do Brasil, em quaisquer dos volumes que venham a integrar a coleção, mediante a indicação da fonte e a citação do meu nome.
4. Oferecer a minha contrapartida ao Atlas Linguístico do Brasil colaborando, se requerido, na transcrição de dados, catalogação e cópia de materiais e em outras atividades que não impliquem a pesquisa de campo.

E por estar de acordo, firmo a presente DECLARAÇÃO que tem, também, o CIENTE do Orientador.

Salvador, 18 de novembro de 2020.

Ludinalva Santos do Amor Divino

Ludinalva Santos do Amor Divino

RG: 0379159597

CPF: 37543040549

CIENTE

REGISTRADO no

Projeto ALiB

Josane Moreira de Oliveira

Orientadora Profa. Dra. Josane Moreira Oliveira