

UNIVERSIDADE ESTADUAL DE PONTA GROSSA
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS DA LINGUAGEM

BRENDA SOARES REZENDE

VOGAL TEMÁTICA EM P4 DE VERBOS REGULARES DE PRIMEIRA CONJUGAÇÃO:
ESTUDO DA VARIAÇÃO MORFOFONOLÓGICA NO NOROESTE PAULISTA

PONTA GROSSA
2025

BRENDA SOARES REZENDE

VOGAL TEMÁTICA EM P4 DE VERBOS REGULARES DE PRIMEIRA CONJUGAÇÃO:
ESTUDO DA VARIAÇÃO MORFOFONOLÓGICA NO NOROESTE PAULISTA

Dissertação apresentada ao Programa de Pós-graduação em Estudos da Linguagem vinculado à Universidade Estadual de Ponta Grossa, como requisito de avaliação parcial para obtenção do título de Mestre em Estudos da Linguagem.

Orientadora: Prof.^a Dr.^a Márcia Cristina do Carmo

PONTA GROSSA
2025

Rezende, Brenda Soares

R467 Vogal temático em P4 de verbos regulares de primeira conjugação: estudo da variação morfofonológica no noroeste paulista / Brenda Soares Rezende. Ponta Grossa, 2025.

98 f.

Dissertação (Mestrado em Estudos da Linguagem - Área de Concentração: Linguagem, Identidade e Subjetividade), Universidade Estadual de Ponta Grossa.

Orientadora: Profa. Dra. Márcia Cristina do Carmo.

1. Teoria - análise linguística. 2. Variação - mudança linguística. 3. Variação morfofonológica. 4. Primeira pessoa do plural. I. Carmo, Márcia Cristina do. II. Universidade Estadual de Ponta Grossa. Linguagem, Identidade e Subjetividade. III.T.

CDD: 415

BRENDA SOARES REZENDE

VOGAL TEMÁTICA EM P4 DE VERBOS REGULARES DE PRIMEIRA CONJUGAÇÃO:
ESTUDO DA VARIAÇÃO MORFOFONOLÓGICA NO NOROESTE PAULISTA

Dissertação apresentada para obtenção do grau de Mestre em Estudos da Linguagem pela Universidade Estadual de Ponta Grossa. Área de concentração: Linguagem, Identidade e Subjetividade. Linha: Estudos linguísticos.

Ponta Grossa, 28 de março de 2025.



Documento assinado digitalmente

MARCIA CRISTINA DO CARMO

Data: 02/04/2025 17:00:10-0300

Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Márcia Cristina do Carmo, Doutora em Estudos Linguísticos –
Universidade Estadual de Ponta Grossa



Documento assinado digitalmente

LOREMI LOREGIAN PENKAL

Data: 31/03/2025 18:20:53-0300

Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Loremi Loregian-Penkall, Doutora em Linguística –
Universidade Estadual do Centro-Oeste do Paraná



Documento assinado digitalmente

MARCOS BARBOSA CARREIRA

Data: 04/04/2025 14:45:51-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. Marcos Barbosa Carreira, Doutor em Letras –
Universidade Estadual de Ponta Grossa

Dedico à minha mãe e principalmente ao meu pai, por todas as vezes que “cantemos” as canções de Adoniran Barbosa juntos.

AGRADECIMENTOS

Agradeço primeiramente e acima de tudo a Deus, por sempre ser o caminho.

Agradeço aos meus pais, Irenice Aparecida Soares e Ricardo Valério Rezende, que me permitiram continuar estudando, me dando todo o suporte e amor.

Também agradeço à minha irmã, Bruna Soares Rezende, por continuar sendo minha fã número 1 e acreditar piamente que eu sou a maior.

À minha Vó Joana Valério Rezende, por sempre orar por mim e por abrir as portas de sua casa para “você poder descansar”.

Às minhas amigas e amigos pelas palavras de apoio e pela presença constante mesmo na distância.

Agradeço também à minha “profi”, minha orientadora Prof.^a Dr.^a Márcia Cristina do Carmo. Palavras não seriam capazes de descrever toda minha gratidão!

Também gostaria de agradecer à Vilma, secretária do Programa de Pós-Graduação em Estudos da Linguagem da UEPG, por sempre atender minhas solicitações com muita paciência e prontidão.

Agradeço também à banca avaliadora, Prof.^a Dr.^a Loremi Loregian-Penkal e Prof. Dr. Marcos Barbosa Carreira, por terem aceitado contribuir para o desenvolvimento desta pesquisa.

Finalmente, à CAPES, por ter financiado este projeto.

No mais, também preciso mencionar minha gratidão à minha cachorrinha Vina, por ter sido minha companheira durante esses dias solitários de pesquisa acadêmica.

Linguagem e sociedade estão ligadas entre si de modo inquestionável. Mais do que isso, podemos afirmar que essa relação é a base da constituição do ser humano.

Tania Maria Alkmim

RESUMO

Esta pesquisa analisa a variação vocálica antecedente ao sufixo número-pessoal /mo(s)/ em verbos regulares e de primeira conjugação da primeira pessoa do plural no modo indicativo, a exemplo: *cheg[ã]mo(s) ~ cheg[ẽ]mo(s)*. A variedade do Português Brasileiro analisada no trabalho é a falada no interior do estado de São Paulo, mais especificamente na região do município de São José do Rio Preto, situado no noroeste do estado. Para tanto, o *cópus* analisado é formado das 151 entrevistas com amostras de fala espontânea coletadas na região e que foram extraídas do banco de dados IBORUNA (Projeto Amostra Linguística do Interior Paulista – ALIP – Gonçalves, 2024 [2007]). Fundamentado teórico-metodologicamente na Teoria da Variação da Mudança Linguística (Labov, 2008; Weinreich; Labov; Herzog, 2006), este trabalho objetiva determinar quais os fatores condicionadores para a aplicação do uso de [ẽ] no lugar de [ã] na posição antecedente ao sufixo número-pessoal nos verbos de primeira conjugação. Desse modo, são investigadas variáveis linguísticas e extralinguísticas, sendo elas: (i) tempo verbal (presente ou pretérito); (ii) -s final da desinência /mo(s)/ (apagamento ou não); (iii) assertividade da sentença (afirmativa ou negativa); (iv) realização de sujeito (explícito ou nulo); (v) elemento temporal (presença ou não); (vi) sexo/gênero (feminino ou masculino); (vii) faixa etária (de 7 a 15 anos, de 16 a 25 anos, de 26 a 35 anos, de 36 a 55 anos ou superior a 55 anos); e (viii) escolaridade (1º ciclo do ensino fundamental, 2º ciclo do ensino fundamental, ensino médio ou ensino superior). A análise quantitativa foi realizada a partir do programa estatístico GoldVarb-X (Sankoff; Tagliamonte; Smith, 2024 [2005]). Como resultados, obteve-se um total de 642 verbos conjugados na primeira pessoa do plural no modo indicativo, sendo 6 com a variante [ẽ]. Apesar do número relativamente baixo de ocorrências, destacam-se os contextos de pretérito perfeito e com apagamento de -s em /mo(s)/ na fala dos homens para a aplicação da variante [ẽ]. Esta pesquisa também compara o fenômeno na variedade estudada com outras variedades do PB a partir de pesquisas publicadas sobre a temática (Costa, 1990; Pereira, 2014; Pereira; Margotti, 2018; Pereira; Lehmkuhl-Coelho; Loregian-Penkal, 2016; Pereira, 2021; Silva, 2023; Zilles; Maya; Silva, 2000), observando que alguns resultados obtidos nessas pesquisas citadas foram semelhantes às obtidas na atual pesquisa, como a baixa produtividade de verbos conjugados em primeira pessoa do plural nos resultados de Pereira (2014), a influência do fator *tempo verbal pretérito* para a aplicação da regra nos resultados obtidos por Costa (1990), Pereira (2021) e Silva (2023) e a constatação de se tratar de um fenômeno característico de zonas rurais (Pereira; Margotti, 2018).

PALAVRAS-CHAVE: teoria e análise linguística; variação e mudança linguística; variação morfofonológica; primeira pessoa do plural.

ABSTRACT

This research analyzes the vowel variation preceding the personal suffix /mo(s)/ in regular verbs of the first conjugation in the first-person plural of the indicative mood, for example: *cheg[ẽ]mo(s) ~ chegu[ẽ̃]mo(s)* ('we arrive'/'we arrived'). The variety of Brazilian Portuguese examined in this study is spoken in the interior of São Paulo state, specifically in the region of São José do Rio Preto, located in the northwest of the state. The corpus analyzed consisted of 151 interviews with spontaneous speech samples extracted from the IBORUNA database (Amostra Linguística do Interior Paulista Project – ALIP – Gonçalves, 2024). Theoretical and methodological foundations are based on the Theory of Linguistic Variation and Change (Labov, 2008; Weinreich; Labov; Herzog, 2006), with the aim of determining the factors that condition the application of the use of [ẽ̃] in place of [ẽ] in the position preceding the personal suffix in first conjugation verbs. Thus, both linguistic and extralinguistic variables are investigated, including: (i) verb tense (present or past); (ii) final -s of the /mos/ suffix (whether deleted or not); (iii) sentence assertiveness (affirmative or negative); (iv) subject realization (explicit or null); (v) temporal element (presence or absence); (vi) sex/gender (female or male); (vii) age group (7 to 15 years old, 16 to 25 years old, 26 to 35 years old, 36 to 55 years old, or over 55 years old); and (viii) education level (1st cycle of elementary education, 2nd cycle of elementary education, high school, or higher education). Quantitative analysis was conducted using the GoldVarb-X statistical software (Sankoff; Tagliamonte; Smith, 2024 [2005]). Results include 642 verbs conjugated in the first-person plural of the indicative mood, 6 of which had the variant [ẽ̃]. Despite the relatively low number of occurrences, the past perfect tense and the deletion of -s in /mo(s)/ in men's speech are highlighted to the application of the [ẽ̃] variant. This research also compares the phenomenon in the variety under analysis with other varieties of Brazilian Portuguese, drawing on previously published studies on the topic (Costa, 1990; Pereira, 2014; Pereira; Margotti, 2018; Pereira; Lehmkuhl-Coelho; Loregian-Penkal, 2016; Pereira, 2021; Silva, 2023; Zilles; Maya; Silva, 2000). It observes that some findings reported in these studies are similar to those obtained in the current research, such as the low productivity of first-person plural verb forms in the results of Pereira (2014), the influence of the past tense on rule application as demonstrated by Costa (1990), Pereira (2021), and Silva (2023), and the identification of the phenomenon as characteristic of rural areas (Pereira; Margotti, 2018).

KEYWORDS: linguistic theory and analysis; linguistic variation and change; morphophonological variation; first person plural.

LISTA DE ILUSTRAÇÕES

Figura 1 – Carta linguística nº 83 – forma verbal de 1PP no Pretérito Perfeito no Indicativo (Koch; Klassman; Altenhofen, 2011)	35
Gráfico 1 - Porcentagem das variantes /a/ e /e/ conforme TMA em 1ª conjugação.....	41
Gráfico 2 – Número de ocorrências e porcentagens de dados no corpus (Silva, 2023)	49
Quadro 1 – Resumo das pesquisas em outras variedades do PB.....	53
Figura 2 – Região metropolitana de São José do Rio Preto dentro da região administrativa	55
Gráfico 3 – Pirâmide etária da região administrativa de São José do Rio Preto.....	56
Figura 3 – Mapa dos municípios componentes do banco de dados Iboruna	60
Quadro 2 – Variáveis controladas na Amostra Censo	62
Quadro 3 – Perfil social dos informantes.....	75
Quadro 4 – Comparativo dos resultados da variedade do noroeste do estado de SP com os de outras variedades.....	85

LISTA DE TABELAS

Tabela 1 – Densidade demográfica dos municípios do <i>cópus</i> de análise.....	57
Tabela 2 – Ocorrências gerais	74
Tabela 3 – Ocorrências em relação à <i>escolaridade</i>	77
Tabela 4 – Ocorrências em relação à <i>faixa etária</i>	78
Tabela 5 – Ocorrências em relação ao <i>sexo/gênero</i>	79
Tabela 6 – Ocorrências em relação ao <i>tempo verbal</i>	81
Tabela 7 – Ocorrências em relação ao <i>apagamento do -s na DNP /mo(s)/</i>	82
Tabela 8 – Ocorrências em relação à <i>assertividade</i>	83
Tabela 9 – Ocorrências em relação à <i>realização do sujeito</i>	84
Tabela 10 – Ocorrências em relação ao <i>elemento temporal</i>	84

LISTA DE ABREVIATURAS E SIGLAS

1PP	1ª pessoa do plural
AC	Amostra Censo
AI	Amostra Interação
ALERS	Atlas linguístico-etnográfico da Região Sul do Brasil
ALIP	Amostra Linguística do Interior Paulista
AM	Amazonas
BDI	Banco de dados Iboruna
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
DMT	Desinência modo-temporal
DNP	Desinência número-pessoal
EF	Ensino Fundamental
EM	Ensino Médio
ES	Ensino Superior
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
IBGE	Instituto Brasileiro de Geografia e Estatística
IBILCE	Instituto de Biociências, Letras e Ciências Exatas
IDEB	Índice de Desenvolvimento da Educação Básica
IPRS	Índice Paulista de Responsabilidade Social
NC	Não canônica
P4	4ª pessoa
PB	Português Brasileiro
PE	Português Europeu
PIB	Produto Interno Bruto
PN	Pessoa-número
PR	Peso relativo
R	Radical
RMSJRP	Região metropolitana de São José do Rio Preto
SF	Sufixo flexional
SMT	Sufixo modo-temporal
SNP	Sufixo número-pessoal
T	Tema
TMA	Tempo modo aspecto
UEPG	Universidade Estadual de Ponta Grossa
UNESP	Universidade Estadual Paulista
VARLINFE	Variação Linguística de Fala Eslava
VARSUL	Variação Linguística na Região Sul do Brasil
VT	Vogal temática

SUMÁRIO

INTRODUÇÃO	13
1 FUNDAMENTAÇÃO TEÓRICA	16
1.1 TEORIA DA VARIAÇÃO E MUDANÇA LINGUÍSTICA	16
1.2 PROCESSO VARIÁVEL INVESTIGADO	22
1.3 PROCESSO VARIÁVEL INVESTIGADO EM OUTRAS VARIEDADES DO PB	25
1.3.1 Variedade Falada no Noroeste do Rio Grande do Sul (Costa, 1990)	25
1.3.2 Variedade Falada na Região de Florianópolis (Pereira, 2014)	28
1.3.3 Variedade Falada na Região Sudeste do Paraná (Pereira; Lehmkuhl-Coelho; Loregian-Penkal, 2016)	31
1.3.4 Levantamento Diatópico nos Estados de Paraná, Santa Catarina e Rio Grande do Sul (Pereira; Margotti, 2018)	34
1.3.5 Variedade Falada nas Regiões Rurais Eslavo-brasileiras no Sudeste do Paraná (Pereira, 2021; Pereira; Agostinho, 2024)	37
1.3.6 Variedade Falada em Panambi e Porto Alegre no Rio Grande do Sul (Zilles; Maya; Silva, 2000)	44
1.3.7 Variedade Falada no Noroeste de São Paulo (Rezende; Carmo; Fujihara, 2024)	45
1.3.8 Variedade Falada na Região Norte do País (Silva, 2023)	49
2 MATERIAIS E MÉTODOS	55
2.1 COMUNIDADE DE FALA	55
2.2 CÓRPUS DE PESQUISA	60
2.3 VARIÁVEIS INVESTIGADAS	64
2.3.1 Variável Dependente	65
2.3.2 Variáveis Independentes	65
2.3.2.1 Variáveis extralinguísticas	65
2.3.2.1.1 <i>Sexo/gênero</i>	66
2.3.2.1.2 <i>Escolaridade</i>	67
2.3.2.1.3 <i>Faixa etária</i>	67
2.3.2.2 Variáveis linguísticas	68
2.3.2.2.1 <i>Tempo Verbal</i>	68
2.3.2.2.2 <i>Presença do -s final da desinência /mos/</i>	69
2.3.2.2.3 <i>Assertividade da sentença</i>	70
2.3.2.2.4 <i>Realização do sujeito</i>	71
2.3.2.2.5 <i>Elemento temporal</i>	71
3 ANÁLISE DOS RESULTADOS	76

3.1	RESULTADOS DA VARIÁVEL <i>ESCOLARIDADE</i>	78
3.2	RESULTADOS DA VARIÁVEL <i>FAIXA ETÁRIA</i>	80
3.3	RESULTADOS DA VARIÁVEL <i>SEXO/GÊNERO</i>	81
3.4	RESULTADOS DA VARIÁVEL <i>TEMPO VERBAL</i>	82
3.5	RESULTADOS DA VARIÁVEL <i>APAGAMENTO DO -S NA DNP /MO(S)/</i>	84
3.6	RESULTADOS DA VARIÁVEL <i>ASSERTIVIDADE</i>	85
3.7	RESULTADOS DA VARIÁVEL <i>REALIZAÇÃO DO SUJEITO</i>	85
3.8	RESULTADOS DA VARIÁVEL <i>ELEMENTO TEMPORAL</i>	86
3.9	RESUMO COMPARATIVO DOS RESULTADOS OBTIDOS EM VARIEDADES DO PB.....	87
4	CONSIDERAÇÕES FINAIS	88
	REFERÊNCIAS	91
	APÊNDICE A – LISTAGEM DOS VOCÁBULOS LEVANTADOS	96

INTRODUÇÃO

A variação é inerente ao sistema linguístico e faz parte da habilidade comunicativa do indivíduo, e mesmo os falantes que compartilham de uma comunidade de fala e que são fluentes em uma língua em comum produzem formas linguísticas diferentes uns dos outros. Essas diferenças em conjunto dão origem às diversas variedades faladas de todas as línguas, as quais podem ser caracterizadas por diferenças regionais, de classes sociais e de contexto comunicativo, por exemplo. Porém, esses usos diferentes são passíveis de análise visto que possuem um ordenamento. Apesar do aparente “caos linguístico” (Tarallo, 2003), essas diferenciações formam padrões, uma vez que estão interligadas a fatores estruturais, ou seja, internos à língua, e a fatores sociais, externos à língua. Assim, estudar a variação linguística é também estudar os fatores sociais de modo inter-relacionado com os fatores linguísticos, explicações que excluem um ou outro não são eficazes para abarcar toda a complexidade dos mecanismos ocorridos na língua (Weinreich; Labov; Herzog, 2006).

Assim surge a necessidade de este trabalho¹ seguir os moldes da sociolinguística laboviana para dar seguimento à metodologia da pesquisa. A partir da Teoria da Variação e Mudança Linguística, toma-se como objeto de estudo o processo que ocorre nas vogais temáticas em posição antecedente à desinência número-pessoal /mo(s)/ de alguns verbos regulares, especificamente nos verbos de primeira conjugação: /amo(s)/ ~ /emo(s)/. Svobodová (2017) descreve esse fenômeno como a variação da vogal temática A para a vogal média alta E.² Já Pereira (2021, p. 17) o denomina como “alternância vocálica precedente ao sufixo de PN de 1ª pessoa do plural em verbos regulares de 1ª e 2ª conjugações”³ e acrescenta que ocorre principalmente em verbos conjugados no presente e no pretérito perfeito do modo indicativo.

Dessa forma, o estudo apresentado neste trabalho se depara com duas variantes, uma dita canônica conforme as normas prescritivas do Português Brasileiro (PB, daqui em diante) e uma não canônica, conforme o exemplo:

(1) nós **andamo(s)** duas vezes nela... que ela é muito legal [BDI – AC – 037 – NE:

¹ Financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES - 88887.884690/2023-00).

² Notação utilizada por Svobodová (2017).

³ Pereira (2021) considera verbos regulares de 1ª e 2ª conjugação como variáveis dependentes; neste trabalho, atemo-nos à análise dos verbos de 1ª conjugação.

L. 24, grifo nosso]: variante canônica⁴

- (2) ele foi e foi... **andemo(s)** uns:: vinte pé e o menino pegô(u) e falô(u) [BDI-AC-063-NE:L.345, grifo nosso]: variante não canônica

Um registro desse fenômeno pode ser encontrado em músicas populares brasileiras que utilizam o dialeto “caipira” para reforçar uma identidade. Músicas como “Trem das Onze” e “Saudosa Maloca” de Adoniran Barbosa servem de exemplo para ilustrar o uso dessa alternância. No trecho de “Saudosa Maloca”: “dim dim donde nois passemos os dias mais feliz de nossas vidas”, além de conter outras variações típicas do dialeto caipira, o verbo *passar* no pretérito perfeito do modo indicativo “passamos” é substituído por “passemos”. Amaral (2020) afirma que essa variação é característica do dialeto “caipira”, ou seja, a variedade falada do PB nas regiões rurais, mais afastadas dos centros urbanos.

Assim sendo, o objetivo geral desta pesquisa é a análise das motivações linguísticas e extralinguísticas para a ocorrência da variação morfofonológica em primeira pessoa do plural de verbos regulares e de primeira conjugação na fala espontânea dos indivíduos da região noroeste do estado de São Paulo. Quanto aos objetivos específicos, a pesquisa propõe-se a: (i) descrever os fatores linguísticos e extralinguísticos que favorecem ou desfavorecem a aplicação da variação morfofonológica de /amo(s)/ ~ /emo(s)/; e (ii) comparar essa variação na variedade do interior paulista com outras variedades do PB, a partir de outras pesquisas que abordam o mesmo objeto.

Inicialmente, a hipótese a ser testada era a de que esse processo variável fosse estigmatizado, esperando-se ser mais recorrente na fala de pessoas com nível de escolaridade mais baixo e do sexo masculino, pois, conforme Labov (2008), é possível observar uma tendência na fala das mulheres de se manterem mais próximas das variantes prestigiadas da língua para preencherem certos espaços sociais ainda não alcançados. Outra hipótese era a de que o processo fosse um caso de variação estável, sendo comum em todas as faixas etárias investigadas. No mais, sobre as variáveis linguísticas e suas influências para a aplicação ou não da regra variável, inicialmente esperava-se que essas não exercessem fortes tendências na dinâmica da variação, exceto as variáveis *tempo verbal*, *apagamento do -s final da DNP* /mo(s)/ e *elemento*

⁴ A formatação dos exemplos segue as normas presentes em: <https://alip.ibilce.unesp.br/citacoes-e-referencias> (acesso em: 28 abr. 2025) das ocorrências extraídas do banco de dados Iboruna (BDI), especificamente da Amostra Censo (AC), em que constam o número do inquérito, a identificação do tipo de texto (NE, NR, RO, RP e DE) e a identificação da linha (L.).

temporal.

A sociolinguística no Brasil tem tornado prática a utilização de bancos de dados, que “são amostras sistematizadas de língua, coletadas em um universo definido como comunidade de fala” (Freitag, 2016, p. 452), justamente por serem uma fonte sistematizada representativa de uma comunidade de fala. Neste trabalho, utilizou-se a Amostra Censo do banco de dados Iboruna⁵ (151 inquéritos) do Projeto Amostra Linguística do Interior Paulista (ALIP - Gonçalves, 2024, FAPESP - Processo 03/08058-6) da UNESP de São José do Rio Preto. O banco é composto por um conjunto de entrevistas estratificadas socialmente, coletadas entre os anos de 2003 e 2007. Essas amostras são representativas do falar espontâneo da comunidade de fala de São José do Rio Preto e seus municípios circunvizinhos, sendo eles: Guapiaçu, Mirassol, Ipiranga, Onda Verde, Bady Bassit e Cedral (Gonçalves, 2019). Destaca-se que o banco de dados utilizado neste trabalho já estava previamente aprovado por Comitê de Ética no período de sua execução, auxiliando a exequibilidade da presente pesquisa.

Com este estudo, busca-se o preenchimento de lacunas do estudo da variação morfêmica de [ã] ~ [ẽ] na variedade investigada. Esta pesquisa vem de um longo percurso que se iniciou em 2019 com uma Iniciação Científica (IC), momento no qual se analisaram 16 inquéritos do banco de dados Iboruna, focalizando apenas as variáveis extralinguísticas. Essa primeira IC se desdobrou em uma segunda IC,⁶ que analisou então 32 inquéritos do banco de dados, já controlando as variáveis linguísticas (Rezende, 2021). Para a atual dissertação de Mestrado, optou-se pela ampliação total, utilizando todos os inquéritos que compõem a Amostra Censo do banco.

Este trabalho está estruturado da seguinte forma: no capítulo 1, é discutida a fundamentação teórica, contendo a descrição da Teoria da Variação e Mudança Linguística (Labov, 2008); do processo variável investigado nesta pesquisa; e do processo variável investigado em outras variedades do PB. No capítulo 2, apresentam-se os passos metodológicos utilizados, descrevendo a variável dependente e as variáveis independentes linguísticas e extralinguísticas envolvidas. No capítulo 3, apresenta-se a discussão dos resultados obtidos. Por fim, no capítulo 4, têm-se as considerações finais, seguidas pelas referências e pelo apêndice.

⁵ Iboruna, termo com origem no tupi-guarani que significa “Rio Preto” (Gonçalves, 2019).

⁶ Conduzida na Universidade Estadual de Ponta Grossa (UEPG), com bolsa PIBIC concedida pela Fundação Araucária (edital PROPESP/UEPG n. 65/2020).

1 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta o embasamento teórico do trabalho, sendo dividido em: detalhamento da Teoria da Variação de Mudança Linguística (seção 1.1); fundamentação e caracterização da variação morfofonológica investigada (seção 1.2); e descrição da variação estudada em outras variedades do PB (seção 1.3).

1.1 TEORIA DA VARIAÇÃO E MUDANÇA LINGUÍSTICA

A Sociolinguística é um ramo dos estudos linguísticos que se difundiu a partir de meados de 1960 com o objetivo de trazer um novo ponto de vista em contraste com certos aspectos das teorias formalistas anteriores (Hora; Matzenauer, 2021). Por exemplo, com a dicotomia *língua* e *fala*, Saussure (2021) considerava que os diferentes modos de dizer a mesma “coisa” tinham relações com aspectos individuais e exclusivamente referentes à fala. Outra teoria, o Gerativismo, fortemente representado por Noam Chomsky, interpreta que a heterogeneidade da língua é relativa ao nível do *desempenho/performance* do falante, porém de uma forma desvinculada das interações entre língua e sociedade, pois, para Chomsky, o interesse é a intuição totalmente individual que o falante tem de um dado linguístico, como afirmado em “para Chomsky, a linguística é propriamente o estudo da competência [...], o real objeto de estudo linguístico é uma comunidade de fala abstrata, homogênea [...]” (Labov, 2008, p. 218).

Labov (2008], p. 16) criticava essa interpretação “individual” dos estruturalistas acerca dos fenômenos de variação da língua, pois segundo ele “[...] o domínio de um falante nativo de estruturas heterogêneas não tem a ver com multidialetalismo nem com o “mero” desempenho, mas é parte da competência linguística monolíngue”. Em outras palavras, as formas diferentes de um falante se comunicar não necessariamente correspondem ao domínio de múltiplos dialetos, pois um único dialeto já contempla estruturas heterogêneas. Além disso, essa heterogeneidade não se trata de um “erro de desempenho”, mas sim uma característica inerente das línguas correlata ao conhecimento linguístico interno dos falantes.

Os estudos linguísticos anteriores à Sociolinguística se concentravam em investigar a língua como um sistema de si mesma por si mesma e explicando as variações de modo puramente linguístico, sem considerar fatores externos à língua, como as situações comunicativas e o perfil social do falante. A Sociolinguística surge então

argumentando que os diferentes modos de produzir um mesmo significado, ou seja, os processos de variação e suas variáveis linguísticas, também estão contidos na língua e não somente na fala e, para além, acrescenta a preocupação em considerar os fatores sociais para os estudos de variação (Hora; Matzenauer, 2021). Dessa forma, conforme Tagliamonte (2006):

A Sociolinguística argumenta que a linguagem existe em um contexto, dependendo do falante que a está usando, e dependendo de onde e do porquê o falante a está usando. Os falantes demarcam suas identidades e histórias pessoais em seus discursos, bem como seus condicionantes geográficos, econômicos e socioculturais no tempo e no espaço (Tagliamonte, 2006, p. 3, tradução nossa).⁷

Os trabalhos de Weinreich, Labov e Herzog (2006) e Labov (2008) representam os maiores precursores das pesquisas na área e constataam que a dinâmica das variantes apresenta ordenamento, sendo passível de projeções probabilísticas dos comportamentos dos padrões sociais e linguísticos. O ordenamento, os autores denominam de “heterogeneidade ordenada”, o que significa que a existência das variações não nega a estrutura da língua e sua sistematização (Weinreich; Labov; Herzog, 2006). Para identificar com mais rigor as variações, o objeto de investigação da Sociolinguística é o “vernáculo do falante”, definido por Labov (2008, p. 244) como “o estilo em que se presta o mínimo de atenção ao monitoramento da fala” e, a respeito de tê-lo como foco das pesquisas, o autor complementa que “oferece os dados mais sistemáticos para a análise da estrutura linguística” (Labov, 2008, p. 244).

Alguns conceitos são primordiais para a compreensão da teoria Sociolinguística, como por exemplo o de *variação linguística*. A variação é a situação da língua em um dado momento do tempo e corresponde às dinâmicas que “maneiras alternativas de dizer ‘a mesma’ coisa” apresentam entre elas, considerando um mesmo contexto e uma mesma comunidade de fala (Labov, 2008, p. 221). A existência de variações em nada compromete a comunicação entre os falantes, pois elas seguem regras e padrões sistemáticos e os falantes possuem a competência necessária para interpretá-los.

Uma regra variável “é uma regra de reescrita sensível ao contexto que relaciona um par de variantes com $x \rightarrow \langle y \rangle$, de modo que, quando a regra se aplica, ocorre “y” e, quando não se aplica, ocorre “x” (Guy; Zilles, 2007, p. 51, grifo dos autores). As variantes

⁷ “Sociolinguistics argues that language exists in context, dependent on the speaker who is using it, and dependent on where it is being used and why. Speakers mark their personal history and identity in their speech as well as their sociocultural, economic and geographical coordinates in time and space” (Tagliamonte, 2006, p. 3).

são então as formas em concorrência na dinâmica da variação e “são idênticas em valor de verdade ou referencial, mas se opõem em sua significação social e/ou estilística” (Labov, 2008, p. 313). O que define se uma regra variável se aplica ou não é um grupo de fatores que formam uma variável. As variações linguísticas são fundamentais para a língua, pois são as precursoras de uma mudança linguística.

A *mudança linguística* se refere a um processo contínuo e natural a todas as línguas, “uma consequência inevitável da dinâmica interna das línguas naturais” (Weinreich; Labov; Herzog, 2006, p. 139). A mudança linguística também se refere à situação de uma língua no decorrer do tempo, visto que não é abrupta, mas sim gradual, e ocorre a partir da dinâmica das variações:

Velhas e novas formas variantes rivalizam num mesmo momento de tempo e essa alternância pode representar uma transição para um outro estado de língua. Essa transição pode ser percebida principalmente no controle social da variação, na sua distribuição pelos diferentes estratos sociais da população analisada (Weinreich; Labov; Herzog, 2006, p. 139).

Sobre a dinâmica existente entre mudança e variação linguística, Weinreich, Labov e Herzog (2006) atestam que nem toda variabilidade e heterogeneidade acarreta mudança, mas toda mudança implica variabilidade e heterogeneidade.

Ainda segundo Labov (2008), o mecanismo da mudança linguística ocorre a partir da generalização de uma variante em um subgrupo social que atribui um determinado valor a essa variante. Posteriormente, o avanço dessa variante para outros subgrupos terá relação com a atitude social dos indivíduos da comunidade, pois quanto mais positiva a avaliação dos indivíduos em relação a uma variante, maiores serão as pressões sociais para suprimir aquelas que não apresentam uma avaliação tão positiva. Isso não quer dizer que as variantes não percebidas pelo crivo social não sofrem ações externas, pois a capacidade dos falantes de avaliarem uma variante não é necessariamente consciente, inclusive Labov (2008, p. 263) argumenta que “a percepção ingênua do nosso próprio comportamento e do dos outros é normalmente categórica”. A percepção que o falante tem de uma variante só deixa de ser inconsciente para ser explícita “se reforçada dentro de um período de tempo, mas que, do contrário, se extingue sem efeito” (Labov, 2008, p. 263).

Esse mecanismo da mudança linguística indica dois critérios: primeiro que a concepção de comunidade de fala é decisiva para a implementação da mudança, visto que é definida como “participação num conjunto de normas compartilhadas; estas normas

podem ser observadas em tipos de comportamento avaliativo explícito” (Labov, 2008, p. 150). Nota-se então que, para Labov (2008), a questão-chave para compreender “comunidade de fala” é compreender também que ela não se refere necessariamente aos falantes dessa comunidade falarem da mesma forma, mas sim compartilharem de um mesmo tipo de avaliação; e em segundo pelo nível de concorrência existente entre as variantes padrão/não padrão; conservadoras/inovadoras; estigmatizadas/de prestígio, que acabam, de certa forma, “competindo” pela sua existência conforme os valores adquiridos na comunidade de fala.

Segundo Labov (2008, p. 65), “os mecanismos usuais da sociedade produziram diferenças sistemáticas entre certas instituições ou pessoas, e que essas formas diferenciadas foram hierarquizadas em *status* ou prestígio por acordo geral”, e são essas diferenças que refletem na classificação das variáveis. Assim, as variantes mais prestigiadas são aquelas utilizadas pelo grupo de pessoas que apresenta um *status* social igualmente de prestígio. Geralmente, em consequência, as variáveis prestigiadas são também consideradas padrão e conservadoras, pois são utilizadas por grupos de posições mais altas na hierarquia social e que possuem maior poder ao definir a avaliação dada a determinadas variantes.

O estudo da mudança linguística pode ser feito de duas formas: *em tempo real* e *em tempo aparente*. O estudo da mudança *em tempo real* coleta dados em diferentes períodos da história e acompanha a evolução da língua no decorrer de anos, portanto pode ser considerado de caráter diacrônico. A análise em tempo real pode ser *de painel*, no qual se investigam os mesmos indivíduos, mas com um intervalo de tempo de no mínimo uma geração; e *de tendência*, no qual se investigam informantes diferentes em um intervalo de tempo, porém representantes da mesma estratificação social e da mesma comunidade de fala. Já a análise *em tempo aparente* é de caráter sincrônico, pois investiga falantes de diferentes gerações de uma comunidade de fala no mesmo período temporal (Paiva; Duarte, 2015). Esses estudos são importantes para compreender o *status* da variação investigada dentro da língua, ou seja, se a variação já foi totalmente estabelecida no inventário dos falantes (*variação estável*) ou se está no caminho para implementação de uma das variantes (*mudança em progresso*).

Ainda sobre a teoria da mudança linguística, Weinreich, Labov e Herzog (2006, p. 140) sinalizam que “deve responder a algumas questões cruciais que envolvem a instalação de uma nova variante: os fatores condicionantes, a transição, o encaixamento, a implementação e a avaliação”. O problema dos *fatores condicionantes* deve dar conta

de selecionar quais as possíveis mudanças em curso na língua, além de compreender quais são as condições para essa mudança e quais os fatores (linguísticos e extralinguísticos) que podem ou não impulsionar o curso dessas mudanças linguísticas. Weinreich, Labov e Herzog (2006) argumentam que nem todas as possibilidades de variáveis e combinações de fatores são estudadas para investigar todos os casos de variação.

Essa questão está diretamente relacionada ao problema de *encaixamento*, pois não é suficiente determinar quais são os fatores influenciadores de um processo de mudança linguística, é preciso compreender como essa mudança linguística se encaixa na língua correlacionada aos aspectos sociais e linguísticos. Assim, o problema de encaixamento é tratado em duas instâncias: (i) encaixamento na estrutura linguística; e (ii) encaixamento na estrutura social. O primeiro se refere a compreender que a mudança linguística não é abrupta, mas sim um processo gradual que ocorre na língua em função da ação das variáveis linguísticas. Ao se encaixar num sistema, a mudança pode desencadear modificações em outras partes desse sistema. Já o segundo se refere a como os fatores sociais pesam sobre a mudança linguística e como ela se encaixa como um todo na comunidade de fala (Weinreich; Labov; Herzog, 2006).

A mudança linguística pode ser observada ainda enquanto ela está ocorrendo, podendo ser identificados os estágios pelos quais está passando. O problema de *transição* trata da identificação desses estágios passados pelos processos de mudança linguística e como esses processos evoluíram de um estágio para outro, conforme Weinreich, Labov e Herzog (2006, p. 122), “entre quaisquer dois estágios observados de uma mudança em progresso, normalmente se tentaria descobrir o estágio interveniente que define a trilha pela qual a estrutura A evolui para a estrutura B”. A transição pode ser observada na medida que, em primeiro lugar, um falante começa a utilizar uma forma alternativa, depois essa forma alternativa coexiste com outras formas linguísticas e, por fim, a mudança se dá quando uma das formas acaba em desuso (Weinreich; Labov; Herzog, 2006).

O problema de *avaliação* envolve as compreensões subjetivas que os indivíduos possuem para o tratamento e percepção de uma dada forma linguística; sobre esse problema, “o grau de avaliação de variações e de mudanças em progresso pode ser compreendido mais convicentemente nas reações e atitudes dos falantes em relação às variantes do que na sua produção” (Weinreich; Labov; Herzog, 2006, p. 146).

Ainda há o problema de *implementação*, no qual se busca compreender o conjunto

de mudanças e condições possíveis para a mudança; e a como ela é implementada levando em conta tanto fatores sociais e linguísticos, ou seja, como a mudança vai se expandindo para distintos contextos sociais e estruturais. Para Weinreich, Labov e Herzog (2006), esse é o verdadeiro “cerne” da teoria da mudança linguística. Identificando quais são os fatores condicionantes agindo na mudança, é possível tecer explicações sobre como essa mudança vai se expandido para o sistema e se implementando na comunidade de fala (Weinreich; Labov; Herzog, 2006, p. 143).

Tais problemas podem ser bem ilustrados com o estudo realizado em Martha's Vineyard, uma das pesquisas canônicas realizadas por Labov, na ilha localizada na costa nordeste dos Estados Unidos. O estudo objetivou analisar a atitude social dos moradores das zonas “ilha baixa” e “ilha alta” em relação à pronúncia da vogal nos ditongos /aj/ e /aw/ de maneira mais centralizada. Labov (2008) pôde perceber que há uma relação direta com o desejo de permanecer na ilha e/ou de reconhecer a identidade como pertencente da cultura da ilha com o nível de centralização da vogal. As pessoas que saíam da ilha e queriam abandonar suas tradições passavam a produzir mais as variantes não centralizadas. Labov (2008, p. 200), discutindo os resultados dessa pesquisa, apontou que o problema de encaixamento “foi abordado inicialmente pela correlação da centralização de /aw/ com /aj/”, ou seja, a centralização de /aw/ se encaixa na estrutura dos ditongos decrescentes, uma vez que a centralização de /aj/ já existia como um traço regional.

O problema de transição foi solucionado a partir de uma análise em tempo aparente (estudo de várias faixas etárias em um mesmo período) e mostrou que o nível de centralização aumentava com falantes mais jovens. Já o problema de avaliação foi abordado a partir da atitude social e subjetiva dos falantes perante a vida na ilha e as variações em questão. Quanto a isso, Labov (2008, p. 202) apontou que “quanto mais um indivíduo se sentisse capaz de reivindicar e manter *status* como um vineyardense nativo mais ele adotaria a centralização (ay) e (aw)”. Essa e outras pesquisas foram essenciais para comprovar que as variáveis não estão em “variação livre”, isso é, não são aplicadas aleatoriamente, mas sim são condicionadas por fatores sociais e linguísticos (Weinreich; Labov; Herzog, 2006).

Na próxima seção, tem-se a descrição do processo morfofonológico da variação da vogal precedente à desinência /mos/ em verbos regulares de primeira conjugação.

1.2 PROCESSO VARIÁVEL INVESTIGADO

Antes da caracterização do processo investigado, cabe levantar alguns apontamentos sobre a estrutura e o funcionamento dos verbos em português, os quais admitem como sufixo uma diversidade de segmentos fônicos denominados *flexões*. Os sufixos flexionais, ou desinências, podem corresponder a tempo e modo (sufixo modo-temporal) e pessoa e número (sufixo número-pessoal), sobre os quais Câmara Júnior (2015, p.85-86) define que:

Figuram duas noções muito diferentes que se completam para flexionar o vocábulo verbal. Uma, para designar o “tempo”, ou ocasião da ocorrência do que o verbo refere, do ponto de vista do momento da comunicação. A outra, que se lhe segue, indica, dentro do vocábulo verbal, a pessoa gramatical do sujeito. No sufixo flexional de tempo verbal, há acumulação da noção de “modo” (indicativo, subjuntivo, imperativo), e, num tempo do pretérito, a do aspecto inconcluso, ou “imperfeito”, do processo verbal referido. Por sua vez, a flexão de pessoa gramatical implica, automaticamente, na indicação do número, singular ou plural, do sujeito.

Segundo Câmara Júnior (2015, p. 103), o *sufixo modo-temporal* (SMT)⁸ e o *sufixo número-pessoal* (SNP) se aglutinam em um *sufixo flexional* (SF). Em relação ao *radical* (R), o autor afirma que, em geral, nos verbos são uma parte invariável que se ocupam de carregar o significado lexical. Unindo o radical a uma *vogal temática* (VT), tem-se o *tema* do verbo (T), de maneira que a estrutura geral dos verbos pode ser expressa pela seguinte fórmula (Câmara Júnior, 2015, p. 104):

$$1. \quad T (R + VT) + SF (SMT + SNP)$$

Com essa estrutura, o presente trabalho objetiva estudar um processo variável que ocorre na vogal temática, unidade morfológica definida como “morfema classificatório”, ou seja, classifica e explicita a conjugação a que determinado verbo pertence, sendo que a vogal -a indica primeira conjugação (como nos verbos “andar”, “falar” e “amar”), -e é indicativo de segunda conjugação (como nos verbos “comer”, “beber” e “viver”) e -i é indicativo de verbos de terceira conjugação (como nos verbos “sorrir”, “partir” e “abrir”).

Neste trabalho, é analisado então um processo morfofonológico que ocorre na vogal temática dos verbos regulares de primeira conjugação no presente e pretérito perfeito do modo indicativo, caracterizado conforme Svobodová (2017, p. 186) como a

⁸ SMT, SNP, SF, T, R e VT são siglas adotadas por Câmara Júnior (2015), que são reproduzidas neste trabalho com a finalidade de simplificar a descrição.

“troca da vogal temática A por E⁹ na desinência da 1.^a ps. pl. do presente do indicativo ou pretérito perfeito simples do indicativo nos verbos da 1.^a conjugação”, ou seja, vogal temática [ã], que passa a ser realizada como a vogal temática média-alta [ẽ], como por exemplo *fal[ã]mo(s) ~ fal[ẽ]mo(s)* e *pass[ã]mo(s) ~ pass[ẽ]mo(s)*. Pode-se observar, então, que esse processo envolve aspectos morfológicos e fonético-fonológicos (Pereira; Margotti, 2018), o que justifica a denominação de variação “morfofonológica” e não somente variação “morfológica”.

Quando ocorre uma elevação das vogais, costuma-se denominar de um caso de *alçamento vocálico*. Segundo Cristófaró Silva (2011, p. 49), esse é um “fenômeno fonológico que envolve a elevação da propriedade de altura da língua das vogais médias-altas [e] e [o] que se realizarão como as vogais altas [i] e [u]”. A definição, portanto, trata de vogais médias-altas alçando para vogais altas, o que não é o caso do processo variável tratado neste trabalho. A elevação de [ã] para [ẽ] corresponde à elevação da vogal média central para vogal média-alta anterior. Utiliza-se essa notação [ã] ~ [ẽ], pois, segundo Battisti (2014, p. 58, grifo da autora), a nasalização da vogal baixa anterior faz com que ela também passe por elevação e centralização, “passando de vogal baixa ao fone vocálico médio central conhecido por *schwa* na literatura especializada e simbolizado por [ə]”. Assim, conforme a autora, não há ocorrência de vogal baixa nasalizada *[ã] na língua portuguesa.

Sobre a estrutura dos verbos no presente e no pretérito perfeito do modo indicativo, Pereira (2014), baseada no estruturalismo de Câmara Júnior, alega que tanto o presente quanto o pretérito apresentam a raiz, a desinência-modo-temporal (DMT) zero (Ø), a vogal temática /a/ e a desinência-número-pessoal (DNP) /mos/, conforme exemplos:

(1) Nós passamos sempre.

(2) Nós passamos ontem.

Nos exemplos, a sentença (1) se encontra conjugada no presente do modo indicativo, enquanto a sentença (2) está conjugada no pretérito perfeito do mesmo modo.

Diante disso, Câmara Júnior (2015, p. 72) nota que existe uma homonímia entre os tempos verbais presente e pretérito, no qual o morfema /mos/ se torna o mesmo em todos os tempos e verbos do português.¹⁰ Ou seja, no PB, ocorre o uso de uma mesma

⁹ Notação utilizada por Svobodová (2017).

¹⁰ No caso do Português Europeu, a distinção entre passado e presente apresenta uma distinção marcada graficamente pelo acento na vogal temática, como será melhor definido mais à frente.

forma verbal para se referir a distintos contextos temporais. Quanto a isso, pode definir como um caso de *neutralização*, um “fenômeno que expressa a perda de contraste fonêmico em ambiente específico” (Cristófaró Silva, 2011, p. 159). Isso pode ocasionar problemas de ambiguidade caso o falante não utilize outros recursos linguísticos, como o uso de adjuntos adverbiais de tempo, para deixar claro a qual tempo verbal está se referindo.

Assim, a motivação para ocorrência dessa alternância pode ser explicada como uma forma intuitiva encontrada pelo falante para demarcar oralmente a distinção entre esses tempos verbais, como afirmado por Pereira (2021, p. 249) ao dizer que:

É preciso evidenciar [que] a *neutralização* provoca ambiguidade de interpretação e, por isso, é possível que certos falantes de português (tanto brasileiro, quanto europeu) recorram a algumas estratégias para diferenciar os dois tempos verbais, com advérbios e outros elementos.

Outra questão levantada por Pereira (2014, p. 52) é que, nos verbos irregulares de primeira conjugação, essa neutralização não ocorre, pois existe uma distinção entre esses diferentes tempos verbais, como no exemplo do verbo *estar*, que no passado é *estivemos*, e do verbo *dar*, que, conjugado no passado, passa para a forma *demos*, como ilustrado nas sentenças levantadas pela autora:

(3) *Nós DAMO(S) comida para a “Amarelinha” quando ela sente fome.*

(4) *Ontem, nós DEMO(S) comida para a “Amarelinha”.*

Em sua tese de Doutorado (Pereira, 2021), a autora aborda novamente essa questão e sugere que uma das motivações por trás da troca de /a/ por /e/ em verbos regulares se dê então por uma assimilação intuitiva que os falantes fazem dos verbos regulares e irregulares de primeira conjugação, na qual os verbos regulares de 1ª conjugação se “apropriam” dessa solução para “resolver” a questão da ambiguidade estrutural.

Outra hipótese levantada por Pereira (2021) para a origem dessa mudança na vogal temática seria a de que os falantes estejam tomando por base a 1ª pessoa do singular do modo indicativo, como *falei*, e, por analogia, produzindo formas não canônicas em 1ª pessoa do plural, como *falemo*.

Em relação ao Português Europeu (PE), a dinâmica desses verbos é um pouco diferente. A questão da ambiguidade entre os tempos verbais pretérito perfeito e presente apresenta outra dinâmica, pois, na estrutura do PE, existe a distinção tímbrica da vogal /a/. Câmara Júnior (2015), tomando por base a pronúncia da variedade lisboeta, expõe

que, em contexto de verbos no presente, os falantes pronunciam a VT com um traço nasalizado que não ocorre quando os verbos estão conjugados no passado. Essa distinção também pode ser marcada na escrita, com o uso do acento em verbos no passado, como em *falámos*.

Pereira e Margotti (2018) levantam a teoria de que a variação na vogal temática no PB pode ter origens nas formas mais arcaicas do PE que influenciam o português rural, como afirmam em “o uso das formas *fiqúemo* e *fiqúemos* seriam possíveis traços de um português europeu mais arcaico, estando presentes em regiões brasileiras de maior conservadorismo linguístico” (Pereira; Margotti, 2018, p. 229), pois muitas formas comuns à fala rural podem ser encontradas também nas descrições dialetológicas do PE.

Ademais, essa variação é descrita na literatura como característica de regiões afastadas da zona urbana e marca da variedade definida como “caipira”. Esse processo já era descrito por Amaral (2020, p. 64, grifo do autor) como registro do dialeto caipira: “nas formas do preter. perf. do indic. dos verbos em *ar*, a tônica muda-se em *e*: *trabaiêmo* = trabalhamos, *caminhêmo* = caminhamos”.

Na próxima seção, parte-se para a descrição do processo variável em variedades distintas do PB.

1.3 PROCESSO VARIÁVEL INVESTIGADO EM OUTRAS VARIEDADES DO PB

As pesquisas relacionadas à primeira pessoa do plural, em grande parte, adotam como variável dependente a alternância entre *nós* ~ *a gente* (Foeger, 2014; Rúbio; Gonçalves, 2012; e outros), enquanto a variação no uso da concordância do plural /mos/ ~ /mo/ e a variação da vogal temática costumam ser analisadas como variáveis independentes internas. Logo, trabalhos que especificam a alternância [ã] ~ [ẽ] ainda são pouco registrados no PB, sendo alguns deles elencados a seguir.

1.3.1 Variedade Falada no Noroeste do Rio Grande do Sul (Costa, 1990)

Costa (1990), em sua tese de Doutorado intitulada “O verbo na fala camponesa: um estudo de variação”, analisou a produção dos verbos na fala dos camponeses descendentes de italianos residentes na Colônia Santo Antônio, uma região rural do município de Ijuí, no noroeste do Rio Grande do Sul. A autora conta que, em um período que vai da década de 30 até a de 80, houve uma substituição do italiano falado na região

pelo português de uma maneira muito rápida, motivada por ações governamentais de interesses políticos e econômicos. O objeto de estudo de Costa (1990), definido por ela como dialeto Italiano originado do Vêneto, é melhor denominado como Talian. Essa língua surgiu na região a partir da necessidade de comunicação dos imigrantes italianos recém-chegados no Brasil, portanto, é uma língua de base Vêneta com influência do PB, mas não é a mesma língua que se fala na Itália (Loregian-Penkal; Balthazar, 2021). O Talian “pertence ao grupo das línguas minoritárias uma vez que é considerado uma língua de imigração e coexiste com o português em algumas comunidades” (Heinen; Loregian-Penkal, 2024, p. 8).

Como o Talian é considerado uma língua de herança, ou seja, uma língua a qual os falantes possuem o sentimento de pertencimento e identificação cultural, pode soar ofensivo aos falantes o título de *dialeto italiano*, afinal o uso de uma língua envolve questões políticas e negar o Talian reforça a negação de uma identidade cultural (Ortale, 2016; Heinen; Loregian-Penkal, 2024). No ano de 2014, o Ministério da Cultura e o Instituto do Patrimônio Histórico e Artístico Nacional reconheceram o Talian como Referência Cultural Brasileira, sendo a primeira língua de imigração a ser oficialmente reconhecida pelo governo brasileiro (Heinen; Loregian-Penkal, 2024).

A pesquisa de Costa (1990) busca determinar se o português usado em diversos contextos na colônia carrega uma herança do Talian e quais são as diferenças encontradas entre esses usos em diferentes contextos. Segundo Costa (1990), descendentes de imigrantes que moram na zona rural são vítimas de discriminação pela população da zona urbana, que tornam alvo de piadas a língua falada por essas comunidades.

A questão da linguagem é um ingrediente fundamental dessas piadas, o que revela que a população urbana não só tem consciência da diferença entre sua fala e a dos camponeses de diversas origens étnicas, mas também que tem formas de expressar sua discriminação (Costa, 1990, p. 4).

Utilizando o modelo teórico-metodológico variacionista, Costa (1990), em pesquisa de campo, teve contato com 30 famílias, o que proporcionou o registro de 27 entrevistas coletadas no período de dezembro de 1984 a setembro de 1985. Isso resultou em uma amostra de 74 informantes, correspondendo a um percentual de 9% do total da população (aproximadamente 800 habitantes) na época. Uma mesma entrevista resultava em mais de um informante, pois “as entrevistas foram feitas com a família em conjunto porque, pelas normas sociais da comunidade, quando a família recebe visitantes, todos se sentam em círculo” (Costa, 1990, p. 21). Um recorte inicial organizou

os informantes em três grupos: 1ª geração (acima de 51 anos), 2ª geração (filhos da primeira geração ou pais da terceira geração) e 3ª geração (até 20 anos), pois “o recorte por gerações nos dá condições de comparar, de uma geração a outra, os estágios por que o grupo passou no processo de substituição do italiano pelo português” (Costa, 1990, p. 19).

A metodologia adotou como primeiro instrumento um questionário com perguntas sobre a situação socioeconômica do informante e sobre seus processos de aprendizagem e de domínio do português e do italiano. O segundo instrumento utilizado na metodologia foram as entrevistas com temas pré-definidos, como, por exemplo: história de vida; organização do trabalho na família; educação das crianças na época; perspectivas para o futuro dos filhos; e relatos de experiências pessoais sobre acidentes ou doenças. Quanto aos fatores sociais, o trabalho determinou a análise de (i) sexo/gênero; (ii) idade; (iii) grau de instrução; (iv) extensão da propriedade; (v) mecanização do trabalho; e (vi) conforto da residência.

Outro objetivo da pesquisa de Costa (1990), além do já mencionado de encontrar marcas no português de sua substituição do italiano, era o de apresentar a caracterização de um aspecto da gramática do dialeto da região, através do estudo das variações nas flexões verbais comuns entre os falantes da colônia. A análise focalizou mais de um tipo de variação que ocorre nos verbos, podendo ser de natureza fonológica, morfológica, sintática ou semântica. Uma delas é a variação foco da presente dissertação, com as variantes /amos/ ~ /amo/ ~ /emo/, definidas pela autora como essencialmente de motivação morfológica e “marcas flexionais variáveis para uma mesma combinação de propriedades de modo, tempo, número e pessoa” (Costa, 1990, p. 78).

Na segunda parte da pesquisa, foram mantidos os grupos de fatores sociais, apenas retirando as variáveis “grau de instrução”, por ter indicado correlação com a idade (quanto mais jovens, mais instruídos), “extensão da propriedade” e “conforto da residência”, por terem indicado correspondência com a “mecanização do trabalho”. Os resultados dessa pesquisa mostraram ligação da forma /emo/ com o dialeto vênето, já que é uma marca da região noroeste do estado o contato interdialeto nos descendentes de italianos. Ademais, os resultados das formas flexionais de primeira conjugação na concordância de 1ª pessoa do plural do presente do indicativo demonstraram três ocorrências da forma canônica *falamos*, a variante com apagamento do -s final da desinência /mos/ apresentou 10 dados e a variante /emos/ obteve 54 ocorrências. Sobre essa mesma análise em verbos no pretérito perfeito, a autora observou 4 ocorrências da

forma padrão *falamos*, 19 ocorrências da forma *falamo* e 223 ocorrências da forma *falemo*. Costa (1990, p. 113) também notou que:

Há uma tendência muito forte a anular a diferença entre as conjugações na primeira pessoa do plural do presente e marcar sistematicamente essa pessoa pelo uso da forma com e-mo. Em 54,3% das ocorrências, a concordância se fez pelo uso dessa marca. A marca e-mo¹¹ só não ocorre com verbos da terceira conjugação, que são raros.

Esses resultados indicam que a forma não canônica com [ẽ] aparenta ter forte influência de verbos no pretérito, mas não se restringe a ocorrer apenas nesse cenário, visto que, no dialeto da Colônia de Santo Antônio, ambas as formas variáveis ocorreram em contextos de presente e de pretérito perfeito, como consta em “tomando-se os verbos regulares da 1ª conjugação, vemos que a forma e-mo ocorre em 79,4% das ocorrências no presente do indicativo e em 90,3% dos casos de pretérito perfeito” (Costa, 1990, p. 127). Os resultados da autora também apontaram uma relação com a ausência do -s final da desinência /mos/ na variedade não canônica.

Na próxima subseção há a descrição do fenômeno da troca da vogal temática [ẽ] ~ [ẽ] nos verbos regulares de 1ª conjugação na variedade de Florianópolis, em Santa Catarina, região sul do Brasil.

1.3.2 Variedade Falada na Região de Florianópolis (Pereira, 2014)

Pereira (2014), em seu artigo “Cuidamo(s) e cuidemo(s): a variação morfêmica na P4 em verbos regulares de 1ª conjugação”, investiga a variação morfêmica na primeira pessoa do plural no presente e no pretérito do modo indicativo. Apesar de Costa (1990) definir que esse fenômeno é característico de regiões com contato interdialeto com italiano, a autora optou por analisar bairros menos urbanos do município catarinense de Florianópolis, para verificar se essa variante apenas se vincula a regiões com esse tipo de contato interlinguístico ou se pode ocorrer em regiões sem esse tipo de contato. Para tanto, utilizou-se do método quantitativo de descrição (Labov, 2008; Weinreich; Labov; Herzog, 2006) e analisou entrevistas estratificadas socialmente retiradas do banco de dados do Projeto Variação Linguística na Região Sul do Brasil (VARSUL). Seu objeto foi a variável dependente binária “amos” ~ “emos”, considerando somente a concordância com pronome *nós* e verbos regulares de primeira conjugação.

¹¹ Notação utilizada por Costa (1990).

Diferentemente de Costa (1990), Pereira (2014) não considerou a ausência do -s final da desinência /mos/ como uma variável dependente, mas sim como uma variável linguística independente. A amostra retirada do banco de dados VARSUL utilizou 16 entrevistas coletadas nos anos de 2010 a 2012. A princípio, a autora pretendia considerar todos os bairros contidos no banco de dados, mas “havia um predomínio de ‘a gente’ e da utilização categórica da forma canônica (sem variação) nos tempos verbais a serem analisados” (Pereira, 2014, p. 58) e por isso decidiu se restringir aos bairros mais afastados dos centros urbanos.

Dessa forma, as entrevistas ficaram divididas entre oito dos bairros “Ratones e Santo Antônio da Lisboa” amalgamados e oito do bairro “Costa da Lagoa”. Quanto aos grupos de fatores considerados na execução da pesquisa, os sociais foram: (i) sexo/gênero (masculino e feminino); (ii) localidade (Ratones com Santo Antônio da Lisboa e Costa da Lagoa); (iii) escolaridade (Ensino Superior e Ensino Fundamental); (iv) idade (mais velhos e mais jovens). Já os grupos de fatores linguísticos analisados foram: (v) assertividade da sentença (sentença afirmativa ou negativa); (vi) realização do sujeito (explícito ou nulo); (vii) tempo verbal (pretérito perfeito ou presente); (viii) apagamento do -s final da DNP (apagado ou não apagado); e (ix) elemento temporal (presença ou não).

Do total de informantes, apenas sete produziram dados com *nós*, pois o restante dos informantes produziu dados de *a gente*. No bairro Costa da Lagoa, as mulheres mais jovens, independentemente da escolaridade, não produziram dados de verbos conjugados em primeira pessoa do plural justamente por essa preferência de uso da variante *a gente*, que altera a concordância verbal, assim como ocorre com os homens mais jovens de maior escolaridade desse mesmo bairro. Sobre Ratones e Santo Antônio da Lisboa, os perfis que não produziram dados de primeira pessoa do plural foram as mulheres mais velhas e mais jovens de maior escolaridade e os homens mais velhos, independentemente da escolaridade e os mais jovens de maior escolaridade.

Os resultados da autora contaram com 30 dados, sendo 12 (40%) da forma não canônica e 18 (60%) da forma canônica. As variáveis linguísticas “assertividade da sentença” e “advérbio de tempo” foram descartadas da análise por não terem demonstrado relevância para a produção do processo variável. Sobre a variável “localidade”, a hipótese inicial da autora era a de que o bairro mais afastado geograficamente, Costa da Lagoa, apresentaria mais produção da forma não canônica, mas os resultados refutaram essa hipótese, pois 26% dos dados de Costa da Lagoa foram da variante não canônica contra 66% dos dados da mesma variante para o bairro

de Ratores e Santo Antônio da Lisboa. Já os resultados para a variável “escolaridade”, Pereira (2014), tomando por base Zilles e Batista (2006), tinha como hipótese inicial que os menos escolarizados seriam responsáveis pela maior produção da forma não canônica por associação a uma fala menos urbana. Os resultados da autora demonstraram que a hipótese inicial era bastante plausível e pôde ser confirmada, pois a forma não canônica não foi encontrada na fala dos indivíduos com maior nível de escolaridade, enquanto os menos escolarizados produziram ambas as formas.

Para a variável “idade”, a hipótese inicial da autora era a de os falantes mais velhos estarem mais propensos a utilizar a forma não canônica, por possuírem menor contato com mídias sociais e outras localidades. Os resultados dessa variável foram condizentes com a hipótese inicial, pois 83% dos dados gerais da fala dos mais jovens foram representativos da forma canônica, já na fala dos mais velhos esse número foi de 23%. Ainda sobre essa variável, a autora atestou que “embora os resultados se apresentem de acordo com nossa hipótese, percebemos que o grupo de fatores “escolaridade” tem um peso maior no uso da variante não canônica” (Pereira, 2014, p. 65), pois os falantes mais escolarizados, independentemente da idade, não produziram a variante [ẽ].

Por fim, sobre os resultados da variável social “sexo”, a hipótese da autora, baseada no pressuposto de Labov (2008) de que as falantes do sexo feminino tendem a utilizar mais as variantes de prestígio em comparação aos do sexo masculino, era a de que os homens seriam responsáveis pela maior produção da forma não canônica. Os resultados foram diferentes do esperado, pois as mulheres obtiveram o maior uso da variante não canônica, sendo 53% dos dados gerais, em contraste a 31% dos dados gerais na fala masculina. Apesar dos resultados divergirem da hipótese inicial para essa variável, Pereira (2014, p. 66) afirma que estão de acordo com outro pressuposto de Labov (2008), “uma vez que, se a forma não canônica estiver propiciando uma mudança linguística no PB, as mulheres estariam sendo precursoras nesse processo”. Outro resultado significativo obtido por Pereira (2014) foi o fato de todas as ocorrências da variante canônica serem produzidas por um único informante.

Partindo para os resultados das variáveis linguísticas, Pereira (2014), abordando o apagamento do -s final na DNP, baseia-se nas pesquisas de Zilles, Maya e Silva (2000, 2006) e Costa (1990), e tinha como hipótese inicial a de que a ocorrência da variante [ẽ] estaria relacionada ao apagamento do -s final da DNP. Em 44% dos dados gerais em contexto de apagamento do -s, a forma não canônica prevaleceu e, em 55% dos dados em mesmo contexto, a forma canônica prevaleceu, porém, em contexto de não

apagamento do -s final da DNP, a variante não canônica ocorreu em 25% dos dados em um contraste com 75% dos dados com a variante canônica. Sobre os resultados da variante /emos/ com a presença do -s final da DNP, Pereira (2014) destaca que obtiveram um dado somente. Dessa forma, além de confirmar sua hipótese inicial, a autora atestou que o apagamento do -s final da DNP ocorre em ambas as variantes na variedade de Florianópolis.

Na análise da variável “realização do sujeito”, Pereira (2014) tinha como hipótese inicial a de que a forma não canônica seria influenciada por sentenças em que o sujeito não estivesse nulo. Os resultados obtidos não confirmaram a hipótese inicial, pois, em contexto de sujeito explícito, o uso da forma canônica foi mais recorrente (64%) do que a forma não canônica (35%). Já com sujeito nulo, as ocorrências de ambas as variantes se deram de forma razoavelmente equilibrada, 47% dos dados gerais foram da variante não canônica e 52% foram da variante canônica.¹²

Por fim, em relação à variável “tempo verbal”, Pereira (2014) tinha por hipótese inicial a de que, em contextos de pretérito perfeito do modo indicativo, haveria uma tendência por parte dos falantes de produzir a variante [ẽ]. Os resultados confirmaram a hipótese inicial da autora, pois, apesar de a variante canônica (55%) ocorrer mais do que a não canônica (45%) no pretérito perfeito, a variante canônica (63%) também ocorre mais do que a não canônica (36%) em contextos de presente. Ou seja, assim como na variedade da Colônia de Santo Antônio no noroeste do Rio Grande do Sul (Costa, 1990), na variedade de Florianópolis (Pereira, 2014) também ocorreu variação nos dois tempos verbais. Apesar disso, os resultados mostraram que as formas no pretérito demonstram impulsionar a forma não canonizada. Outros fatores apontados como influentes para o uso da variante não canônica foram o apagamento do -s final da DNP e a escolaridade mais baixa.

Na próxima subseção, tem-se a descrição do fenômeno variável da alternância da vogal temática em verbos regulares na variedade da região sudeste do Paraná.

1.3.3 Variedade Falada na Região Sudeste do Paraná (Pereira; Lehmkuhl-Coelho; Loregian-Penkal, 2016)

Pereira, Lehmkuhl-Coelho e Loregian-Penkal (2016), no artigo “Variação na concordância verbal de nós no presente e pretérito perfeito em verbos regulares de 1ª e

¹² Apresentam-se, nesta dissertação, as porcentagens conforme indicadas no artigo.

2ª conjugação: produtiva no sudeste paranaense?”, analisam a variação morfológica em primeira pessoa do plural em verbos regulares de 1ª e 2ª conjugação no presente e pretérito perfeito do modo indicativo, como /amos/ ~ /emos/ e /emos/ ~ /imos/. A metodologia foi qualitativa, fundamentada na sociolinguística de Labov (2008) e Weinreich, Labov e Herzog (2006), e utilizou dados de fala do banco de dados Variação Linguística de Fala Eslava (VARLINFE) e de escrita do acervo das cidades de Irati e Mallet, na região sudeste do Paraná. O banco de dados VARLINFE contém um acervo de amostras de fala das seguintes cidades do sudeste paranaense: Irati, Mallet, Ivaí, Prudentópolis, Rebouças e Rio Azul. Em cada uma dessas cidades, foram coletadas 24 entrevistas estratificadas socialmente, sendo dois informantes para cada perfil social, ou seja, o banco é composto por 12 perfis sociais diferentes constituídos pelo cruzamento das seguintes variáveis extralinguísticas: (i) 2 sexos (feminino e masculino); (ii) 3 níveis escolaridade (1 a 4 anos; 5 a 8 anos; 9 anos ou mais de escola); e (iii) 2 faixas etárias (25 a 49 anos; e acima de 50 anos).

Pereira, Lehmkuhl-Coelho e Loregian-Penkal (2016), dando continuidade à pesquisa de Pereira (2014), optaram por analisar um banco de dados mais representativo do dialeto interiorano rural, uma vez que os resultados anteriores na variedade de Florianópolis se demonstraram pouco produtivos para o uso da variante não canônica com [ẽ], como *cheg[ẽ]mo(s)*, provavelmente por conta da urbanização da região. Inicialmente, Pereira, Lehmkuhl-Coelho e Loregian-Penkal (2016) teorizavam que esse processo variável da alternância da vogal temática em verbos regulares estaria dependente de regiões com influência de dialetos italianos, mas, conforme as autoras puderam notar com Vasconcelos (1970), esse fenômeno já tinha registros no século XIX na variedade do PE. As perguntas principais que regeram a pesquisa foram: “em que contexto geográfico ocorre essa variação?”; “existem também registros escritos dessa variação na variedade em questão?”; “essa variação é produtiva na variedade do sudeste paranaense?”.

Assim, a metodologia delimitou a amostra aos municípios de Irati e Mallet, utilizando uma entrevista sociolinguística de cada um desses municípios. Além disso, a amostra também foi composta por documentos escritos formais de Irati; um diário escrito de um informante de Mallet e depoimentos orais transcritos¹³ de informantes dessas

¹³ O fator é “oral transcrito”, pois se trata de um depoimento oral de um homem do campo da região, transcrito no livro “Memórias dos povos do campo no Paraná Centro-sul” (2013), organizado por Liliana Porto, Jefferson de Oliveira Salles e Sônia Maria dos Santos Marques.

mesmas localidades. Esses textos, além de contarem com diferentes modalidades (escrito, oral, oral transcrito) e gêneros discursivos (documentos, diário, depoimento e entrevista sociolinguística), também se dividiram quanto a estilo (formal e informal) e período (início do século XX e início do século XXI). Os grupos de fatores linguísticos analisados na pesquisa de Pereira, Lehmkuhl-Coelho e Loregian-Penkall (2016) foram: (i) saliência fônica (apagamento ou não do -s final da DNP); (ii) tempo verbal (presente ou pretérito perfeito); e (iii) conjugação verbal (1ª ou 2ª conjugação).

A hipótese inicial de Pereira, Lehmkuhl-Coelho e Loregian-Penkall (2016, p. 495) era a de que “a variante NC [não canônica] vá se estabelecendo cada vez mais conforme o menor nível de formalidade, no caminho também para a oralidade”. Assim, as entrevistas do VARLINFE teriam o maior número de dados da variante não canônica, pois são amostras representativas de fala espontânea da comunidade e, portanto, menos formais; em contrapartida, os documentos de estilo mais formais não teriam essas “marcas não canônicas da oralidade”. Analisando os documentos formais, escritos e do começo do século XX, retirados da obra “Iratí (1974)”, de José Maria Orreda, as autoras observaram que não utilizavam o pronome “nós” e suas formas verbais de concordância (pretérito e presente do indicativo), mas havia uma prevalência de uso de voz passiva nas atas, imperativo nas leis e verbos em 1ª pessoa do singular em depoimentos. Apesar disso, ainda foram encontrados registros da forma canônica em ambos os tempos verbais em depoimentos de 1968 e 1967. Também nas transcrições de textos jornalísticos, os resultados só apresentaram dados da variante canônica com o -s final da DNP não apagado.

Já sobre os registros na obra “Memórias dos povos do campo no Paraná Centro-Sul (2013)”, foram encontrados registros das duas formas e variação na parte respectiva ao diário (escrito e informal). Segundo Pereira, Lehmkuhl-Coelho e Loregian-Penkall (2016, p. 497), “isso demonstrou que o uso das formas não canônicas parece ocorrer com maior expressividade na fala de um sujeito morador do campo”. A variante não canônica encontrada no diário estava em um contexto de passado, enquanto a variante canônica estava em um contexto de presente. Conforme as autoras, isso poderia indicar que os falantes do sudeste do Paraná utilizam a variação morfofonológica para demarcar a diferença entre esses tempos verbais.

Por fim, acerca dos resultados sobre as entrevistas sociolinguísticas do VARLINFE, além de ser observada a variação em 1ª conjugação, diferentemente das outras amostras, também foi possível observá-la em verbos de 2ª conjugação. A variante

não canônica se deu exclusivamente com o apagamento do -s final da DNP e em contextos de pretérito perfeito.

Os resultados dessa pesquisa mostram mais produtividade do processo do que a pesquisa anterior e, diferentemente dela, considerou que o apagamento do -s final não é um condicionante da forma não canônica, mas sim um fator comum na fala e escrita do vernáculo rural. Os resultados também indicam que a forma não canônica em verbos de segunda conjugação ficou restrita a textos escritos do tipo informal. Também apontam para uma tendência dos falantes do sudeste do Paraná a produzirem as diferentes variantes como ferramentas distintivas de tempo verbal, utilizando a variante não canônica para demarcar passado e indicando que a região possa estar passando por um processo de mudança linguística no que se refere ao morfema para passado. Outra consideração feita no trabalho foi o fato dessa variação parecer ser característica da mesorregião do interior do Paraná, por até então não observarem obras que tratem desse fenômeno em outras regiões do Brasil.

Na próxima subseção, há a descrição do estudo diatópico dos usos variáveis de primeira pessoa do plural na região sul do país.

1.3.4 Levantamento Diatópico nos Estados de Paraná, Santa Catarina e Rio Grande do Sul (Pereira; Margotti, 2018)

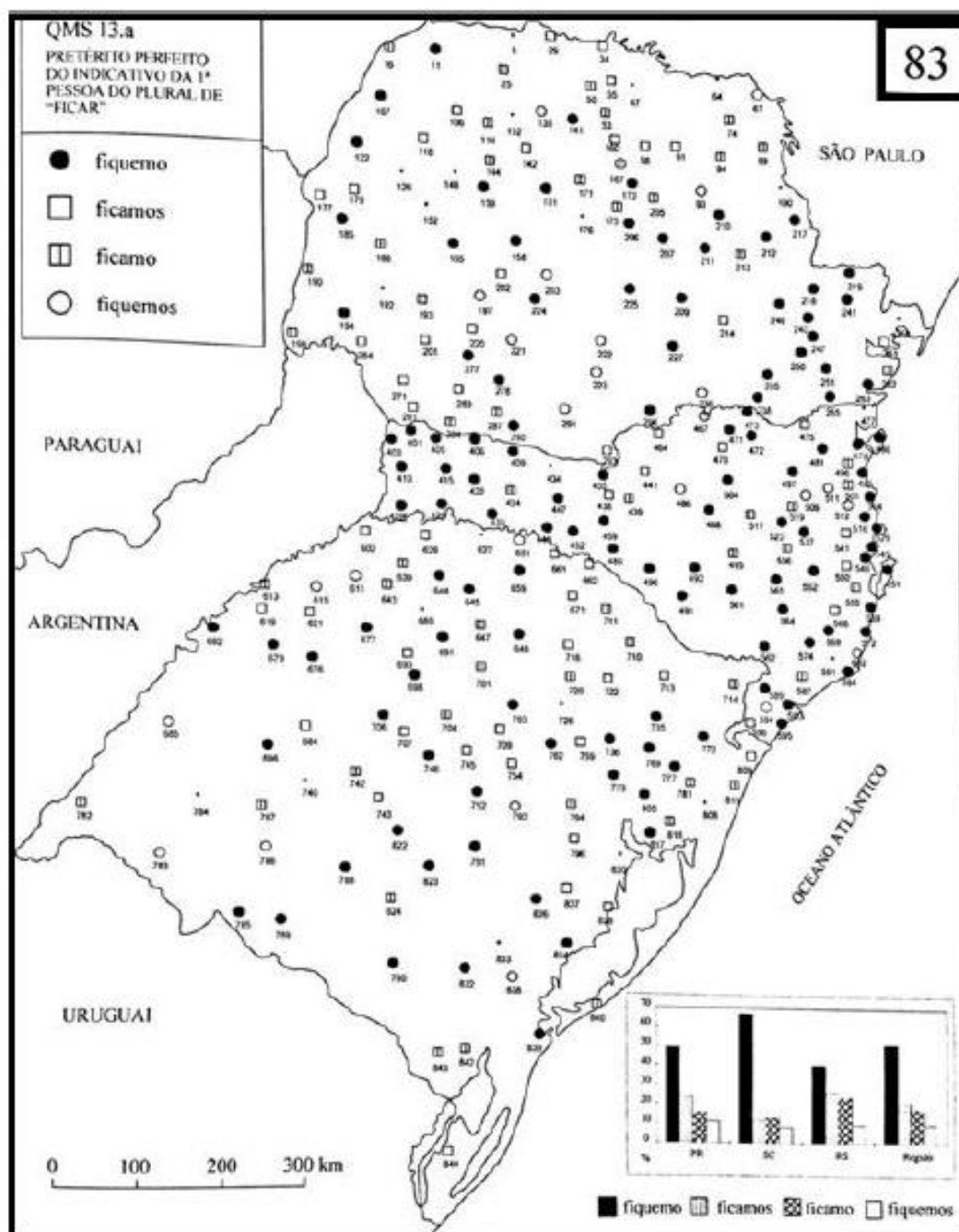
Pereira e Margotti (2018) fazem um levantamento diatópico do uso das variantes de primeira pessoa do plural no pretérito perfeito do modo indicativo. O *cópus* analisado foi a carta linguística nº 83 do Atlas Linguístico-Etnográfico da Região Sul do Brasil (ALERS), que registra o item lexical *ficar*. Portanto, as formas variantes dos autores foram: *ficamos*, *ficamo*, *fiquemos* e *fiquemo*. O ALERS é um atlas produzido a partir de registros de fala advindos de 294 entrevistas retiradas de 294 localidades, sendo 102 do Rio Grande do Sul, 86 de Santa Catarina e 106 do Paraná. Os informantes apresentam pouca escolaridade e idade entre 28 e 58 anos, dividindo-se em 3 informantes para áreas urbanas e 1 para áreas rurais. A carta utilizada do Atlas foi retirada do questionário morfossintático. A variante *ficamos* foi definida como canônica devido à sua proximidade com as regras prescritivas do PB, enquanto as outras três ficaram designadas como não canônicas. Porém, os autores defendem que, entre essas três consideradas não canônicas, é possível determinar uma escala da mais canônica para a menos canônica, conforme a seguinte ordem: *ficamos* > *ficamo* > *fiquemos* > *fiquemo* (Pereira; Margotti,

2018). Sobre o uso de *fiqemos* sem o apagamento do -s final da DNP, Pereira e Margotti (2018, p. 235) argumentam que:

Talvez possa se pensar que o informante de um contexto rural produza a forma *fiqemos* acreditando que ela está de acordo com o padrão, isto é, se a DNP não apagada significa um grau maior de monitoramento, ou seja, quando um falante do campo e pouco escolarizado a produz associada a uma mudança da VT, isso pode significar que ele esteja agindo com hipercorreção em relação ao alçamento vocálico. Não faria muito sentido ele monitorar seu uso em uma desinência (DNP) e não monitorar em outro morfema (VT ou DMT).

O objetivo do trabalho de Pereira e Margotti (2018) foi analisar e descrever essas variantes em jogo em diferentes localidades dos estados do sul do país: Paraná, Santa Catarina e Rio Grande do Sul. A hipótese inicial dos autores era a de que a variante não canônica estivesse relacionada a regiões mais rurais. Como já mencionado, a metodologia utilizou a carta número 83 do atlas linguístico (ilustrada a seguir) e considerou apenas os 275 pontos rurais, sendo eles: 100 pontos no Paraná; 80 pontos em Santa Catarina; e 95 no Rio Grande do Sul.

Figura 1 - Carta linguística nº 83 - forma verbal de 1PP no Pretérito Perfeito no Indicativo



Fonte: Koch, Klassmann e Altenhofen (2011, p. 403)

Conforme é possível observar na Figura 1, os autores identificaram que a variante *fiquemo* está espalhada pelos três estados, e isso pode significar que essa variante não parece ser influenciada pelo aspecto geográfico, mas sim pelo fato de os pontos selecionados serem todos rurais. Pode ser notado que o uso de *fiquemo* aparece em 47% do mapa, seguido de *fícamos* com 19%, *fícamo* com 16% e *fiquemos* com 9%. Para

os autores, isso demonstra como o uso do morfe /a/ está associado ao uso do morfe normativo /mos/ e o uso do morfe não canônico /e/ está associado ao uso do morfe não normativo /mo/ (Pereira; Margotti, 2018).

Os resultados obtidos confirmaram as hipóteses iniciais de Pereira e Margotti (2018) de que as variantes consideradas não canônicas *ficamo*, *fiquemos* e *fiquemo* carregam um traço mais rural e menos urbano do que a variante considerada canônica *ficamos*. As mesorregiões Oeste de Santa Catarina e Centro-sul e sudeste do Paraná foram as que apresentaram o maior índice de uso das formas não canônicas.

1.3.5 Variedade Falada nas Regiões Rurais Eslavo-brasileiras no Sudeste do Paraná (Pereira, 2021; Pereira; Agostinho, 2024)

Pereira (2021), em sua tese “‘O Pirogue, nós Aprendimo da Mãe’ e ‘Agora nós mudemo o Borscht’: Variação Morfofonológica em Comunidades Rurais Eslavo-brasileiras no Sudeste do Paraná”, investiga a alternância vocálica não padrão antecedente a DNP em 1ª pessoa do plural em verbos de 1ª e 2ª conjugações no presente e no pretérito perfeito do indicativo. Pereira (2021) alega que a origem desse trabalho se deu ao notar a existência de uma outra forma para desneutralizar o presente e o pretérito perfeito do indicativo /imo/. Pereira (2021), assim como em Pereira, Lehmkuhl-Coelho e Loregian-Penkal (2016), utilizou como corpus de pesquisa o banco de dados VARLINFE e selecionou como corpus de pesquisa todas as 168 entrevistas do banco de dados representativo da fala de descendentes de eslavos no Paraná.

A pergunta central a ser respondida por Pereira (2021) foi: “os falantes de uma comunidade linguística eslava fazem alternância não canônica da vogal precedente ao sufixo PN na 1ª pessoa do plural em verbos regulares de 1ª e 2ª conjugação?” (Pereira, 2021, p. 22). O objetivo foi analisar e descrever os aspectos sociais, históricos, geográficos e linguísticos envolvidos no processo morfofonológico estudado. Para tanto, a autora faz um estudo de interface entre Morfologia, Fonética e Fonologia, Sociolinguística Variacionista, Contato Linguístico e Dialetoleologia, baseando-se na Teoria da Variação e Mudança Linguística.

O recorte e a categorização dos informantes das entrevistas do banco de dados VARLINFE se deu da seguinte forma: (i) sexo (feminino e masculino); (ii) faixa etária (até 50 anos e mais de 50 anos); e (iii) escolaridade (1 a 4 anos de escolarização, 5 a 8 anos de escolarização e 9 a 11 anos de escolarização); além de (iv) etnia (mista eslava,

polonesa, eslava com outras, ucraniana); ademais, outros condicionadores externos investigados foram: (v) diatopia (Irati, Rio Azul, Cruz Machado, Prudentópolis, Ivaí, Mallet, Rebouças); (vi) nível de bilinguidade (grau I, grau II, grau III); (vii) ocupação; (viii) grau de etnicidade (grau I, grau II, grau III); (ix) grau de mobilidade (baixa, média, alta); e (x) grau de localismo (pouco integrado, mais ou menos integrado e bem integrado). Já as variáveis linguísticas investigadas foram: (xi) produção fonética da vogal ([ẽ], [a], [e], [ɛ] e [i]); (xii) grau de ruralidade (baixo, médio e alto); (xiii) realização do sufixo PN (/mos/ e /mo/); (xiv) tempo-modo-aspecto (presente do indicativo e pretérito perfeito do indicativo); e (xv) item lexical.

Os resultados do trabalho de Pereira (2021) mostraram uma ocorrência geral de 1.858 dados de verbos de 1ª e 2ª conjugação, sendo 1.631 (88%) de 1ª conjugação e 227 (12%) de 2ª conjugação. Com relação aos dados de 1ª conjugação, dos 1.631 totais, 900 foram da variante canônica (55,2%) e 731 da variante não canônica (44,8%). Já os resultados de 2ª conjugação apresentam 227 dados, sendo 110 (48,5%) da variante canônica e 117 (51,5%) da não canônica.

Para executar as análises quantitativas, a autora utilizou o programa estatístico GoldVarb-X, e os fatores linguísticos selecionados como pertinentes para a variação foram, em ordem decrescente: 1) tempo-modo-aspecto (TMA); 2) grau de ruralidade; 3) escolaridade; 4) faixa etária; 5) etnia; 6) diatopia; 7) grau de etnicidade; 8) realização do sufixo de PN; e 9) nível de bilinguidade. Os resultados constituem um uso significativo da forma não canônica, frequente em contexto de uso do sufixo /mo/, em pretérito perfeito do indicativo e por grupos sociais caracterizados por serem falantes de Irati, com menor nível de escolaridade e alto grau de ruralidade.

Em relação aos resultados da variável “tempo-modo-aspecto” obtidos pela pesquisa de Pereira (2021), demonstraram que sentenças em contexto de passado exercem mais influência para o uso da forma não canônica, confirmando a hipótese inicial de Pereira (2021), que previa essa relação. A forma com /e/ ocorreu em 74,4% (peso relativo, daqui em diante PR, = 0.922) das ocorrências gerais de pretérito perfeito e em somente 1,7% (PR = 0.026) das ocorrências de presente. Já o uso da forma com /a/ ocorreu em 98,3% das ocorrências gerais do presente do indicativo e em 25,6% das ocorrências em pretérito perfeito. A interpretação de Pereira (2021) para esses resultados demonstra que o uso de /e/ está atrelado a fazer a marcação do contexto de passado na sentença.

Os resultados da variável “grau de ruralidade” também vão ao encontro das

hipóteses iniciais da pesquisa de Pereira (2021), confirmando-as, pois a hipótese era a de que indivíduos com maior grau de ruralidade falariam mais a variante não canônica. O grau de ruralidade “alto” indicou 479 dados da variante não canônica, correspondendo a 52,9% (PR = 0.646) das ocorrências gerais, enquanto o grau de ruralidade “baixo” obteve somente 27 ocorrências, sendo 17,4% (PR = 0.155).

Dando continuidade aos resultados obtidos por Pereira (2021), ao falar da variável “escolaridade”, a autora, tomando por base Zilles, Maya e Silva (2000) e Pereira (2014), tinha como hipótese inicial a de que a escolaridade seria um fator significativo, visto que, segundo esses autores, pessoas com um menor nível de escolaridade maiores chances teriam de produzir a forma não canônica. Os dados mostraram que falantes do ensino fundamental I produzem mais a forma com /e/, correspondendo a 57,2% dos casos gerais (PR = 0.604).

Outra variável selecionada como relevante pelo programa estatístico para a variedade falada pelos descendentes de escravos do Paraná foi a “faixa etária”. A hipótese inicial de Pereira (2021) era a de que pessoas mais velhas possuiriam maior tendência de reproduzir a variante não canônica. Os resultados obtidos para essa variedade mostraram que a faixa etária acima de 50 anos apresentou a maior porcentagem de aplicação do fenômeno (47,1%) e a faixa etária até 50 anos apresentou uma porcentagem de 43,3% para a mesma variante [ẽ]. Porém, examinando os pesos relativos, o índice do fator até 50 anos foi maior (PR = 0.568) do que o do fator acima de 50 anos (PR = 0.397). Dessa forma, Pereira (2021) conclui que não são os mais velhos que estão proporcionando essa variação e, segundo ela, o que pode explicar essa diferença é a quantidade de dados gerais obtidos em ambas as faixas etárias, pois, enquanto os de idade acima de 50 anos produziram 645 dados, os de até 50 anos produziram 986.

Sobre a variável “etnia”, também selecionada pelo programa, Pereira (2021) observa que os indivíduos de etnia “mista eslava” favoreceram a aplicação da forma não canônica do processo variável investigado nessa região do Paraná, com uma porcentagem de 50,2% e PR = 0.691, seguidos pelos indivíduos poloneses, com 46,1% e PR = 0.534.

Outra variável selecionada foi “diatopia”, que Pereira (2021) defende ter sido selecionada por conta de o corpus ser representativo de cidades de características rurais, e a cidade indicada como a que mais está condicionando os usos da variante com [ẽ], com PR = 0.751, foi Irati, município com alto grau de ruralidade e com uma porcentagem

de 60,3% de ocorrências da forma não canônica, seguido pela cidade de Rio Azul, com 54,1% e $PR = 0.608$. Pereira (2021) afirma que, apesar de Irati ocupar o 4º lugar no *ranking* das cidades mais urbanizadas da região sudeste do Paraná, o que teoricamente deveria não propiciar o uso da forma não canônica, o que a faz exercer tal influência é o inter cruzamento com a variável “grau de ruralidade”, no qual é possível observar que são os indivíduos de maior grau de ruralidade moradores de Irati que estão impulsionando o uso da variante [ẽ].

A próxima variável selecionada pelo programa estatístico para a variedade do PB estudada por Pereira (2021) foi “grau de etnicidade”, sendo de grau III aquele indivíduo que mais se aproxima, identifica-se e expressa costumes ligados a determinada etnia. Os resultados revelaram que, quanto maior o grau de etnicidade, menor o uso da variante não canônica, pois essa variante ocorreu em apenas 37,40% dos casos de indivíduos de grau III e em 56,80% dos casos de indivíduos de grau I. Conforme Pereira (2021), esse resultado não demonstra que essa variante possui ligação com a identidade eslava e indica que não é um fenômeno vinculado à etnia polonesa.

A penúltima variável selecionada pelo programa estatístico foi “realização de PN”, e, sobre ela, Pereira (2021), tomando por base pesquisas anteriores como Bortoni-Ricardo (2011), Zilles, Maya e Silva (2000), Pereira (2014), Pereira e Margotti (2018) e outros, esperava que essa variável estivesse relacionada com o uso da forma não canônica quando a desinência fosse produzida como /mo/. De 1.579 dados de /mo/, 726 aplicaram a regra variável, ou seja, 46% ($PR = 0.514$); por sua vez, houve 52 dados de /mos/, dos quais somente 5 aplicaram a regra, ou 9,6% ($PR = 0.162$). Porém, Pereira (2021) acrescenta que o uso de /mo/ é bastante generalizado, visto que as 1.579 ocorrências registradas advêm de um corpus de 1ª conjugação com 1.631 dados. Esse resultado foi evidenciado quando feito o inter cruzamento dessa variável com a escolaridade, no qual Pereira (2021) obteve o resultado de que, quanto maior o nível de escolaridade, menor o uso da variante /mo/.

Por fim, a última variável selecionada pelo programa foi “grau de bilinguidade”. Os resultados da pesquisa de Pereira (2021) para essa variável inicialmente mostraram que os indivíduos de grau 3 (indivíduos que entendem e falam a língua) de bilinguidade encabeçariam o uso da variante canônica, obtendo uma porcentagem de 60,7% contra 39,3% da variante não canônica. Os indivíduos de grau 2, com uma porcentagem de 53,6% da variante não canônica e 46,4% da variante canônica, então, propiciariam o uso da variante com /e/. Já o de grau 1 estaria no intermédio, com valor equilibrado de 49,3%

do uso da variante /e/ contra 50,7% de uso da variante /a/. Porém, ao gerar os resultados de pesos relativos, Pereira (2021) obteve outra interpretação da análise dos dados. A partir das porcentagens, os indivíduos do grau III seriam os selecionados como influentes para o uso da variante canônica, porém o programa estatístico os selecionou como mais propensos ao uso da variante não canônica, devido a apontar $PR = 0.635$ para esse fator.

Segundo Pereira (2021), o que ocorre com essa variável é um problema de assimetria, no qual um conjunto não equilibrado na distribuição de dados nas categorias compromete a análise estatística. Pereira (2021) identificou uma discrepância significativa nas ocorrências gerais de verbos em primeira pessoa em cada grau de bilinguidade. O grau I obteve 643 ocorrências gerais com 317 aplicações da variante /e/, gerando $PR = 0.350$. O grau II resultou em 179 ocorrências gerais com 96 aplicações e $PR = 0.431$. Por fim, o grau III obteve 809 ocorrências gerais com 318 aplicações da regra variável, gerando assim o maior $PR, = 0.635$.

Pereira (2021) afirma que esse problema pode ocorrer quando se trabalha com variáveis mais “abertas”, ou seja, não tão estratificadas como escolaridade ou faixa etária. Assim, “quando tratamos de bancos de fala que apresentam características mais identitárias, como é o caso do VARLINFE, podemos nos deparar com essa questão da não ortogonalidade das variáveis” (Pereira, 2021, p. 370). Uma solução apontada por Pereira (2021) seria decompor o grau III em mais uma categoria, consequentemente também decompondo o número de ocorrências gerais. Isso faria com que o número de ocorrências gerais dos graus de bilinguidade ficassem mais equilibrados. Porém, como o programa estatístico selecionou a variável “grau de bilinguidade” por último, Pereira (2021) decidiu não executar essa alternativa de solução.

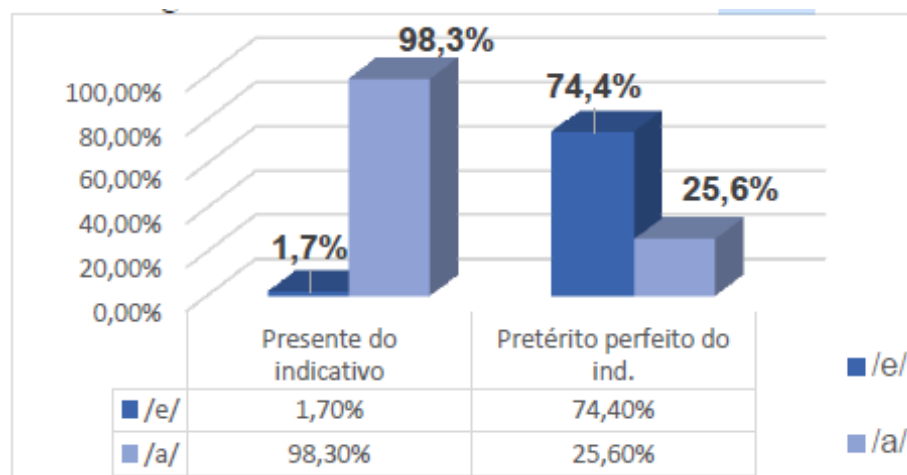
Pereira e Agostinho (2024), no artigo “‘Analisemo’ e ‘descrevimo’: um caso de alçamento vocálico de caráter morfofonológico”, estudam a alternância vocálica de /amos/ ~ /emos/ e /emos/ ~ /imos/ a partir dos dados coletados e resultados obtidos na pesquisa acima descrita (Pereira, 2021). O objetivo geral desse estudo foi defender que essa alternância da vogal temática é um caso de variação morfofonológica, e não um caso de alçamento vocálico prototípico como defendem outros autores. Para tanto, Pereira e Agostinho (2024) buscam descrever o fenômeno de “alçamento vocálico” e o confrontam com as características da variação da vogal temática.

A metodologia de pesquisa adotada foi quantitativa-qualitativa (Labov, 2008, 1982, 1994; Weinreich; Labov; Herzog, 2006; Guy, 1981, 2000) e utilizou todas as entrevistas do banco de dados VARLINFE e, como já explicitado na pesquisa de Pereira (2021), as

rodadas estatísticas foram realizadas pelo programa computacional GoldVarb-X. As variáveis linguísticas e extralinguísticas controladas por Pereira e Agostinho (2024) foram as mesmas da pesquisa de Pereira (2021), porém apenas os resultados da variável linguística TMA foram de interesse para comprovar que o fenômeno não se trata de alargamento vocálico. As autoras explicam que essa variável foi focalizada no artigo por ter sido selecionada pelo programa como significante tanto para o fenômeno em verbos de 1ª conjugação quanto em verbos de 2ª conjugação, além de, segundo Pereira e Agostinho (2024, p. 12), “ser a que nos interessa para a argumentação de que o tipo de alargamento vocálico encontrado em nossos dados é de caráter morfofonológico motivado por tempo, modo e aspecto”.

Retomando os resultados obtidos para a variável TMA referente à 1ª conjugação, Pereira e Agostinho (2024) analisaram os 1.631 dados gerais dos verbos conjugados em P4 de 1ª conjugação. A porcentagem tanto da variante /e/ para pretérito perfeito quanto da variante /a/ para presente do indicativo foram um indicador da preferência da comunidade linguística da região em utilizar a variante /e/ para marcar contextos de passado, como pode ser observado no gráfico 1:

Gráfico 1 – Porcentagem das variantes /a/ e /e/ conforme TMA em 1ª conjugação



Fonte: Pereira (2021, p. 351)

As informações de PR do efeito da variável TMA para a aplicação da variação morfofonológica confirmaram ainda mais essa tendência de uso da variante /e/, visto que o valor do PR para o fator “presente do indicativo” foi = 0.026, em um contraste significativo com o valor de PR = 0.922 para “pretérito perfeito”.

Com esses resultados, Pereira e Agostinho (2024) confrontaram o fenômeno da

variação morfofonológica objeto de estudo com os critérios que definem um fenômeno como alçamento vocálico. Para tanto, as pesquisadoras tomaram por base três pressupostos que caracterizam um alçamento: (i) o alçamento vocálico ocorre com vogais médias-altas que alçam para altas, por exemplo /e/ que alça para [i/I]; (ii) ocorre mais comumente em sílabas pretônicas e postônicas; (iii) tem como resultado alofonia sem distinção de significado (Pereira; Agostinho, 2024).

Sobre o primeiro pressuposto, Pereira e Agostinho (2024) argumentam que o fenômeno da alternância vocálica corresponde a um alçamento de /a/ para /e/, como em *analizamos ~ analisemos*, ou seja, uma vogal baixa alçando para média-alta, subindo dois níveis de elevação, o que seria então um caso de alçamento não prototípico no PB, já que o comum são vogais médias-altas alçando para vogais altas.

Já no que tange ao segundo pressuposto, os dados coletados retirados do banco de dados VARLINFE mostram que esse fenômeno de alternância vocálica ocorre em sílabas tônicas, um contexto incomum para o alçamento vocálico. Além disso, Pereira e Agostinho (2024), com base em outras pesquisas já realizadas na variedade, notaram que, na comunidade linguística investigada, o alçamento vocálico não ocorre em contextos pretônicos e postônicos, nos quais o alçamento costuma acontecer no PB. Logo, se o alçamento vocálico não ocorre nos contextos esperados para essa variedade, há um estranhamento notar que ocorre em contextos em que não é esperado. Para responder essa questão, Pereira e Agostinho (2024, p. 3) afirmam que “o contexto não esperado de alçamento se dá pela necessidade de desambiguar um contexto em que há a mesma forma verbal para situações diferentes de TMA”. Porém, essa afirmação entra em confronto com o último pressuposto.

Como indicado nos resultados da variável TMA, existe uma tendência dos falantes dessa variedade do sudeste do Paraná em produzir a variante não canônica /e/ no pretérito perfeito e quase nunca a utilizando no presente. Esses resultados permitem a Pereira e Agostinho (2024) concluir que a variável TMA está diretamente relacionada a gerar oposição distintiva de significado, ou seja, opor presente do pretérito perfeito de forma que a variante /a/ indique presente e a variante /e/ indique passado. Nos casos prototípicos de alçamento vocálico no PB, as vogais alçam provocando apenas casos de alofonia, sem provocar mudanças de significado, diferentemente do que parece ocorrer com o fenômeno variável em questão.

Justamente essa distinção, o uso da variante /a/ para presente e /e/ para passado, permite dizer que esse fenômeno também apresenta teores fonológicos, dessa forma

“seria possível assumir formas fonológicas diferentes da do PB padrão para {emo} e {imo}, em que a vogal fonológica passa a ser de fato /e/ e /i/ nessa variedade”¹⁴ (Pereira; Agostinho, 2024, p. 20). Em resumo, as autoras definem que essa variação na vogal temática precedente a /mos/ é diferente de um alçamento vocálico prototípico do PB, pelos seguintes critérios: (i) ocorre com uma vogal baixa que alça para média-alta; (ii) ocorre em sílaba tônicas; (iii) resulta em oposição distintiva de TMA; e (iv) ocorre em uma variedade que apresenta taxas baixas de ocorrência de alçamento vocálico prototípico em pretônicas e postônicas (Pereira; Agostinho, 2024).

Dessa forma, com base nesses argumentos, Pereira e Agostinho (2024) defendem que, se esse fenômeno da alternância vocálica for considerado caso de alçamento vocálico, seria necessário rever toda a definição tradicional do conceito na literatura especializada. Assim, as autoras propõem que esse fenômeno na variedade do sudeste do Paraná se trate de um caso de alçamento vocálico não prototípico em sílaba tônica com motivação morfofonológica (Pereira; Agostinho, 2024).

1.3.6 Variedade Falada em Panambi e Porto Alegre no Rio Grande do Sul (Zilles; Maya; Silva, 2000)

Zilles, Maya e Silva (2000) estudam a variação na VT como uma variável independente. A variável dependente desse trabalho é a variação na desinência número-pessoal em primeira pessoa do plural (/mos/ ~ /mo/) de duas comunidades do Rio Grande do Sul, Panambi e Porto Alegre, e o objetivo é discutir em que estágio esse processo de variação se encontra, pois há o questionamento se a desinência número-pessoal está caminhando para não ser mais produzida dentro da variedade estudada.

A amostra utilizou 32 entrevistas sociolinguísticas do banco de dados do Projeto VARSUL. As três variantes consideradas foram a desinência zero, a desinência padrão /mos/ e a desinência não padrão /mo/. As variáveis sociais consideradas foram: (i) escolaridade (primário e segundo grau); (ii) sexo¹⁵ (feminino e masculino); (iii) idade (mais de 50 anos e menos de 50 anos); e (iv) comunidade (Panambi e Porto Alegre). Já as linguísticas foram: (i) conjugação do verbo (primeira, segunda e terceira conjugação); (ii)

¹⁴ Pereira e Agostinho (2024) estendem suas discussões sobre alçamento vocálico e casos de variação morfológica em verbos de 2ª conjugação. Como o objeto desta pesquisa são os verbos de 1ª conjugação, restringimo-nos a esses casos.

¹⁵ Terminologia adotada por Zilles, Maya e Silva (2000).

tempo e modo verbal (indicativo, subjuntivo, imperativo, infinitivo e perífrase ir + infinitivo); (iii) realização do sujeito (pronomes *nós* explícito, nulo e sintagma nominal); (iv) estrutura verbal (simples e composta); (v) tipo de discurso (reportado e não-reportado); (vi) contexto seguinte (vogais, consoantes e pausas); (vii) posição do sujeito em relação ao verbo (posposição, anteposição direta, distância entre sujeito e verbo de uma a três sílabas e distância entre sujeito e verbo de mais de três sílabas); (viii) alternância da vogal temática ([ẽ] e [ẽ̃]); e (ix) posição do acento no verbo (proparoxítonas e paroxítonas).

As hipóteses levantadas pelos autores são a de que a forma verbal proparoxítona favorece a omissão da desinência número-pessoal /mos/; a de que as formas não padrão (zero e /mo/) são favorecidas na fala dos informantes com menor grau de escolaridade; e a de que as formas não padrão são mais favorecidas em Panambi, em decorrência de aquisição tardia do português pelos falantes bilíngues dessa comunidade (Zilles; Maya; Silva, 2000, p. 204). A metodologia seguida nessa pesquisa foi a teoria sociolinguística de Labov e a análise quantitativa foi feita pelo programa VARBRUL.

Os resultados dos autores evidenciaram que o uso da DNP em primeira pessoa do plural ainda estava ocorrendo na região estudada. Sobre a desinência zero, indicam uma relação dessa variante com palavras proparoxítonas. Os autores também encontraram diferenças entre os resultados de Panambi e Porto Alegre, talvez por influência do bilinguismo alemão-português de Panambi. Sobre a escolaridade, foi verificado com os resultados que o uso da DNP padrão /mos/ é favorecido por maiores níveis de escolaridade.

Quanto à variável independente linguística *alternância da vogal temática*, os dados mostraram PR = 0.93 para a relação do apagamento do –s final da DNP com a forma /emos/, que ocorreu em 95% dos casos. Já o uso da forma canônica do PB se mostrou menos importante para o apagamento do -s final da DNP, pois ocorreu em 25% dos casos e com PR = 0.37. Em contrapartida, os resultados dos autores não podem confirmar a influência dos verbos no pretérito para a forma não canônica, pois, em um intercruzamento dessa variável com o *tempo verbal*, os autores obtiveram 22 ocorrências da variante não canônica no presente do indicativo e 22 ocorrências da mesma variante no pretérito perfeito.

1.3.7 Variedade Falada no Noroeste de São Paulo (Rezende; Carmo; Fujihara, 2024)

As investigações sobre a variação morfofonológica em primeira pessoa do plural

de verbos regulares do presente e do pretérito perfeito do modo indicativo na variedade da região noroeste do estado de São Paulo tiveram início em 2019 com um projeto de pesquisa de Iniciação Científica vinculado à Universidade Estadual de Ponta Grossa. Os resultados desse projeto resultaram na publicação do artigo “Alternância Morfofonológica em P4 de verbos regulares de primeira conjugação no noroeste paulista”, por Rezende, Carmo e Fujihara (2024). No artigo, os autores estudam a alternância da vogal temática antecedente à DNP ([ẽ]mos ~ [ē]mos) de verbos regulares de primeira conjugação à luz da Teoria da Variação e Mudança Linguística (Labov, 2008) e de outras pesquisas que investigaram o mesmo fenômeno, mas em outras variedades do PB (Pereira, 2014, 2021; Pereira; Margotti, 2018; Pereira; Lehmkuhl-Coelho; Loregian-Penkal, 2016; Zilles; Maya; Silva, 2000).

O objetivo geral dessa investigação na variedade do interior do estado de São Paulo foi analisar as motivações linguísticas e sociais para a aplicação da regra de variação morfofonológica em P4 na fala espontânea dos indivíduos dessa região. Já os objetivos específicos foram descrever os fatores linguísticos e extralinguísticos em jogo na dinâmica dessa variação, bem como comparar os resultados obtidos para a variedade do noroeste do estado de São Paulo com o de outras variedades do PB.

Essa pesquisa na variedade do interior do estado de São Paulo utilizou como corpus 32 entrevistas do banco de dados Iboruna (Projeto ALIP – Gonçalves, 2024) e o tratamento quantitativo dos dados utilizou o programa estatístico GoldVarb-X (Sankoff; Tagliamonte; Smith, 2024). A motivação para se estudar esse fenômeno nessa variedade se deu pela observação de uma lacuna na descrição desse fenômeno nessa região (Rezende; Carmo; Fujihara, 2024).

As variáveis controladas por Rezende, Carmo e Fujihara (2024) foram: (i) sexo/gênero (masculino e feminino); (ii) faixa etária (16 a 25 anos; 26 a 35 anos; 36 a 55 anos; e superior a 55 anos); (iii) escolaridade (1º ciclo do Ensino Fundamental; 2º ciclo do Ensino Fundamental; Ensino Médio; e Ensino Superior); (iv) tempo verbal (presente e pretérito perfeito); (v) -s final da desinência /mos/ (presente e ausente); (vi) assertividade da sentença (positiva e negativa); (vi) realização do sujeito (nulo e explícito); (vii) elemento temporal (presença e ausência).

A hipótese inicialmente levantada sobre esse processo era a de que se trataria de um fenômeno estigmatizado, sendo mais frequente nos inquiridos dos indivíduos com níveis de escolaridade mais baixo e do sexo/gênero masculino. Outra hipótese era a de que esse processo variável fosse um caso de variação estável, apresentando ocorrências

em todas as faixas etárias, no mais também era esperado que a variável *tempo verbal* fosse influente para a aplicação da regra variável (Rezende; Carmo; Fujihara, 2024).

Com a rodada dos dados levantados no programa estatístico GoldVarb-X, Rezende, Carmo e Fujihara (2024) obtiveram 130 ocorrências gerais (3,8%) de verbos na primeira pessoa do plural do presente e do pretérito perfeito do indicativo. Dessas, 5 ocorrências aplicaram a variação de [ẽ] ~ [ẽ̃], sendo elas: dois casos de *cheg[ẽ̃]mos* e 1 caso de *and[ẽ̃]mos*, *levant[ẽ̃]mos* e *peg[ẽ̃]mos*. Esses resultados indicam uma baixa produtividade de dados, o que, conforme Rezende, Carmo e Fujihara (2024), pode ser explicado pelo uso da forma pronominal *a gente*, e tal escolha “acaba inviabilizando o contexto necessário para a ocorrência da variante estudada (Rezende; Carmo; Fujihara, 2024, p. 12).

Quanto aos resultados das variáveis extralinguísticas, para o *sexo/gênero*, a pesquisa de Rezende, Carmo e Fujihara (2024) observa que todas as ocorrências com aplicação ocorreram por informantes homens, confirmando suas hipóteses iniciais de que as mulheres estariam mais propensas a utilizar a forma mais prestigiada.

Já os resultados da variável *faixa etária* permitiram identificar 4 dados para informantes de 26 a 35 anos e 1 dado para um informante com mais de 55 anos. A hipótese inicial não pôde ser confirmada, pois, segundo Rezende, Carmo e Fujihara (2024), a baixa produtividade dos dados não permite conclusões apuradas sobre o *status* da variação.

Com a análise da variável *escolaridade*, obtiveram-se todos os dados com aplicação na fala de informantes do 1º ciclo de Ensino Fundamental, e essa informação pode ser indicativa de que essa variável exerce uma influência na aplicação da regra variável (Rezende; Carmo; Fujihara, 2024). Os resultados para essa variável se aproximaram da hipótese inicial de que esse fenômeno seria mais recorrente na fala de pessoas menos escolarizadas, demonstrando indícios de estigma social (Rezende; Carmo; Fujihara, 2024).

Os resultados obtidos por Rezende, Carmo e Fujihara (2024) sobre a variável *tempo verbal* constataram que todas as ocorrências da variante [ẽ̃] foram restritas a sentenças no pretérito, assim essa variável aparenta mostrar influência na ocorrência da variação, confirmando as hipóteses preliminares. Esses resultados foram semelhantes aos obtidos por Costa (1990) e Pereira (2014), e essa tendência, conforme explicam Rezende, Carmo e Fujihara (2024), pode ser explicada pela necessidade de desneutralizar os verbos no presente e no pretérito perfeito do indicativo (Rezende;

Carmo; Fujihara, 2024).

A variável *realização do sujeito* obteve 3 ocorrências com aplicação em contexto de sujeito nulo e 2 ocorrências em contexto de sujeito explícito. As ocorrências, tanto as gerais, quanto as com aplicação, obtiveram uma distribuição equilibrada de dados, o que, conforme Rezende, Carmo e Fujihara (2024), indica que essa variável não demonstra ser influente para a variação morfofonológica ocorrer nessa variedade falada do PB (Rezende; Carmo; Fujihara, 2024).

Os apuramentos de Rezende, Carmo e Fujihara (2024) na análise do apagamento do -s final na DNP /mo(s)/ verificaram que todas as 5 ocorrências com aplicação ficaram restritas a contextos em que o -s final não foi pronunciado pelo informante. Semelhante ao que foi obtido na pesquisa de Pereira (2014) e de Zilles, Maya e Silva (2000), esse resultado pode ser um indicativo de que o apagamento do -s final da DNP é determinante para a aplicação da regra variável.

No que diz respeito à variável *assertividade da sentença*, Rezende, Carmo e Fujihara (2024) verificaram uma distribuição assimétrica nas ocorrências gerais, sendo 126 ocorrências gerais em sentenças negativas (com todos os 5 casos de aplicação) e apenas 4 dados em contexto de sentenças negativas (Rezende; Carmo; Fujihara, 2024).

Por fim, os resultados de Rezende, Carmo e Fujihara (2024) sobre *presença ou não de elemento temporal na sentença* revelaram que as 5 ocorrências com aplicação se deram em sentenças com ausência do elemento temporal. Com isso, Rezende, Carmo e Fujihara (2024) discutem que isso pode ser um indicativo de que o fator “ausência” pode exercer influência para o uso da variante [ẽ], pois “a presença de elemento temporal na sentença pode inibir a necessidade da demarcação do pretérito perfeito por meio da variante [ẽ]” (Rezende; Carmo; Fujihara, 2024, p. 15).

Em relação às hipóteses iniciais, ao final da pesquisa, Rezende, Carmo e Fujihara (2024) verificaram que grande parte pode ser confirmada, apenas a variável *faixa etária* não permitiu identificar o *status* da variação morfofonológica em decorrência da baixa produtividade dos dados. O fato de os resultados exibirem maior produção da variante não canônica na fala dos informantes homens de níveis de escolaridade mais baixa parece ser um indicador afirmativo da hipótese de que se trata de um processo estigmatizado. Já sobre as hipóteses iniciais das variáveis linguísticas, pode ser confirmado que a variável *tempo verbal* está relacionada à variação da vogal temática, assim como as variáveis *presença do elemento temporal* e *apagamento do -s final da DNP /mo(s)/*.

Comparando os resultados de Rezende, Carmo e Fujihara (2024) com outras pesquisas conduzidas em outras variedades do PB (Pereira, 2014; Pereira; Margotti, 2018; Pereira; Lehmkuhl-Coelho; Loregian-Penkal, 2016; Zilles; Maya; Silva, 2000), os autores constataram semelhanças do padrão dessa variação com o de outras variedades, como, por exemplo, a baixa ocorrência de dados gerais e comportamentos semelhantes das variáveis linguísticas como condicionadoras do fenômeno variável.

1.3.8 Variedade Falada na Região Norte do País (Silva, 2023)

Na pesquisa intitulada “Pronto, professora, demoremo mas cheguemo: Um estudo sociolinguístico sobre a variação morfofonológica em verbos da 1ª e 2ª conjugações na fala de Tefé-Amazonas”, Silva (2023) analisa a variação morfofonológica em primeira pessoa do plural em verbos regulares de 1ª e 2ª conjugação no presente e no pretérito perfeito (*chegamos ~ cheguemos, nascemos ~ nascimos*), seguindo os preceitos da Teoria da Variação e Mudança Linguística (Labov, 2008; Weinreich; Labov; Herzog, 2006). A partir do levantamento feito por Silva (2023), trata-se de uma região ainda pouco explorada linguisticamente. Há uma quantidade significativa de pesquisas realizadas nos municípios mais próximos do rio Solimões, porém as relativas ao falar dos moradores das regiões interioranas não são muito conhecidas (Silva, 2023).

Diante disso, Silva (2023) investigou a variedade do português falado pela comunidade do município interiorano de Tefé, localizado na mesorregião do centro amazonense. A escolha do objeto de estudo se deu pela observação da recorrência da variação morfofonológica de P4 nessa comunidade. O objetivo geral da pesquisa de Silva (2023, p. 18) é “investigar a ocorrência de variantes da vogal temática da P4, em verbos regulares de 1ª e 2ª conjugação na fala tefeense (AM), especificamente nos tempos Presente e Pretérito Perfeito do Indicativo” e os objetivos específicos são: (i) descrever as variantes morfofonológicas de primeira pessoa do plural em verbos regulares de 1ª e 2ª conjugação na fala tefeense; e (ii) analisar os grupos de fatores linguísticos e extralinguísticos envolvidos na variação da vogal temática.

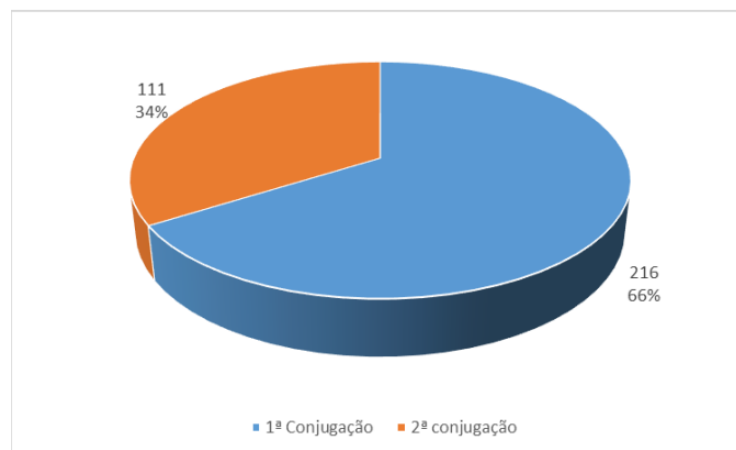
As hipóteses iniciais levantadas por Silva (2023) eram a de que a alternância de [a] para [ẽ] na primeira pessoa do plural, em se tratando de verbos da 1ª conjugação, ocorreria para demarcar a diferença entre o presente e o pretérito perfeito do modo indicativo, assim como também seria a motivação para a alternância de [ẽ] para [i] em verbos de 2ª conjugação. Outra hipótese levantada era a de que as alternâncias de [a]

para [ê] e [ẽ] para [i] seriam mais frequentes na fala de pessoas mais velhas, menos escolarizadas e do sexo masculino.

A pesquisa na variedade de Tefé utilizou uma amostra constituída por 12 entrevistas de fala espontânea coletadas entre 2 e 20 de maio de 2023 em nove bairros da cidade. Os informantes foram estratificados socialmente conforme sexo, escolaridade e faixa etária. Assim, as variáveis extralinguísticas controladas foram: (i) sexo (6 informantes do sexo feminino e 6 informantes do sexo masculino); (ii) escolaridade (Ensino Fundamental e Ensino Médio); e (iii) faixa etária (de 9 a 21 anos, de 22 a 49 anos, e acima de 50 anos). Já as variáveis linguísticas controladas foram: (iv) tempo verbal (presente e pretérito); (v) item lexical; (vi) formas de realização do pronome (*nós* e *a gente*); (vii) apagamento ou não do -s na desinência (apagado e não apagado). A metodologia adotada por Silva (2023) analisa os dados a partir de percentuais e frequência de ocorrência dos dados, mas não submete os dados coletados em programas estatísticos.

Os resultados de Silva (2023) mostraram que a variedade obteve um total de 327 ocorrências de verbos conjugados em primeira pessoa do plural, incluindo as variantes canônicas e não canônicas. Desse total, 216 foram dados de 1ª conjugação e 111 foram dados de 2ª conjugação, conforme é representado no gráfico a seguir:

Gráfico 2 - Número de ocorrências e porcentagens de dados no corpus (Silva, 2023)



Fonte: Silva (2023, p. 84).

Assim como Pereira (2021), Silva (2023) contou com um número mais significativo de verbos de 1ª conjugação do que de 2ª conjugação. Um argumento apontado pela autora, tomando por base Mattos e Silva (2006), para justificar essa diferença é o de que,

desde o latim até atualmente, o paradigma mais produtivo e mais numeroso é o de vogal temática /a/, ou seja, naturalmente se espera que as ocorrências em 1ª conjugação sejam mais numerosas.

Os resultados dos verbos regulares de 1ª conjugação para a variedade de Tefé obtiveram 216 dados, sendo 183 da ocorrência da variante canônica [a] (84,72%) e 33 da ocorrência da variante não canônica [ẽ] (15,28%).

Inicialmente, em conformidade com Labov (2008), era esperado por Silva (2023) que o comportamento linguístico das mulheres fosse mais conservador, com maior frequência de uso da forma canônica [a], porém o obtido para a variedade de Tefé nessa variável não possibilitou confirmar a hipótese inicial, pois os dados destacaram que as formas verbais canônicas foram mais frequentes em relação às formas verbais não canônicas tanto para os homens (94,44%) quanto para as mulheres (79,86%), entretanto, os resultados da forma verbal não canônica [ẽ] foram mais preponderantes no modo de falar das mulheres (20,14%) quando comparados com as ocorrências obtidas no falar dos homens (5,56%). Silva (2023) se baseia em Labov (2008) para justificar essa tendência das mulheres de Tefé de utilizarem a variante [ẽ], pois, segundo Labov (2008), quando uma variável não canônica estiver gerando um processo de mudança linguística, as mulheres geralmente tendem a ser as precursoras.

Silva (2023) tinha como hipótese inicial acerca da variável *escolaridade* que os informantes com níveis mais baixos de instrução utilizariam com mais frequência a variante não canônica /emos/, assim como demonstraram os resultados de Pereira (2014, 2021). As porcentagens para essa variável confirmaram a hipótese inicial levantada. Silva (2023) notou que os informantes da escolaridade Ensino Fundamental estão favorecendo o uso da variante [ẽ], pois indicaram uma porcentagem de uso de 21,84%, enquanto os informantes do nível de escolaridade do Ensino Médio indicaram uma porcentagem de uso dessa mesma variante de 10,85%.

As porcentagens obtidas por Silva (2023) para o uso da variante canônica também corroboraram sua hipótese inicial, pois os indivíduos com nível de formação correspondente ao Ensino Médio obtiveram uma porcentagem de uso de 89,15% em contraposição à porcentagem de uso de 78,16% dos informantes de nível de escolaridade do Ensino Fundamental. Silva (2023) conclui a partir da análise desses resultados que a utilização da variante não canônica por informantes de nível mais baixo de escolarização se justifica pela influência da escola para mudanças de usos linguísticos, pois com a escolarização os falantes “se tornam mais preocupados em

utilizarem as formas padrão da língua” (Silva, 2023, p. 92).

Analisando o grupo de fatores *faixa etária*, Silva (2023) tinha como hipótese inicial que os informantes mais velhos estariam encabeçando o uso da variante não canônica /emo/, baseando-se na teoria de que os jovens possuem mais contato com as mídias, redes sociais e redes de informação. Os resultados de Silva (2023) ressaltaram que, na faixa etária de 9 a 21 anos, 94,44% produziram a variante canônica e 5,56% produziram a variante não canônica. Na faixa etária de 22 a 49 anos, 86,75% dos informantes produziram a forma canônica, contra 13,25% que produziram a forma não canônica. Já na faixa etária acima de 50 anos, 79,38% dos informantes produziram a variante canônica e 20,62% produziram a variante não canônica. Os dados percentuais indicam que, quanto mais velhos, menor o uso da variante /amos/ e maior o uso da variante /emos/.

No que tange à variável *tempo verbal*, Silva (2023), fundamentada em trabalhos anteriores (Costa, 1990; Pereira, 2014; Pereira; Coelho; Loregian-Penkal, 2016; Pereira; Margotti, 2018; Pereira, 2021), assume como hipótese inicial que o uso da variante não canônica é evidenciado em contextos de uso de sentenças no pretérito perfeito do indicativo. Os resultados obtidos por essa autora para a variedade de Tefé apontaram um resultado categórico, ou seja, todas as 33 ocorrências da variante com [ẽ] se deram em contexto de pretérito perfeito. Apesar de os resultados aparentarem confirmar a hipótese inicial da autora, Silva (2023) define que não é possível afirmar que o processo de variação é produzido numa tentativa do falante de solucionar a ambiguidade estrutural entre presente e pretérito perfeito do indicativo, tendo em vista o baixo número de dados coletados.

A hipótese inicial de Silva (2023) sobre a variável *item lexical* era a de que as variantes não canônicas estariam mais relacionadas ao uso de itens lexicais mais frequentes em uso não canônico, como por exemplo os verbos *ficar*, *casar* e *pescar*. Para testar sua hipótese, Silva (2023) realizou o mapeamento dos verbos de primeira pessoa do plural mais recorrentes em seu corpus de análise, e seus resultados localizaram 6 verbos, sendo eles: *chegar* com 5 ocorrências (15,15%), *ficar*, *passar*, *casar* e *pescar* com 3 ocorrências cada (9,09%) e *começar* com 2 ocorrências (3,03%). Segundo Silva (2023, p. 97), “estes casos refletem processos de variação implementados lexicalmente”, e são verbos que representam a realidade e o cotidiano dos informantes.

A penúltima variável investigada por Silva (2023) foi *formas de realização de pronome*. Com esse estudo, o objetivo foi verificar quais formas de realização do pronome estão envolvidas no preenchimento do sujeito quando em contexto de aplicação da

variação da vogal temática. Os fatores controlados por Silva (2023) foram o uso de *nós*, *a gente* e o não preenchimento. A hipótese inicial de Silva (2023) para essa variável era que os informantes preencheriam com mais frequência o sujeito com o pronome *nós* quando houvesse o uso da variante não canônica, pois, com base em seus próprios trabalhos, Silva (2023) verificou que são os informantes mais velhos que utilizam a variante /emo(s)/. Os resultados de Silva (2023) identificaram 33 dados da forma não canônica, sendo 11 (33,33%) quando o sujeito não está preenchido, 17 (51,52%) quando o sujeito é preenchido por *nós*, e 5 (15,15%) quando o sujeito é preenchido pelo pronome *a gente*. Assim, esses resultados confirmaram a hipótese inicial de que o pronome *nós* é o mais recorrente no repertório linguístico.

Por fim, a última variável no estudo de Silva (2023), *apagamento ou não do -s na desinência*, teve o objetivo de verificar a influência do apagamento do -s para o uso ou não da variante [ẽ]. Tomando por base a pesquisa de Zilles, Maya e Silva (2000), Silva (2023) teve como hipótese inicial para essa variável que o apagamento do -s final da desinência estaria influenciando a aplicação da variante [ẽ]. Os resultados demonstraram que 100% dos casos em que foram utilizados a variante /emo(s)/ ocorreram em contextos nos quais houve o apagamento do -s final da desinência.

Com base nessas informações sobre variedades distintas do PB, tem-se um quadro síntese das pesquisas acerca da variação morfofonológica da VT.

Quadro 1 – Resumo das pesquisas em outras variedades do PB

	Costa (1990)	Pereira (2014)	Pereira, Lehmkuhl-Coelho e Loregian-Penkal (2016)	Zilles, Maya e Silva (2000)	Silva (2023)
Variedade	Italianos de uma zona rural de RS	Florianópolis - SC	Sudeste do Paraná	Panambi e Porto Alegre	Tefé - AM
Variável dependente	[ẽ]mo[s] ~ [ẽ]mo	[ẽ]mos ~ [ẽ]mos	[ẽ]mos ~ [ẽ]mos e [ẽ]mos ~ [i]mos	/mos/ ~ /mo/ ~ Ø	[ẽ]mos ~ [ẽ]mos e [ẽ]mos ~ [i]mos
Variáveis selecionadas como relevantes	Tempo verbal pretérito	Escolaridade baixa; Faixa etária; Sexo feminino; Tempo verbal pretérito; e apagamento do –s final.	Tempo verbal pretérito; apagamento do –s final; oralidade; contextos informais.	[ẽ]mos ~ [ẽ]mos	Escolaridade baixa; Sexo feminino; Tempo verbal pretérito; apagamento do –s final.
Ocorrências gerais da forma não canônica	54,3%	40%	-	4,25%	15,28%

Fonte: elaboração própria

Na próxima seção, tem-se a descrição da metodologia empregada na execução da presente pesquisa. Dessa maneira, é apresentado mais detalhadamente o banco de dados Iboruna, assim como o perfil social e econômico da comunidade de fala da região de São José do Rio Preto, foco desta pesquisa.

2 MATERIAIS E MÉTODOS

Esta pesquisa investiga a variação morfofonológica que ocorre na VT precedente à DNP /mo(s)/ em P4 de verbos regulares de primeira conjugação no presente e no pretérito perfeito do modo indicativo na variedade do PB falado na região de São José do Rio Preto. Portanto, é necessário compreender as características sociais e demográficas da região, pois estas estão relacionadas ao comportamento linguístico da variação e, conforme Labov (2008, p. 150), “por meio de observações do comportamento linguístico, é possível fazer estudos detalhados da estrutura de estratificação de classe numa dada comunidade”.

Neste capítulo, é feita toda a descrição dos passos metodológicos utilizados na execução desta pesquisa, tomando por base a metodologia laboviana (Labov, 2008). Assim sendo, o capítulo se divide em: caracterização da comunidade de fala da variedade investigada (seção 2.1); do banco de dados Iboruna (seção 2.2); das variáveis dependentes e independentes investigadas (seção 2.3); e, por fim, dos passos metodológicos empregados na execução desta pesquisa e baseados na Teoria da Variação e Mudança Linguística, de Labov (2008) (seção 2.4).

2.1 COMUNIDADE DE FALA

O banco de dados Iboruna conta com amostras de fala espontânea coletadas na região de São José do Rio Preto e abrange falantes dos municípios de Bady Bassit, Cedral, Guapiaçu, Ipiguá, Mirassol, Onda Verde e, como já mencionado, São José do Rio Preto.

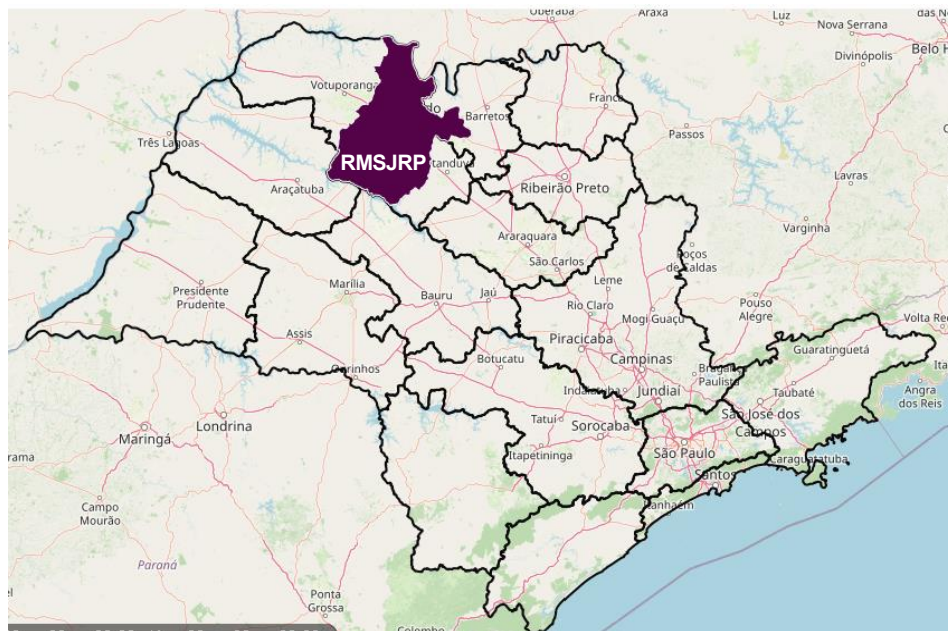
A cidade-polo da região está localizada a uma distância de 442 km da capital paulista. Conforme o censo de 2022 do Instituto Brasileiro de Geografia e Estatística (IBGE), a cidade apresenta um território estimado de 431,944 km² e, conforme os dados de 2019, 124,79 km² do território total correspondem à área urbanizada. Ainda conforme o último censo realizado (2022), a população estimada do município é de 480.393 pessoas, um crescimento de 17,68% em relação ao censo demográfico anterior (2010), colocando a cidade na posição de 45^a mais populosa do país e 10^a do estado.

A história da fundação da cidade tem origem em 1840, quando mineiros passaram a se fixar na região noroeste do estado com intuito de explorar agropecuariamente a área. Mais tarde, no dia 19 de março de 1952, após um fazendeiro doar parte de seu território

para a construção da cidade, é fundado o município de São José do Rio Preto, nomeado assim pela junção do nome do santo patrono (São José) e do nome do principal rio que cruza a cidade (Rio Preto).¹⁶

Com relação às suas divisões geográficas, a cidade principal compõe, juntamente com outras 95, uma região administrativa do estado. Além do mais, em 2021, a Assembleia Legislativa do Estado de São Paulo aprovou a criação da região metropolitana de São José do Rio Preto (RMSJRP), composta por 37 municípios. São José do Rio Preto é o polo principal da região metropolitana por conta de sua influência regional e nacional.

Figura 2 – Região metropolitana de São José do Rio Preto dentro da região administrativa



Fonte: GeoSEADE. Disponível em: <https://portalgeo.seade.gov.br/i3geo/interface/osm.php>. Acesso em: 29 abr. 2025. Destaque da autora.

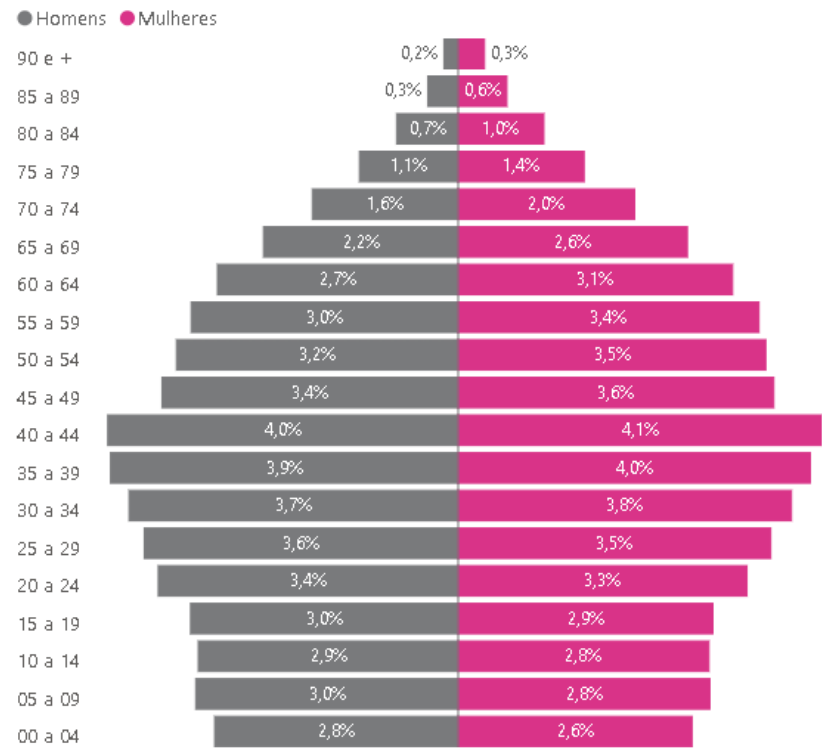
Economicamente, destacam-se na região a produção agropecuária juntamente com a atividade industrial. Na agropecuária, a região tem como base a produção de cana de açúcar, laranja e carne bovina e, na indústria, a economia se baseia na produção de alimentos, de biocombustível, de móveis, de borracha e de produtos do segmento têxtil. O setor terciário também se destaca na região, principalmente no município de São José

¹⁶ Dados retirados do site da prefeitura de São José do Rio Preto. Disponível em: <https://www.riopreto.sp.gov.br/historia-simbolos>. Acesso em: 28 abr. 2025.

do Rio Preto, pois a estrutura de serviços desenvolvida exerce influência e atração geográfica nas localidades vizinhas e é o setor responsável por gerar o maior número de cargos empregatícios na região.¹⁷ Em relação à população, dados disponibilizados pelo IBGE de 2022 demonstram que a média salarial da região é de 2,6 salários-mínimos e que 42,85% da população possui algum tipo de ocupação.

Conforme o último censo demográfico do IBGE divulgado pelo SEADE CENSO 2022, a população da região administrativa aqui citada é de 1.605.475 habitantes, demonstrando um percentual de crescimento de 0,92% no período de 2010-2022. Desse total de habitantes da região, 823.288 são mulheres (51,28%) e 782.187 são homens (48,72%), estabelecendo uma razão de sexo de 105,3% (número de mulheres para cada 100 homens). Ainda pelo último censo divulgado pelo IBGE, a maior parte da população se encontra entre a faixa de 35 a 44 anos, conforme pode ser observado pela pirâmide etária a seguir:

Gráfico 3 – Pirâmide etária da região administrativa de São José do Rio Preto.



Fonte: IBGE, Censos demográficos; Fundação SEADE. Disponível em: <https://censo2022.seade.gov.br/populacao-por-idade-e-sexo/>. Acesso em: 29 abr. 2025.

¹⁷ Dados retirados do portal *Desenvolve SP*. Disponível em: <https://www.desenvolvesp.com.br/mapadaeconomia paulista/ra/sao-jose-do-rio-preto/>. Acesso em: 28 abr. 2025.

Um olhar mais específico para o município central indica uma população de aproximadamente 52,28% de mulheres em contraste com 47,72% de homens, e com a faixa etária predominante na população semelhante à da região (30 e 44 anos).¹⁸ A densidade demográfica bem como a proporção entre homens e mulheres dos outros municípios estão organizadas na seguinte tabela:

Tabela 1 – Densidade demográfica dos municípios do corpus de análise

Município	Pop. Masc.	Pop. Fem.	Total de hab.
Bady Bassit	13.369	13.891	27.260
Cedral	6.222	6.396	12.618
Guapiaçu	10.759	10.952	21.711
Ipiguá	3.400	3.361	6.761
Mirassol	30.702	32.635	63.337
Onda Verde	2.400	2.371	4.771

Fonte: elaboração própria, com base nos dados do IBGE, censo demográfico, e fundação SEADE.

Sobre os fatores socioeconômicos ligados a essa região administrativa, o último Índice Paulista de Responsabilidade Social (IPRS, 2018) fornece alguns dados.¹⁹ O IPRS é um conjunto de indicadores que medem o nível de desenvolvimento dos municípios paulistas a partir de quatro critérios: riqueza, longevidade, escolaridade e classificação dos municípios considerando os três anteriores. O nível de desenvolvimento de cada um desses fatores é avaliado por uma pontuação que varia de uma escala de 0 a 100.

Pelo fator *riqueza*, o IPRS do município central da região atingiu uma pontuação de 38 pontos e, pelos valores de referência, municípios com uma pontuação de até 38 pontos são considerados de baixo índice de riqueza. Porém, em 2021, conforme dados do IBGE, cálculos indicavam que o PIB *per capita* do município era R\$ 44.679,93; em comparação com os outros 644 municípios do estado, São José do Rio Preto ficava alocado na posição 185 e na posição 1.191 em comparação com os 5.570 municípios do país.

Já para o fator *longevidade*, medido pela taxa de mortalidade dentre as diversas

¹⁸ Dados retirados do censo do IBGE de 2022. Disponível em: <https://cidades.ibge.gov.br/brasil/sp/sao-jose-do-rio-preto/panorama>. Acesso em: 28 abr. 2025.

¹⁹ Dados disponíveis em: <https://iprs.seade.gov.br/#> e https://iprs.seade.gov.br/downloads/pdf/metodologia_do_iprs_2018.pdf. Acesso em: 28 abr. 2025.

faixa etárias, o IPRS do município atingiu uma pontuação de 77 pontos, para os valores de referência (superiores a 70), a pontuação classifica a cidade como de alto nível de longevidade.

Por fim, no quesito *escolaridade*, São José do Rio Preto atingiu uma pontuação de 66 e uma qualificação na categoria *alto índice*, por apresentar uma pontuação acima de 61 pontos. Pelo IBGE, em 2010 (dado mais recente), a taxa de escolarização de pessoas entre 6 e 14 anos era de 98% e, pelo Índice de Desenvolvimento da Educação Básica (IDEB) de 2021, a cidade recebeu nota 6,2 para os anos iniciais e nota 5,4 para os anos finais do Ensino Público Fundamental. Em 2014, São José do Rio Preto recebeu do Ministério da Educação o selo de município livre de analfabetismo, gratificação dada às cidades que apresentam índice igual ou superior a 96% de alfabetização, índice que foi superado pela cidade ao atingir a marca de 96,8% de taxa de alfabetização conforme os dados do levantamento do IBGE de 2010.²⁰

No ano de 2024, com base no levantamento do censo de 2022, o IBGE divulgou que São José do Rio Preto se manteve com ótimos índices de população apta a ler e escrever, com uma porcentagem de 97,84%, superando inclusive a média nacional de 93% e estadual de 96,89%. Ainda no censo de 2022, o IBGE divulgou que a taxa média de pessoas não alfabetizadas acima de 15 anos no estado de São Paulo era de 5,14%. São José do Rio Preto mais uma vez apresentou bons índices ao ter a taxa de pessoas analfabetas abaixo dessa média (2,16%). O município de Bady Bassitt também ficou abaixo da média estadual de analfabetismo, com uma porcentagem de 2,44%.²¹

O IPRS também classifica os municípios do estado pela comparação dos 3 parâmetros anteriores, assim podendo ser considerados municípios vulneráveis (baixa riqueza e baixos indicadores sociais), em transição (baixa riqueza e médio indicadores sociais), equitativos (baixa riqueza e alto indicadores sociais), desiguais (alta riqueza e baixos indicadores sociais) ou dinâmicos (alta riqueza e altos indicadores sociais). Quanto a isso, apenas 0,11% (um município) da população da região administrativa de São José do Rio Preto está em situação de vulnerabilidade; 2,48% (sete municípios) da população

²⁰ Informação divulgada por matéria do portal G1 Rio Preto e Araçatuba. Disponível em: <https://g1.globo.com/sao-paulo/sao-jose-do-rio-preto-aracatuba/noticia/2014/06/rio-preto-recebe-selo-de-municipio-livre-de-analfabetismo-dado-pelo-mec.html#:~:text=Do%20G1%20Rio%20Preto%20e%20Ara%C3%A7atuba.%20S%C3%A3o,%C3%ADndice%20igual%20ou%20superior%20a%2096%%20de>. Acesso em: 28 abr. 2025.

²¹ Informação divulgada por matéria do portal G1 São José do Rio Preto e Araçatuba. Disponível em: <https://g1.globo.com/sp/sao-jose-do-rio-preto-aracatuba/noticia/2024/06/08/mais-de-70percent-das-cidades-do-noroeste-de-sp-tem-a-taxa-de-pessoas-nao-alfabetizadas-maior-que-a-media-do-estado-aponta-ibge.ghtml>. Acesso em: 28 abr. 2025.

da região são considerados em situação de desigualdade; 4,5% (14 municípios) estão em transição; 41,12% da população estão equitativas; e 51,79% são dinâmicas. Esses resultados demonstram que a região apresenta bons parâmetros de qualidade de vida para a população.

Na próxima seção, passamos para a descrição do corpus de pesquisa utilizado no desenvolvimento desta pesquisa.

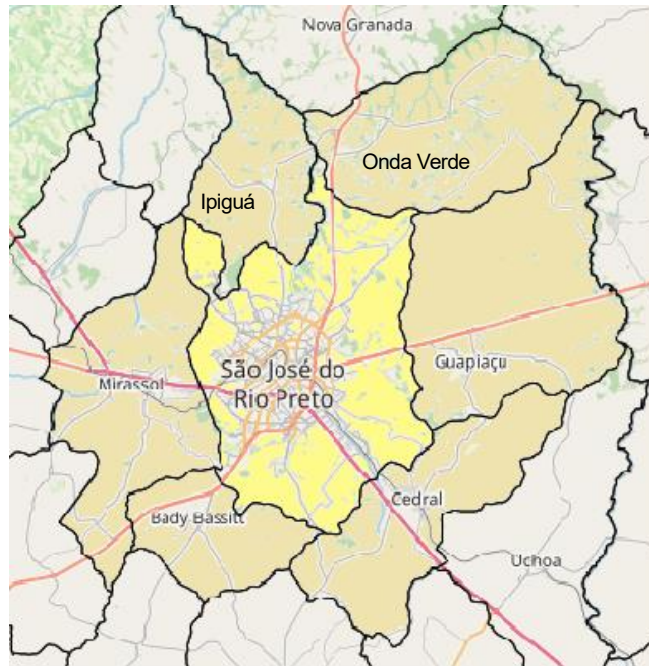
2.2 CÓRPUS DE PESQUISA

Conforme a metodologia laboviana, os processos variáveis também são determinados por fatores sociais (Weinreich; Labov; Herzog, 2006, p. 126). Sendo assim, é necessário que o corpus de pesquisa levantado seja representativo da comunidade de fala. Dessa forma, neste trabalho foi utilizado o banco de dados Iboruna (Rio Preto, em Tupi Guarani), que conta com amostras de fala em contextos reais de uso representativas do PB falado na região de São José do Rio Preto, especificamente abrangendo as cidades que fazem fronteira com a cidade central, a saber: Bady Bassit, localizada 12 km ao sul da cidade central; Cedral, localizada 14 km ao sul; Guapiaçu, localizada 16 km ao leste; Ipiguá, localizada 18 km ao norte; Mirassol, localizada 14 km a oeste; e Onda Verde, localizada 25 km ao norte, conforme pode ser visualizado na figura 3.

O banco de dados Iboruna²² é um dos principais resultados do Projeto Amostra Linguística do Interior Paulista (ALIP) (FAPESP 03/08058-6 – Gonçalves, 2024) e foi elaborado com o objetivo de fornecer um acervo sistematicamente organizado e disponível digitalmente da língua em contexto real de fala, favorecendo as pesquisas linguísticas que tomam por base esse tipo de registro. Os inquéritos que compõem o banco foram coletados entre os anos de 2003 e 2007 e foram construídos a partir dos esforços de um grupo de pesquisadores funcionalistas da UNESP de São José do Rio Preto (Instituto de Biociências, Letras e Ciências Exatas – UNESP/IBILCE). Coordenado pelo Professor Doutor Sebastião Carlos Leite Gonçalves, o banco é considerado de médio porte e conta com aproximadamente 130 horas de gravação de fala que estão disponíveis para acesso juntamente com a transcrição ortográfica dessas gravações e as fichas sociais dos informantes (Gonçalves, 2007).

²² O banco de dados pode ser acessado em: <https://alip.ibilce.unesp.br/bancos-de-dados/banco-de-dados-iboruna>. Acesso em: 28 abr. 2025.

Figura 3 – Mapa dos municípios componentes do banco de dados Iboruna



Fonte: Adaptado do Portal Geo SEADE. Disponível em: <https://portalgeo.seade.gov.br/i3geo/interface/osm.php>. Acesso em: 29 abr. 2025.

Composto por dois tipos de amostra, o banco de dados se divide em: *Amostra Censo* ou *Amostra Comunidade* (AC) e *Amostra de Interação* (AI). O primeiro tipo é o mais utilizado em pesquisas sociolinguísticas, pois controla os fatores sociais relacionados aos informantes, sendo eles: (i) sexo/gênero (feminino e masculino); (ii) faixa etária (de 7 a 15 anos; de 16 a 25 anos; de 26 a 35 anos; de 36 a 55 anos e acima de 55 anos); (iii) escolaridade (1º ciclo do Ensino Fundamental; 2º ciclo do Ensino Fundamental; Ensino Médio; e Ensino Superior); e (iv) renda familiar (até 5 salários mínimos; de 6 a 10 salários mínimos; de 11 a 24 salários mínimos; e acima de 25 salários mínimos). Já o segundo é mais propício para pesquisas relacionadas à análise de discursos em interface com aspectos gramaticais, pois não controla variáveis sociais e objetiva centralizar a interação verbal dentro de uma situação comunicativa (Gonçalves, 2019). Ambos os tipos de amostra, além de serem formados pelos arquivos sonoros, também apresentam arquivos de transcrição ortográfica, fato inédito na época em um projeto de construção de banco de dados sociolinguístico (Gonçalves, 2005).

A AI é composta por 11 amostras de interação dialógica, cada uma com até 5 informantes, contendo um total de 28 informantes. O objetivo dessa amostra é a coleta de conversas rotineiras com assuntos que circundam o cotidiano do indivíduo, como, por exemplo, um diálogo entre duas vizinhas no portão de suas residências. Assim, para

garantir a naturalidade dos diálogos coletados, os informantes obtinham ciência da gravação somente após o final da entrevista, momento no qual poderiam dar o aval ou não do uso do material para a constituição do banco de dados. Pelo mesmo motivo, garantir que os informantes não estivessem atentos aos seus discursos, a metodologia foi flexível quanto às situações comunicativas e temáticas coletadas (Gonçalves, 2019).

Já a AC é composta por 151 entrevistas, com uma média de 50 minutos de duração cada uma. De cada informante, foram coletados durante a gravação das entrevistas 5 tipos de textos: (i) *narrativa de experiência pessoal*, envolvendo relatos pessoais de primeira pessoa; (ii) *narrativa recontada*, solicitando que o informante relatasse um fato ou uma história ocorrida com uma terceira pessoa sem seu envolvimento pessoal; (iii) *texto descritivo*, com a descrição de um local; (iv) *relato de procedimento*, envolvendo o relato de atividades que exijam um processo ordenado, como uma receita, por exemplo; e (v) *relato de opinião*, solicitando que o informante opinasse sobre assuntos variados, como política ou religião (Gonçalves, 2019).

Durante a construção da AC, o inter cruzamento das variáveis sociais resultou em 152 perfis sociais, porém o grupo de pesquisadores não conseguiu identificar um representante do sexo feminino, do 1º ciclo de Ensino Fundamental na faixa etária entre 26 e 35 anos, o que ocasionou a ausência da entrevista número 60 no corpus (Gonçalves, 2005), tornando-o composto por 151 entrevistas. Gonçalves (2007) aponta que uma possível explicação para a dificuldade em encontrar perfis sociais das faixas etárias intermediárias com cruzamento de baixa escolaridade pode se dar “ao fato de a população da região de São José do Rio Preto apresentar um alto índice de escolaridade em relação à média do país” (Gonçalves, 2007, p. 20). Porém, a ausência do AC-060 não compromete a representatividade do banco sobre a comunidade de fala, pois, conforme Labov (2008), não é necessária uma quantidade muito alta de registros de fala para determinar as características linguísticas da comunidade de fala estudada, pois a variação é padronizada (Gonçalves, 2007).

Esta pesquisa, por ser uma continuação de pesquisas anteriores (Rezende, 2021; Rezende; Carmo; Fujihara, 2024) que utilizaram esse banco de dados para o estudo da variação em voga, ampliou o corpus de pesquisa para todas as 151 entrevistas que compõem a AC, conforme ilustradas no quadro seguinte:

Quadro 2 – Variáveis controladas na Amostra Censo

Renda / gênero Faixa etária / escolaridade		25+ salários mínimos		11- 24 salários mínimos		6-10 salários mínimos		5- salários mínimos	
		Masc	Fem	Masc	Fem	Masc	Fem	Masc	Fem
7 a 15 anos	1º. Ciclo EF	001	002	003	004	005	006	007	008
	2º. Ciclo EF	009	010	011	012	013	014	015	016
	Ens. Médio	017	018	019	020	021	022	023	024
16-25 anos	1º. Ciclo EF	025	026	027	028	029	030	031	032
	2º. Ciclo EF	033	034	035	036	037	038	039	040
	Ens. médio	041	042	043	044	045	046	047	048
	Superior	049	050	051	052	053	054	055	056
26-35 anos	1º. Ciclo EF	057	058	059	060	061	062	063	064
	2º. Ciclo EF	065	066	067	068	069	070	071	072
	Ens. médio	073	074	075	076	077	078	079	080
	Superior	081	082	083	084	085	086	087	088
36-55 anos	1º. Ciclo EF	089	090	091	092	093	094	095	096
	2º. Ciclo EF	097	098	099	100	101	102	103	104
	Ens. médio	105	106	107	108	109	110	111	112
	Superior	113	114	115	116	117	118	119	120
55+ anos	1º. Ciclo EF	121	122	123	124	125	126	127	128
	2º. Ciclo EF	129	130	131	132	133	134	135	136
	Ens. médio	137	138	139	140	141	142	143	144
	Superior	145	146	147	148	149	150	151	152
Legenda			BAD		OND		UPI		QUA
			CED		MIR		SJP		

Fonte: Gonçalves (2019, p. 286)

Como já mencionado, a escolha pelo uso da AC se dá pela necessidade de controlar também as variáveis extralinguísticas relacionadas com os informantes. Cada uma das cédulas do quadro representa um(a) informante de um perfil social resultante do cruzamento dessas variáveis. Assim, o informante número 084, por exemplo, é representante do perfil social de uma mulher de 26 a 35 anos com o ensino superior completo ou em andamento. Vale ressaltar que a localidade de origem dos entrevistados não foi considerada uma variável, assim a decisão da quantidade de indivíduos representantes de cada município se deu por proporção em relação à densidade demográfica. Dessa forma, o corpus conta com: 4 informantes de Bady Bassit, representados pela cor azul no quadro; 2 informantes de Cedral, representados pela cor verde; 5 informantes de Guapiaçu, representados pela cor amarela; 1 informante de

Ipiguá, representado pela cor rosa; 16 informantes de Mirassol, representados pela cor laranja; 2 informantes de Onda Verde, representados pela cor cinza; e 122 informantes de São José do Rio Preto, que são o restante das células em branco no quadro. Tanto a exclusão da localização geográfica quanto a escolha por ter um informante por célula social foram decisões metodologicamente tomadas para garantir a exequibilidade do projeto (Gonçalves, 2019).

A constituição do banco de dados Iboruna propiciou a condução de várias pesquisas sobre variação e mudança na variedade do noroeste paulista.²³ Dentre as de cunho fonético-fonológico, têm-se, por exemplo, os estudos de: (i) Pavezi (2006), sobre haplologia; (ii) Silveira (2008), Carmo (2009, 2013, 2014, 2018, 2019), Carmo e Tenani (2013) e Carmo e Carlos (2019), acerca do alçamento vocálico de vogais médias pretônicas; (iii) Ramos (2009), que analisa o alçamento e a síncope de vogais médias postônicas mediais; (iv) Ferreira (2010), sobre o apagamento de /d/ em morfema de gerúndio; (v) Carmo e Machado (2023), que versam sobre monotongação do ditongo oral decrescente [ej]; (vi) Dias e Carmo (2021), acerca da metátese; (vii) Carmo e Taborda (2021), sobre apagamento de /r/ em coda silábica; (viii) Bugai (2021, 2024), que investiga o rotacismo; (ix) Rezende e Carmo (2023), sobre monotongação do ditongo oral decrescente [ow]; (x) Prodelik e Carmo (2024), acerca da ditongação diante de /S/; e (xi) o já apresentado artigo de Rezende, Carmo e Fujihara (2024), sobre alternância morfofonológica em P4 de verbos regulares de primeira conjugação. Avança-se em relação ao último estudo por, como detalhado nesta seção, esta pesquisa de Mestrado analisar todos os inquéritos que compõem a AC do banco de dados Iboruna.

Na próxima seção, passa-se para a descrição e discussão das variáveis investigadas.

2.3 VARIÁVEIS INVESTIGADAS

Neste tópico, apresentam-se a variável dependente e as variáveis independentes linguísticas e extralinguísticas analisadas no desenvolvimento desta pesquisa de Mestrado, bem como as justificativas para a escolha de tais variáveis e as hipóteses iniciais a elas relacionadas.

²³ Disponível em: <https://alip.ibilce.unesp.br/producao-cientifica/banco-de-dados-iboruna>. Acesso em: 28 abr. 2025.

2.3.1 Variável Dependente

Este trabalho apresenta como variável dependente a alternância morfofonológica da vogal temática em primeira pessoa do plural de verbos regulares de primeira conjugação do presente e do pretérito perfeito no modo indicativo do PB. Esse processo é explicado como a troca da vogal temática baixa [ẽ] pela realização da média-alta [ê], conforme pode ser visto no exemplo a seguir, retirado do *cópus* de pesquisa:

“ele foi e foi... **andemo(s)** uns:: vinte pé e o menino pegô(u) e falô(u)” [BDI-AC-063-NE:L.345, grifo nosso]

“gente andô(u) num monte de brinquedo **andamo(s)** no na montanha russa do escuro...” [BDI – AC-037- NE: L. 19, grifo nosso]

2.3.2 Variáveis Independentes

As variáveis independentes extralinguísticas investigadas foram: (i) sexo/gênero; (ii) escolaridade; e (iii) faixa etária. Já as variáveis independentes linguísticas consideradas foram: (iv) tempo verbal; (v) apagamento do -s final da desinência /mos/; (vi) assertividade da sentença; (vii) realização de sujeito; e (viii) elemento temporal.

2.3.2.1 Variáveis extralinguísticas

Antes de apresentar as variáveis sociais, é necessário justificar a decisão de excluir a variável *renda familiar* da metodologia desta pesquisa. Durante a construção do banco de dados Iboruna, os pesquisadores do projeto ALIP decidiram ser mais flexíveis quanto a essa variável devido à dificuldade que encontraram ao tentarem preencher determinados perfis sociais. Inicialmente, os pesquisadores hipotetizaram que tal dificuldade era resultado do cruzamento das variáveis *renda* e *escolaridade*, porém, no decorrer da construção do banco de dados, os pesquisadores notaram que a dificuldade de encontrar indivíduos representativos de certos perfis sociais se dava pelo cruzamento das variáveis *escolaridade* e *faixa etária*, possivelmente por conta do alto nível de escolaridade da população da região (Gonçalves, 2007).

Desse modo, as variáveis extralinguísticas controladas na metodologia da pesquisa são: (i) *sexo/gênero* (masculino e feminino); (ii) *escolaridade* (1º ciclo do Ensino Fundamental, 2º ciclo do Ensino Fundamental, Ensino Médio e Ensino Superior); e (iii)

faixa etária (7 a 15 anos, 16 a 25 anos, 26 a 35 anos, 36 a 55 anos e superior a 55 anos).

2.3.2.1.1 *Sexo/gênero*

O controle da variável *sexo/gênero* é recorrente nas pesquisas sociolinguísticas para determinar as relações entre língua e sociedade. Em Trudgill (2000), a concepção é de que as diferenças na forma de falar e usar a língua com relação a *sexo/gênero* se dá porque a língua é um fenômeno social que reflete atitudes sociais. Nos estudos mais basilares da sociolinguística (Trudgill, 2000; Labov (2008), compreende-se que é comum às mulheres, independentemente de outras categorias sociais, utilizarem mais as formas mais prestigiadas da língua em comparação aos homens da mesma classe socioeconômica a que pertencem, justamente por possuírem diferentes papéis sociais que produzem diferentes padrões de comportamento. Sobre esse “consenso” da teoria metodológica da Sociolinguística, Freitag (2015, p. 18) levanta o seguinte questionamento:

Se a Sociolinguística tem como premissa, em tendência ampla, o estudo da relação entre língua e sociedade, precisa considerar que a sociedade muda; se a sociedade muda, as explicações do modelo teórico metodológico deveriam, também, mudar: a explicação de as mulheres preferirem as formas padrão ou não estigmatizadas, por conta de seu papel como mães e educadoras, talvez fosse válida e pertinente nos anos 1960; hoje, não se pode dizer que é este o papel das mulheres na sociedade.

Freitag (2015) também questiona o binarismo *sexo/gênero*, refletindo que atualmente as questões relativas a essa categoria apresentam uma diversidade de possibilidades e identificações que vão além daquelas relativas ao sexo biológico, o qual de modo geral vem sendo utilizado e considerado pelos bancos de dados de fala, e assim questiona o uso da notação *sexo/gênero*, pois une em uma única variável conceitos distintos. Embora a par dessa discussão, optou-se por ainda assim utilizar a nomenclatura *sexo/gênero* para manter o padrão de uso de nomenclatura do banco de dados Iboruna. Nesta pesquisa, com o objetivo de verificar a influência dessa variável para o processo investigado, mantém-se como base a premissa laboviana (2008) de que as mulheres utilizam menos as formas estigmatizadas em comparação aos homens. Dessa forma, a hipótese é a de que os resultados desta pesquisa irão ao encontro da perspectiva de Labov (2008), obtendo maior ocorrência da variante menos prestigiada na fala masculina.

2.3.2.1.2 *Escolaridade*

A variável *escolaridade* foi estratificada com os seguintes fatores: (i) *1º ciclo do Ensino Fundamental*; (ii) *2º ciclo do Ensino Fundamental*; (iii) *Ensino Médio*; e (iv) *Ensino Superior*. Dessa forma, é possível ter um parâmetro amplo do comportamento dos falantes com o processo variável em distintos níveis de escolarização. Essa variável também é comumente selecionada para estudos de variação, pois falantes com um maior nível de escolaridade costumam apresentar distinções no uso da língua em comparação aos falantes com menor nível de escolaridade. Quanto mais tempo de estudo o indivíduo tiver, mais ele tende a reproduzir as variantes mais prestigiadas do PB. O papel da escola tem grande influência nessa tendência, pois a escola é uma promotora das formas linguísticas que circulam nos setores mais intelectualizados da sociedade (Schwindt *et al.*, 2007). Tal resultado pode ser atestado em Pereira (2014, 2021), que indicou que a *escolaridade* foi influente para o uso da variante não canônica, visto que esta foi mais reproduzida por indivíduos com índices mais baixos de escolaridade. A hipótese deste trabalho é a de que os resultados sejam semelhantes aos das pesquisas mencionadas, ou seja, é esperado que a variante [ẽ] seja mais recorrente na fala de indivíduos que estudaram até o 1º ou 2º ciclo do Ensino Fundamental.

2.3.2.1.3 *Faixa etária*

A variável *faixa etária* foi estratificada da seguinte forma: (i) 7 a 15 anos; (ii) 16 a 25 anos; (iii) 26 a 35 anos; (iv) 36 a 55 anos; e (v) superior a 55 anos. Assim como as outras duas variáveis sociais descritas anteriormente, a faixa etária também é recorrentemente controlada em pesquisas sociolinguísticas, pois a idade que o falante possui pode refletir diretamente na variação e mudança da língua. A partir da análise desse grupo de fatores, é possível determinar se a variação está estabilizada ou em processo de mudança. Essa variável permite esse tipo de conclusão, pois, conforme Naro (2013), o processo de aquisição da linguagem se encerra por volta do início da puberdade e, a partir desse momento, a gramática do indivíduo não sofre maiores alterações. Logo, a faixa etária, justamente por conta dessa estabilidade, possibilita analisar a evolução da língua no tempo. Como já mencionado, este trabalho faz uma análise em tempo aparente, e a hipótese inicial é a de que a variante não canônica esteja presente em todas as faixas etárias.

2.3.2.2 Variáveis linguísticas

As variáveis linguísticas controladas na metodologia desta pesquisa e que são mais bem descritas a seguir são: (i) *tempo verbal* (presente e pretérito perfeito); (ii) *presença do -s final da desinência /mos/* (presente e ausente); (iii) *assertividade da sentença* (positiva e negativa); (iv) *realização do sujeito* (nulo e explícito); (v) *elemento temporal* (ausente e presente). A escolha dessas variáveis e seus fatores se deu com base na pesquisa de Pereira (2014) acerca do fenômeno na região sul do país, e a hipótese inicial sobre elas era a de que não exerceriam forte influência para a produção da variação, salvo a variável *tempo verbal*, pois, com base em pesquisas anteriores, esperava-se que sentenças em contexto de pretérito seriam mais determinantes para a aplicação da regra de variação e do apagamento do -s final da DNP /mo(s)/.

2.3.2.2.1 Tempo Verbal

No PB, alguns verbos dispõem da mesma estrutura para expressar flexões de tempo distintas, isso significa que certos verbos regulares podem apresentar o mesmo radical, a mesma vogal temática e o mesmo sufixo número-pessoal para indicar presente e passado, por exemplo. Essa particularidade ocorre nos verbos objeto de estudo desta pesquisa, pois tanto o presente quanto o pretérito perfeito do modo indicativo apresentam a mesma estrutura morfológica (Câmara Júnior, 2015). Segundo Monteiro (2002), ocorre *neutralização*, definida pela impossibilidade de opor duas estruturas que se tornam homônimas. Gonçalves (2019, p. 84) especifica a *homomorfia*, que ocorre quando “há uma única forma associada a diferentes instâncias de significado”, mas a nível morfológico, como é o caso do morfema investigado nesta pesquisa. A questão problematizadora da neutralização ou homonímia é a *polissemia*, uma mesma forma podendo representar significados diferentes, ocasionando uma ambiguidade na interpretação, o que pode fazer os falantes buscarem uma alternativa para contornar essa questão.

Pesquisas como as de Amaral (1920), Costa, (1990), Zilles, Maya e Silva (2000), Pereira, (2014), Pereira (2021) e Pereira, Lehmkuhl-Coelho e Loregian-Penkal (2016) demonstraram a relação existente entre a variação morfofonológica /amo(s)/ ~ /emo(s)/ com o pretérito perfeito do modo indicativo, justamente como uma tentativa de driblar a

ambiguidade existente. Assim, a hipótese inicial para a variável *tempo verbal* é a de que as sentenças em contexto de passado estariam propiciando o uso da variante não canônica. Desse modo, os fatores que constituem essa variável são:

- *Pretérito*

- (1) ah... a gente chegô(u) na CHÁcara... eu e o I. pegô(u) trocô(u) de RO(u)PA...nadô(u) um PO(u)CO... é::... fomo(s) lá no campo de futebol:: **brinCAmo(s)**... tomamo(s) o café da maNHÃ... [BDI-AC-03-NE: L. 6, grifo nosso]

- *Presente*

- (2) uma coisa que nós **gostamo(s)** muito de fazê(r) né?... é pescá(r) [BDI-AC-11-DE: L. 142, grifo nosso]

2.3.2.2.2 Presença do -s final da desinência /mos/

No PB, a concordância verbal de primeira pessoa do plural pode ser feita como /mos/, forma considerada canônica, ou como /mo/, forma considerada não canônica (*nós falamos* ~ *nós falamo*). Ainda há uma terceira forma, com o apagamento de toda a DNP (*nós falamos* ~ *nós fala*) (Bortoni-Ricardo, 2011). Logo, a concordância verbal com 1ª pessoa do plural é uma regra variável no PB e há diversos estudos que a analisam como variável dependente.

Com base em trabalhos anteriores (Bortoni-Ricardo, 2011; Costa, 1990; Foeger, 2014; Pereira, 2014; Pereira; Margotti, 2018; Pereira, 2021; Zilles; Maya; Silva, 2000), é possível se observar uma forte correlação da concordância verbal em P4 com a elevação da vogal temática, principalmente quando associada à variedade caipira, conforme afirma Bortoni-Ricardo (2011, p. 234): “na variedade caipira, a variante /mu/ geralmente coocorre com a mudança da vogal temática (/a/>/e/) nos pretéritos da primeira conjugação”. Assim, nossa hipótese inicial segue o que foi constatado nessas pesquisas, e espera-se que o apagamento do -s final esteja correlacionado ao uso da forma não canônica.

Ademais, Pereira (2021) questiona o uso da expressão *apagamento do -s* para se referir à regra variável, pois, como a ausência do -s tem sido descrita como marca do dialeto caipira, a autora considera que os falantes aprendem desde muito novos a utilizar

a variante /mo/, tratando-se assim de uma realização não padrão ou não prototípica, e não um *apagamento*, conforme afirmado em “acreditamos que é mais provável que os falantes da comunidade de fala investigada nesta pesquisa adquiram a forma não-padrão (sem o /s/) e, apenas depois, mediante escolarização, insiram este fonema no sufixo modo temporal” (Pereira, 2021, p. 310). Então, os fatores para essa variável são:

- Presença do -s final da desinência
- (3) ela cheGÔ(u)...e **entramos** no::... restaurante... [BDI-AC-13-NE: L. 19, grifo nosso]
- Ausência do -s final da desinência
- (4) ela... me agradeceu tal nós **ficamo(s)** tudo de boa de novo... [BDI-AC-19-NE: L. 14, grifo nosso]

2.3.2.2.3 Assertividade da sentença

Seguindo a metodologia de Pereira (2014), a variável *assertividade da sentença* foi investigada com a finalidade de averiguar em que grau sentenças afirmativas ou negativas poderiam afetar a troca da vogal temática. Porém, já com base nos próprios resultados de Pereira (2014), a hipótese inicial era a de que essa variável não seria selecionada como relevante para o fenômeno, por conta de outra hipótese, a de que essa variação traz influências de uma desneutralização de tempo verbal. Os fatores dessa variável linguísticas são:

- Sentença negativa
- (5) aí terminamo(s) **nunca mais conversamo(s)**... [BDI-AC-22-NE: L.115, grifo nosso]
- Sentença positiva
- (6) tudo era novida::de **passeamo(s)** bastante comemo(s) muito pe(i)xe... [BDI-AC-24-NE: L. 70, grifo nosso]

2.3.2.2.4 Realização do sujeito

Conforme a língua está evoluindo, gradativamente o PB vem aos poucos perdendo sua tendência de não preenchimento do sujeito. Uma hipótese da motivação dessa nova tendência pode ser pelo encaixamento na estrutura linguística da mudança do sistema flexional verbal que vem ocorrendo no PB já há algum tempo (Duarte, 1995). Com a substituição de pronomes tais quais *tu* e *nós* para *você* e *a gente*, respectivamente, houve uma redução no sistema flexional do PB. De cinco formas verbais distintivas, passa-se a ter três, ou até mesmo duas, dependendo da realização da concordância verbal em terceira pessoa do plural (Duarte, 1995). Isso ocorre porque as variantes *você* e *a gente* demandam uma flexão verbal para gerar concordância diferente das demandadas pelas variantes *tu* e *nós*. Em consequência, perde-se a informação de número-pessoa, que é recuperada com a evidenciação do sujeito na sentença. Guy (2006, p. 141) afirma que “em situações onde está envolvida uma redução no quadro flexional, pode-se observar uma tendência ao preenchimento [do sujeito]”.

Baseando-se em Pereira (2014), a hipótese inicial para essa variável seria a de que a sentença com sujeito preenchido favoreceria o uso da variante /e/, pois ambas são consideradas não canônicas no PB. Por tal motivo também se espera que esses contextos estejam interrelacionados com o nível de escolaridade mais baixo, pois o sujeito expresso é mais facilmente implementado na fala de pessoas menos escolarizadas (Duarte, 1995). Os fatores da variável *realização do sujeito* são:

- Sujeito explícito

(7) aí **nós mandamo(s)** rancá(r) o café:... [BDI-AC-114-DE: L. 398, grifo nosso]

- Sujeito nulo

(8) e:... **trocamo(s)** o piso do... do banhe(i)ro... e da sala... [BDI-AC-119-DE: L. 154, grifo nosso]

2.3.2.2.5 Elemento temporal

Por fim, a variável *elemento temporal* tem o objetivo de investigar a presença de adjuntos adverbiais de tempo ou advérbios de tempo nas sentenças. Essa variável pode estar correlacionada com a alternância da vogal temática, pois o uso de elementos que

demarcam temporalização no espaço pode auxiliar o falante a desneutralizar as formas em ambiguidades de presente e pretérito perfeito do modo indicativo. A hipótese inicial para essa variável é a influência do uso da variante não canônica quando não houver elemento indicador temporal na sentença. O falante, ao não produzir nenhum indicador de tempo, adota outra estratégia para demarcar a diferença, no caso a variação morfofonológica objeto de estudo desta pesquisa. Os fatores que compõem essa variável são:

- Presença do elemento temporal

(9) uma ditadura aí que nós já **passamo(s) naquela época** né?... [BDI- AC-144-RO: L. 722, grifo nosso]
- Ausência do elemento temporal

(10) O um o dois... e o três... nós **moramo(s)** no dois... aqui no bairro tem tudo que cê precisa...[BDI-AC-136-DE: L. 139, grifo nosso].

2.4 PASSOS METODOLÓGICOS

A metodologia adotada para a execução deste trabalho envolvendo a variação morfofonológica em questão, como já mencionado anteriormente, é a Teoria da Variação e Mudança Linguística, a qual tem como seu maior representante William Labov (2008), que considera que os estudos de mudança linguística devem tomar como pressuposto a heterogeneidade da língua e que a língua sofre ação de fatores de cunho social e linguístico.

Segundo Oushiro (2014), o sociolinguista apresenta as seguintes tarefas: (i) a coleta de dados, comumente através de gravações de entrevistas com falantes de uma comunidade prelimitada; (ii) a transcrição das gravações coletadas; (iii) a definição de uma variável sociolinguística e de seus contextos variáveis; (iv) o levantamento de ocorrências no corpus de entrevistas; (v) o levantamento de hipóteses sobre fatores sociais e linguísticos envolvidos no processo; (vi) a codificação das ocorrências de acordo com as hipóteses levantadas; (vii) a análise quantitativa dos dados, geralmente no GoldVarb-X ou RBrul; e, por fim, (viii) a interpretação de resultados obtidos. Segundo Brescancini (2002), a metodologia laboviana contém os seguintes passos: (i) definição da variável dependente; (ii) definição

das variáveis independentes linguísticas e extralinguísticas; (iii) recorte da amostra de dados; (iv) levantamento dos dados; (v) transcrição dos dados; (vi) codificação e quantificação dos dados; e (vii) interpretação dos resultados obtidos.

Dessa forma, a variável dependente analisada é a variação da vogal que antecede o sufixo número-pessoal em verbos regulares de 1ª conjugação do presente e do pretérito perfeito do modo indicativo, isto é, a variação morfofonológica /amos/ ~ /emos/. Esse processo variável foi escolhido porque se observou uma lacuna de estudos produzidos acerca de tal objeto na variedade analisada e há pouca produção de trabalhos sobre o tema em outras variedades no PB em geral.

Já a definição das variáveis independentes se deu pelo embasamento de pesquisas anteriores correlacionadas ao fenômeno, como a pesquisa de Pereira (2014), que analisa a variação morfêmica na primeira pessoa do plural no presente e no pretérito do modo indicativo, utilizando o banco de dados VARSUL.

A amostra utilizada, como também já mencionado, foi a do banco de dados Iboruna, representativo da variedade do PB falado na região de São José do Rio Preto. Como essa pesquisa é decorrente de trabalhos anteriores (Rezende, 2021; Rezende; Carmo; Fujihara, 2024), optou-se por dar continuidade, ampliando o *cópus* de análise, consequentemente, esta pesquisa utiliza todos os 151 inquéritos presentes no banco de dados, com a finalidade assim de também ampliar os resultados obtidos, uma vez que os trabalhos anteriores resultaram na baixa ocorrência do processo variável.

Prosseguindo os passos metodológicos, após a ampliação do *cópus*, foi realizado o levantamento de todos os contextos que possibilitam a aplicação da variação, ou seja, foi feito o levantamento de todos os verbos regulares do presente e pretérito perfeito do modo indicativo, pois “para fazer uma generalização razoável, é necessário considerar todas as observações relevantes, não só um subconjunto selecionado de forma enviesada” (Guy; Zilles, 2007, p. 81).

A próxima etapa foi a análise de oitiva dos inquéritos. Esse passo metodológico foi auxiliado pela transcrição ortográfica das entrevistas, que servem de guias construídos no mesmo período de elaboração do banco de dados Iboruna e estão disponíveis para acesso juntamente com os arquivos de áudio e as fichas sociais dos informantes. Então, os dados coletados foram codificados, tabelados e categorizados quanto às variáveis internas e externas.

Dando seguimento, o próximo passo corresponde à execução da análise quantitativa dos dados, a realização desse tipo de análise “possibilita o estudo da

variação linguística, permitindo ao pesquisador apreender sua sistematicidade, seu encaixamento linguístico e social e sua eventual relação com a mudança linguística” (Guy; Zilles, 2007, p. 73). Foi o advento de métodos estatísticos que permitiu que os estudos de variação deixassem de ser considerados secundários e aleatórios e passassem a serem entendidos como instrumentos de compreensão de dinâmicas sociais e passíveis de serem cientificamente compreendidos (Guy; Zilles, 2007).

A análise quantitativa deste trabalho foi realizada a partir de rodadas estatísticas no programa computacional GoldVarb-X (Sankoff; Tagliamonte; Smith, 2020). Esse programa faz avaliações das relações estatísticas existentes entre as variáveis em diferentes contextos, fornece informações como o percentual e a probabilidade de ocorrência do fenômeno em relação aos fatores de cada variável, o PR para cada um desses fatores, além de indicar quais dos fatores favorecem ou desfavorecem a aplicação das regras. Assim, a análise envolve “a contagem das ocorrências da variável, a descrição de tendências e da extensão da variabilidade, bem como das restrições ou fatores que a influenciam, mediante métodos estatísticos” (Guy; Zilles, 2007, p. 50).

O PR é uma medida quantitativa que indica a força que um fator linguístico ou extralinguístico possui na dinâmica da variação e a probabilidade desse evento ocorrer dentro de um contexto específico. Esse indicador permite que o sociolinguista verifique os fatores influentes para a aplicação da regra variável e assim teste suas hipóteses e identifique padrões. Cada fator recebe um peso, que é “um número entre 0 e 1, que caracteriza o efeito deste fator sobre a regra variável em questão” (Guy; Zilles, 2007, p. 51). Quanto maior esse número, ou seja, mais próximo do 1, maior a chance de a regra variável ser aplicada em função de determinado fator (Guy; Zilles, 2007). Por exemplo, Pereira (2021), analisando a variação da vogal temática [ã] ~ [ẽ] em verbos regulares de P4, controlou a variável *grau de ruralidade*. O PR do fator *grau alto de ruralidade* foi 0.646, enquanto o fator *grau de ruralidade baixo* obteve PR = 0.155, isso significa que, na dinâmica dessa variação, a probabilidade de o informante aplicar a variação da vogal temática e utilizar a variante [ẽ] é maior quando ele advém de um contexto de maior grau de ruralidade. O peso é relativo ao valor geral de ocorrências da variação, o *input*.

O *input* é um indicador quantitativo que, em conjunto dos pesos relativos, corresponde aos efeitos dos fatores. O valor do *input* “representa o nível geral de uso de determinado valor da variável dependente” (Guy; Zilles, 2007, p. 238). Quando seu valor se distancia da taxa geral, isto é, a porcentagem geral de ocorrência da aplicação da regra em análise no total de informantes, pode indicar que a distribuição dos dados nos

fatores não está equilibrada (Guy; Zilles, 2007).

Por fim, foram realizadas a análise e a interpretação dos resultados obtidos pelo programa estatístico, descritas no próximo capítulo.

3 ANÁLISE DOS RESULTADOS

Este capítulo apresenta a análise dos dados coletados no banco de dados Iboruna, com o objetivo de investigar a variação morfofonológica que pode ocorrer nas vogais temáticas de verbos regulares de primeira conjugação, na forma de primeira pessoa do plural, nos tempos verbais presente e pretérito perfeito do modo indicativo. A pesquisa foi conduzida a partir de um corpus constituído por falantes da variedade linguística falada na região noroeste do estado de São Paulo. A análise buscou identificar padrões de uso e de fatores sociais e linguísticos que condicionam a ocorrência da variação na VT, para contribuir na compreensão das dinâmicas da variação linguística no PB dessa localidade.

Com a análise de oitiva das 151 entrevistas do banco de dados Iboruna, foram levantados 642 dados de verbos conjugados na primeira pessoa do plural. Desses dados gerais, 6 (0,9%) apresentaram a variante [ẽ], conforme a tabela:

Tabela 2 – Ocorrências gerais

	Ocorrências	Porcentagens
Aplicação	6	0,9%
Não aplicação	636	99,1%
Total	642	100%

Fonte: elaboração própria

Os dados encontrados foram duas ocorrências de *cheg[ẽ]mo(s)* e um caso de *and[ẽ]mo(s)*, *levant[ẽ]mo(s)*, *peg[ẽ]mo(s)* e *fiq[ẽ]mo(s)*. Esses 6 dados foram retirados dos seguintes contextos nos inquéritos do banco de dados:

- ele foi e foi... **andemo(s)** uns:: vinte pé e o menino pegô(u) e falô(u) [BDI-AC-063-NE:L.345, grifo nosso]
- fui lá... a S. foi junto... e eu e ela... **cheguemo(s)** lá... colocô:: [BDI-AC-063-NE:L.341- 342, grifo nosso]
- aí nos voltô(u) pa:: trás aí meu tio foi com nós **cheguemo(s)** lá [BDI-AC-063-NR:L.603, grifo nosso]
- embolô(u) em cima de mim... e nós caiu po chão **levantemo(s)** e em vez de chegá(r) [até em casa] nós voltô(u) pra trás [BDI-AC-063-NR:L.599-600, grifo nosso]

- aí **pegamo(s)**²⁴ e descemo(s) aqui pra ba(i)xo... ela morava aqui na rua::... a:: São Bento né?... [BDI-AC-127-NE:L.20-21, grifo nosso]
- e fomo(s) com o pessoal da terceira idade tam(b)ém... um político lá::... **fiquemo(s)** lá tres di::as ali::... a turma né?... [BDI-AC-140- NR: L. 250, grifo nosso]

Das 151 entrevistas, 3 apresentaram dados com aplicação, sendo 4 delas pelo informante 63, uma pelo informante 127 e uma pelo 140. No quadro a seguir, é possível observar o perfil social desses informantes:

Quadro 3 – Perfil social dos informantes

	AC - 063	AC – 127	AC - 140
Sexo/gênero	Masculino	Masculino	Feminino
Faixa etária	36 anos	64 anos	60 anos
Escolaridade	1º Ciclo do EF	1º Ciclo do EF	Ensino Médio (supletivo)

Fonte: elaboração própria

Três desses quatro itens lexicais aparecem também no corpus sem a aplicação do fenômeno, ou seja, com a manutenção da vogal [ã], como nos seguintes exemplos:

- eu lembro perfeitamente como se fosse hoje... da praia da/ do mar::... e a areia assim também que nós... **andamos** [BDI-AC-064-DE:L.92-93, grifo nosso]
- aí a gente foi no convênio... **chegamos** no convênio tal eu desespera::do... [BDI- AC-055-NE:L.48-49, grifo nosso]
- pa... viajá(r) no dia seguinte... então no dia seguinte... **pegamos** o ô::nibus tu::do nós fomos pra São Paulo [BDI-AC-087-NE:L.12-13, grifo nosso]

Ainda conforme a tabela 2, os resultados mostram que houve uma baixa produtividade de dados. A variedade do PB falada em Florianópolis também apresentou uma baixa produtividade dos dados, 30 ocorrências no geral e 12 da variante não canônica. Uma possível explicação pode ser a preferência dos falantes pelo uso da variante *a gente* para primeira pessoa do plural, assim como ocorreu com os resultados obtidos por Pereira (2014) na variedade dos bairros afastados dos centros urbanos da

²⁴ Apesar de a transcrição desse dado estar em conformidade com a variante canônica, a metodologia tomou por base a análise de oitiva e não essa transcrição; assim, esse dado corresponde à variante não canônica. Contudo, foi decidido não alterar a transcrição original do banco de dados Iboruna, a fim de manter a padronização desses dados.

cidade de Florianópolis. Os seguintes exemplos retirados do *cópus* de pesquisa ilustram a afirmação:

- ele levô(u) eu pra casa da mãe dele e **a gente** ficô(u) moran(d)o lá [BDI-AC-032-NE:L.48-49, grifo nosso]
- esperei UM ANO pra ficá(r) com ele... mas aí **a gente...** ficô(u) junto né?... [BDI-AC-056-NE:L.5-6, grifo nosso]
- [...] aí quando foi no dia... **a gente** combinô(u) de encontrá(r) lá na praça e tal daí encontramo(s) lá... [BDI-AC-127-NE:L.17-18, grifo nosso]

Ao fazer essa escolha, o falante tende a não produzir o contexto linguístico necessário para que ocorra a variação na vogal temática, pois a concordância com o uso de *a gente* geralmente se dá em verbos conjugados na terceira pessoa do singular.

Outra possibilidade para a baixa de dados de Pereira (2014) é o fato desse processo variável ser característico do dialeto rural (Amaral, 2020; Bortoni-Ricardo, 2011), e ruralidade não é uma característica da variedade de Florianópolis, por esta conter traços mais urbanos. A pesquisa de Pereira e Margotti (2018) pode confirmar ainda mais essa possibilidade, pois, com os resultados obtidos a partir da carta linguística n.º 83 do ALERS, os autores identificaram que, na variedade da região Sul, a variante não canônica é mais produzida nas regiões rurais, principalmente na região oeste de Santa Catarina e a região centro-sul e sudeste do Paraná (Pereira; Margotti, 2018). Já na variedade de Tefé – AM, a produtividade de dados foi um pouco mais alta, apresentando 216 dados gerais de verbos de primeira conjugação em P4, porém comparando os dados da variante canônica e não canônica, a variedade também obteve uma diferença significativa, pois 15,28% dos dados aplicaram a regra variável, ou seja, 33 dados, enquanto 84,72% não aplicaram a regra.

Nos próximos subcapítulos, são apresentados os resultados obtidos para cada variável após a rodada dos dados no programa GoldVarb-X.

3.1 RESULTADOS DA VARIÁVEL *ESCOLARIDADE*

Após a rodada, o programa indicou vários nocautes (resultados categóricos de 0% ou 100%). A tabela 3 exhibe os resultados obtidos para a variável *escolaridade*. Os fatores 2º Ciclo do Ensino Fundamental e Ensino Superior não obtiveram casos de aplicação da regra variável, ou seja, caracterizam casos de nocaute. Porém, pensando na hipótese inicial de que as ocorrências da forma menos prestigiada seriam mais comuns na fala de

indivíduos menos escolarizados, os resultados aqui obtidos aparentam caminhar nessa direção, visto que 5 dados (3,1%) de dois informantes diferentes do 1º Ciclo EF apresentaram a aplicação da regra variável, enquanto somente um dado (0,6%) de um outro informante do EM²⁵ obteve o uso da mesma variante.

Tabela 3 – Ocorrências em relação à escolaridade

	Aplicação/ocorrências	Porcentagens
1º Ciclo EF	5/163	3,1%
2º Ciclo EF	0/120	0%
Ensino Médio	1/173	0,6%
Ensino Superior	0/186	0%
Total	6/642	0,9%

Fonte: elaboração própria

Ademais, também em relação ao informante do nível de escolaridade Ensino Médio que produziu um único dado da variante não canônica, é possível notar que também produziu um dado da variante canônica, como pode ser visto no trecho retirado do corpus:

- aí nós fomos viajô(u) de madrugada até clarean(d)o o dia **chegamo(s)** lá...
chegan(d)o lá tem... [BDI-AC-140- DE: L. 221, grifo nosso].

Os resultados da variedade de Florianópolis (Pereira, 2014) respectivos à variável *escolaridade* apresentaram características semelhantes às obtidas para a variedade de São José do Rio Preto. Os informantes mais escolarizados não produziram nenhum dado da variante não canônica, assim como os informantes do Ensino Superior desta pesquisa. A mesma tendência pode ser verificada na variedade das regiões rurais eslavo-brasileiras no sudeste do Paraná, em que Pereira (2021) identificou maior uso da variante não canônica no repertório linguístico dos informantes com escolaridade *até fundamental I*, enquanto o nível *ensino médio* propiciou o uso da variante canônica. A variedade de Tefé também não se distanciou muito dos resultados obtidos por esta pesquisa e de outras variedades, já que os resultados de Silva (2023) indicaram uma influência da escolaridade mais baixa para o uso da variante não canônica, assim 21,84% dos dados dessas variantes foram de informantes do Ensino Fundamental, enquanto 10,35% dos dados de aplicação foram dos informantes do Ensino Médio. As justificativas giram em torno do papel da escola na formação dos indivíduos, pois é um espaço de promoção

²⁵ A ficha social desse informante esclarece que o nível de escolaridade Ensino Médio foi concluído por ele na modalidade Ensino de Jovens e Adultos.

das variantes mais prestigiadas da língua.

3.2 RESULTADOS DA VARIÁVEL *FAIXA ETÁRIA*

Com relação à *faixa etária*, os nocautes ocorreram nos fatores *7 a 15 anos*; *16 a 25 anos*; e *36 a 55 anos*, pois não apresentaram dados com aplicação da regra variável. Dos seis dados totais, quatro (3,1%) ocorreram em falantes da faixa etária *26 a 35 anos* e dois (1,6%) em falantes da faixa etária *superior a 55 anos*. A hipótese inicial levantada era a de que o processo variável poderia ser encontrado em todas as faixas etárias.

Tabela 4 – Ocorrências em relação à *faixa etária*

	Aplicação/ocorrências	Porcentagens
7 a 15 anos	0/83	0%
16 a 25 anos	0/96	0%
26 a 35 anos	4/131	3,1%
36 a 55 anos	0/206	0%
+ 55 anos	2/126	1,6%
Total	6/642	0,9%

Fonte: elaboração própria

Os resultados da variedade de Florianópolis (Pereira, 2014) indicaram que a *faixa etária* pode estar exercendo forças para a aplicação da regra variável, pois os mais jovens foram responsáveis pela produção de 83% dos dados da variante canônica. Apesar disso, seus resultados mostram uma força muito maior para a influência da variável *escolaridade*, pois seus resultados mostraram que os mais escolarizados independentemente da idade não produziram a variante não canônica, semelhante ao que se obteve nesta pesquisa da variedade de São José do Rio Preto, no qual os informantes com nível superior não produziram a forma não canônica. Na variedade de Tefé, são informantes mais velhos responsáveis pelo maior uso da variante [ê], diferentemente do que parece ocorrer na variedade de São José do Rio Preto e na variedade do sudeste do Paraná.

Na variedade das regiões de influência eslava do sudeste do Paraná (Pereira, 2021), como mencionado, os resultados foram relativamente diferentes. Apesar dos mais velhos demonstrarem uma porcentagem de uso maior da variante não canônica (47,10%) do que os mais jovens (43,30%), os pesos relativos indicaram uma força de atuação maior da faixa etária mais jovem (até 50 anos) para a produção do fenômeno. A

justificativa de Pereira (2021) para os mais jovens gerarem esse peso gira em torno do fato de que essa variedade do sudeste do Paraná tem características de uma variedade rural e, portanto, os informantes possuem um estilo de vida diferente dos indivíduos que vivem em zonas mais urbanizadas. As ocupações profissionais das pessoas do campo geralmente não demandam o uso das variantes mais prestigiadas, como Pereira (2021, p. 354) afirma em “precisamos também levar em conta a profissão que exercem, considerando-se que a estrutura social no campo difere expressivamente em relação à composição social de realidades urbanas”.

3.3 RESULTADOS DA VARIÁVEL SEXO/GÊNERO

No que se refere aos resultados obtidos relativos a *sexo/gênero*, visíveis na Tabela 5, é possível observar que cinco dos seis casos da variante [ẽ] não canônica ocorreram na fala masculina e um na fala feminina. Um olhar inicial indica uma tendência comum nos resultados envolvendo esses fatores, uma vez que o próprio Labov (2008) teoriza sobre a tendência de as mulheres se aterem a produzir as formas mais prestigiadas da língua por busca de adentramento a espaços sociais historicamente acessados pelo gênero masculino.

Tabela 5 – Ocorrências em relação ao sexo/gênero

	Aplicação/ocorrências	Porcentagens
Masculino	5/357	1,4%
Feminino	1/285	0,4%
Total	6/642	0,9%

Fonte: elaboração própria

A variedade de Florianópolis (Pereira, 2014) demonstrou resultados que vão em direção contrária ao que foi obtido na variedade de São José do Rio Preto, pois as mulheres dessa comunidade foram responsáveis pela maior produção da variante não canônica (53,3%), enquanto os homens produziram mais a variante canônica (68%). Porém, realizando o cruzamento com a variável *escolaridade*, Pereira (2014) identificou que todos os dados da variante canônica do sexo/gênero masculino foram produzidos pelo mesmo informante homem.

Já para a variedade do sudeste do Paraná (Pereira, 2021), essa variável não foi selecionada pelo programa estatístico como atuante na produção da variação

morfofonológica. Os resultados obtidos para essa variável foram de certa forma equilibrados. Dos dados gerais da variante canônica, 52,9% foram produzidos por homens e 57,3% por mulheres, assim como para a variante não canônica, os homens foram responsáveis por produzir 47,1% dos dados e as mulheres 42%.

Na variedade de Tefé no Amazonas (Silva, 2023), os resultados mostraram que tanto o repertório linguístico dos homens quanto o das mulheres contam com o uso da variante canônica de forma mais preponderante, porém ambos os gêneros produziram também a forma não canônica. Os resultados da variedade de Tefé se assemelham ao que foi averiguado na variedade de Florianópolis, pois em Tefé também foram as mulheres que produziram com maior frequência a variante não canônica.

Tanto Silva (2023) quanto Pereira (2014) se baseiam em Labov (2008) para justificar essa tendência de as mulheres em ambas as variedades utilizarem a variante menos prestigiada, pois, quando a variação linguística estudada está sendo precursora de uma mudança linguística, geralmente são as mulheres as mais envolvidas para motivar tal mudança.

Silva (2023) ainda discute o papel da mulher na sociedade atualmente e especialmente no município de Tefé. O resultado de que as mulheres mais velhas são as que fazem o uso da variante não canônica pode indicar uma mudança no papel da mulher nessa região. Assim, além das funções relacionadas aos serviços domésticos e familiares, com o avanço das leis trabalhistas e dos movimentos feministas, as mulheres também estão cada vez mais em espaços mais produtivos e ocupando outras instâncias da sociedade a partir de sua força de trabalho (Silva, 2023).

3.4 RESULTADOS DA VARIÁVEL *TEMPO VERBAL*

Os resultados das ocorrências relativas a tempo verbal também revelaram nocaute. Todos os dados com aplicação da regra variável se deram em contexto de sentenças em *pretérito perfeito*. Verificando as ocorrências gerais dessa variável, é possível perceber que a quantidade de dados no pretérito perfeito também foi maior do que a quantidade de dados no presente.

Tabela 6 – Ocorrências em relação ao *tempo verbal*

	Aplicação/ocorrências	Porcentagens
Pretérito Perfeito	6/576	1,0%
Presente	0/66	0%
Total	6/642	0,9%

Fonte: elaboração própria

A hipótese inicial, tomando por bases outros estudos (Amaral, 1920; Costa, 1990; Zilles; Maya; Silva, 2000; Pereira, 2014; Pereira, 2021; Pereira; Coelho; Loregian-Penkal, 2016), era a de que o fator *pretérito perfeito* exerceria mais influência para o uso da variante com [ẽ], corroborando a teoria de que essa variação morfofonológica tem origem na desneutralização das formas de presente e pretérito, para que seja possível fazer a demarcação dessa diferença. Os resultados para a variedade da região de São José do Rio Preto parecem seguir essa tendência, mas a baixa produtividade de dados dificulta traçar explicações mais afirmativas.

Na variedade do noroeste do Rio Grande do Sul, Costa (1990), apesar de ter identificado a variação morfofonológica em ambos os contextos temporais, notou que sentenças no contexto do pretérito perfeito são as que favorecem a aplicação da regra variável; de forma similar, a variedade de Florianópolis (Pereira, 2014) apresentou resultados com as mesmas características, pois em ambos os contextos verbais a variante menos prestigiada pode ser identificada, mas foi o contexto de pretérito perfeito que favoreceu o uso da forma [ẽ] por ter resultado em uma maior porcentagem de uso (45%) em comparação ao presente (36%).

A variedade do sudeste do Paraná (Pereira, 2021) também obteve a variante [ẽ] em ambos os contextos, mas o uso mais expressivo dessa variante também se deu no pretérito perfeito, com uma porcentagem de 74,40%. Esse valor demonstrou alto contraste com a porcentagem de uso da mesma variante não canônica para o tempo presente (1,7%). Sobre esses resultados, Pereira (2021) afirma então que essa variação nessa variedade pode estar relacionada com a demarcação da diferença entre pretérito e presente na sentença para desambiguar ambos os contextos. Outro fato observado por Pereira (2021) é que, para o pretérito, as formas que parecem estar em variação realmente são a alternância entre [ã] ~ [ẽ], mas, devido ao uso quase categórico da variante [ã] para presente, as formas variáveis para esses contextos parecem ser, por exemplo, *fala* ~ *falamos* em decorrência do uso de *a gente* como sujeito (Pereira, 2021, p. 342).

Por sua vez, os resultados da variedade de Tefé foram muito semelhantes ao da variedade de São José do Rio Preto, pois os dados da variante não canônica se deram categoricamente em contexto de pretérito perfeito, não havendo casos dessa ocorrência no presente do indicativo. Porém, igualmente ao que ocorre na variedade estudada nesta pesquisa, a baixa produtividade de dados obtida na variedade de Tefé não possibilita afirmar que o que vem atuando nessa variação é um processo de desneutralização entre presente e pretérito perfeito do indicativo, demandando assim um estudo mais aprofundado dessa variável e uma ampliação dos dados.

3.5 RESULTADOS DA VARIÁVEL APAGAMENTO DO -S NA DNP /MO(S)/

Os dados da variável *apagamento do -s na DNP /mo(s)/* resultaram em um nocaute no fator *não apagado*, pois não houve produção da variante [ẽ] quando o -s da DNP estava presente na concordância verbal. Todas as ocorrências da variante não canônica se deram em contexto de -s na DNP /mo(s)/ apagado.

Tabela 7 – Ocorrências em relação ao *apagamento do -s na DNP /-mo(s)/*

	Aplicação/ocorrências	Porcentagens
Apagado	6/427	1,4%
Não apagado	0/215	0%
Total	6/642	0,9%

Fonte: elaboração própria

Em conformidade com os resultados dessa variável no estudo de outras variedades do PB, era esperado inicialmente que fosse significativa para a variação morfofonológica ocorrer no repertório dos indivíduos da variedade de São José do Rio Preto.

Ao estudar a variedade do noroeste do Rio Grande do Sul, Costa (1990) considerou as variantes /amos/ ~ /amo/ ~ /emo/ e não considerou a variante /emos/; ao fazer essa escolha, Costa (1990) compreende que a ocorrência de [ẽ] se dá juntamente com o apagamento do -s final da DNP. Da mesma forma, Pereira, Lehmkuhl-Coelho e Loregian-Penkal (2016) também consideram que o apagamento do -s é um traço característico da região e não um condicionante da variante não canônica. Pereira (2021) também aceita que o apagamento do -s é uma característica do dialeto caipira de uso quase categórico e, portanto, é aprendido desde cedo pelos falantes, que só durante o

período de escolarização vão passar a utilizar com mais frequência a variante /mos/. Em contrapartida, esta pesquisa e as de Pereira (2014), Zilles, Maya e Silva (2000) e Silva (2023) apontaram o apagamento do -s final da DNP como um fator motivador para a produção da variante não canônica.

3.6 RESULTADOS DA VARIÁVEL ASSERTIVIDADE

Analisando a variável *assertividade*, notou-se um número mais expressivo de sentenças positivas do que sentenças negativas. Assim, as ocorrências da variante [ẽ] ficaram todas restritas aos contextos de sentenças afirmativas, não havendo dados dessa variante para contexto de sentenças negativas.

Tabela 8 - Ocorrências em relação à *assertividade*

	Aplicação/ocorrências	Porcentagens
Positivas	6/626	1,0%
Negativas	0/16	0%
Total	6/642	0,9%

Fonte: elaboração própria

Então, dado o número desequilibrado de ocorrências gerais, não é possível afirmar categoricamente que a variável *assertividade* exerce ou não motivações para o uso variável da vogal temática. Comparando com as outras variedades, apenas a pesquisa de Pereira (2014) sobre os bairros de Florianópolis considerou essa variável e, após a rodada dos dados, esse conjunto de fatores foi descartado pelo programa estatístico como influente para a variação da vogal temática.

3.7 RESULTADOS DA VARIÁVEL REALIZAÇÃO DO SUJEITO

Sobre os resultados referentes à *realização do sujeito*, quatro dos seis dados com [ẽ] ocorreram em contexto de sujeito nulo, e dois dados ocorreram em contexto de sujeito explícito. Devido à distribuição dos dados entre os fatores, inicialmente essa variável não aparenta ter grandes efeitos sobre o processo.

Tabela 9 - Ocorrências em relação à *realização do sujeito*

	Aplicação/ocorrências	Porcentagens
Nulo	4/357	1,1%
Explícito	2/285	0,7%
Total	6/642	0,9%

Fonte: elaboração própria

Na variedade de Florianópolis (Pereira, 2014), os resultados foram similares, pois o uso da forma não canônica foi favorecido pelo sujeito nulo. Porém, Pereira (2014) faz o cruzamento dessa variável com *apagamento do -s final da DNP* e obtém que não houve ocorrência da variante menos prestigiada quando o sujeito é nulo e com presença do -s final, mas que houve duas ocorrências da variante mais prestigiada também com sujeito nulo e presença de -s. Desse modo, para Pereira (2014), o que pode estar ocorrendo é uma omissão do sujeito quando o -s final não é suprimido, independentemente da variação da vogal temática.

3.8 RESULTADOS DA VARIÁVEL *ELEMENTO TEMPORAL*

Os resultados da variável linguística presença/ausência de *elemento temporal* apontaram que cinco dados com aplicação ocorreram em contextos em que o elemento temporal não estava demarcado na sentença, e somente um em contexto em que estava.

Tabela 10 - Ocorrências em relação ao *elemento temporal*

	Aplicação/ocorrências	Porcentagens
Ausência	5/543	0,9%
Presença	1/99	1,0%
Total	6/642	0,9%

Fonte: elaboração própria

Pela ausência de algum marcador temporal, era esperado que o uso da forma não canônica fosse privilegiado para gerar uma forma alternativa de marcar o tempo ao qual a sentença se refere. Os dados obtidos nessa variável realmente apontaram para mais ocorrências do uso da variante [ê] em contextos em que não há um demarcador de tempo na sentença, o que poderia ser um indício afirmativo da hipótese inicial, porém, verificando o número de ocorrências gerais, nota-se um desequilíbrio relativamente grande entre os dois fatores dessa variável, o que dificulta a conclusão.

Na variedade de Florianópolis (Pereira, 2014), essa variável foi descartada pelo programa por não ser relevante para a variação. Nas outras variedades, como a do

sudeste do Paraná (Pereira, 2021) e a de Tefé, no Amazonas (Silva, 2023), essa variável não foi investigada.

A partir da análise dos dados, é feito, na próxima seção, um quadro comparativo com resultados obtidos em outras variedades do PB.

3.9 RESUMO COMPARATIVO DOS RESULTADOS OBTIDOS EM VARIEDADES DO PB

Quadro 4 – Comparativo dos resultados da variedade do noroeste do estado de SP com os de outras variedades

Autores	Variedade	Resultados	Atual pesquisa
Costa (1990)	Italianos de uma zona rural do RS	Variante [ẽ] com ocorrências no presente e no pretérito, com destaque para contexto de pretérito e apagamento do -s final. Maior ocorrência da forma não canônica do que da canônica.	Variante [ẽ] com menor ocorrência e produtiva somente nos contextos de pretérito. Destaque para contexto de pretérito e apagamento do -s final.
Pereira (2014)	Florianópolis - SC	Baixa produtividade do fenômeno. Variáveis <i>assertividade</i> e <i>advérbio de tempo</i> descartadas. Maiores ocorrências na fala de mulheres e de menos escolarizados. Destaque para sujeito explícito, contexto pretérito e apagamento do -s final.	Baixa produtividade do fenômeno. Variáveis <i>assertividade</i> e <i>advérbio de tempo</i> descartadas. Maiores ocorrências na fala de mulheres e de menos escolarizados. Destaque para sujeito explícito, contexto pretérito e apagamento do -s final.
Pereira, Lehmkuhl-Coelho e Loregian-Penkall (2016)	Sudeste do Paraná	Apagamento do -s final da DNP considerado uma marca do vernáculo rural. Destaque para sentenças no passado. Mais comum na oralidade e em contextos informais, apesar de também ter registros na escrita. Característica da mesorregião do PR.	Indícios de que também ocorre em zonas rurais do estado de SP. Destaque para contextos de sentenças no pretérito como diferenciação temporal.
Pereira (2021)	Regiões rurais eslavo-brasileiras no sudeste do Paraná	Maior ocorrência de dados gerais e da forma não canônica. Destaque para apagamento do -s final da DNP e tempo verbal pretérito. Maior frequência na fala de indivíduos menos escolarizados com maior nível de ruralidade.	Baixa produtividade de dados gerais e da forma não canônica. Variável <i>sexo/gênero</i> com aparente influência. Destaque para menor escolaridade.
Silva (2023)	Tefé - AM	Menor produtividade da variante [ẽ]. Maior ocorrência na fala de mulheres e de menos escolarizados. Uso categórico da variante [ẽ] em contexto de pretérito e de apagamento do -s final.	Baixa produtividade da variante [ẽ]. Maior ocorrência na fala dos homens e de menos escolarizados. Uso categórico da variante [ẽ] em contexto de pretérito e de apagamento do -s final.

Fonte: elaboração própria

4 CONSIDERAÇÕES FINAIS

Baseada teórico-metodologicamente na Teoria da Variação e Mudança Linguística (Labov 2008; Weinreich; Labov; Herzog, 2006), esta pesquisa investigou tanto variáveis linguísticas quanto variáveis extralinguísticas para a descrição da variação morfofonológica na vogal temática precedente à DNP /mo(s)/ em verbos regulares de primeira conjugação no presente e no pretérito perfeito do modo indicativo. Assim sendo, as variáveis analisadas foram: (i) *sexo/gênero*; (ii) *faixa etária*; (iii) *escolaridade*; (iv) *tempo verbal*; (v) *apagamento do -s final da DNP /mo(s)/*; (vi) *assertividade da sentença*; (vii) *realização do sujeito*; e (viii) *presença de elemento temporal*.

Esta pesquisa também tomou como base o estudo da variação morfofonológica da vogal temática em outras variedades do PB a partir de trabalhos já realizados por outros pesquisadores, tendo assim o objetivo de comparar as características do processo na variedade do São José do Rio Preto com outras variedades para identificar semelhanças ou divergências entre elas.

A partir dos resultados desta pesquisa e em comparação aos resultados obtidos em outras variedades do PB, foi possível observar que a troca da vogal temática [ã] ~ [ẽ] em verbos regulares de 1ª conjugação no presente e no pretérito perfeito do modo indicativo pode ser considerada uma variação morfofonológica, pois não envolve somente aspectos morfológicos, mas também fonético-fonológicos (Pereira; Margotti, 2018). De modo geral, essa variação ainda é pouco estudada no PB, apesar de já ter registros de sua identificação há muitos anos (Amaral, 2020; Costa, 1990; Bortoni-Ricardo, 2011) e de demonstrar ter suas origens nas variedades mais arcaicas do PE que influenciam nas variedades rurais do PB. Por conta disso, esse processo tem sido descrito como uma marca da zona rural e característico do falar “caipira”, o que pode ser um indício de que essa variação carrega marcas de estigma social.

Com o resultado do levantamento dos dados nos inquéritos do banco de dados Iboruna, foi possível observar que, assim como na variedade de Florianópolis (Pereira, 2014), o uso de [ẽ] no lugar de [ã] possui baixa produtividade, apresentando somente 6 aplicações da variante [ẽ] de um total de 642 ocorrências gerais de verbos no presente e no pretérito perfeito do modo indicativo (0,9%). Tais ocorrências foram extraídas de inquéritos de 3 informantes de um total de 151 que compõem o banco de dados. Em decorrência dessa baixa porcentagem obtida, o programa estatístico indicou nocautes nas variáveis: (i) *escolaridade*; (ii) *faixa etária*; (iii) *tempo verbal*; (iv) *presença do -s final*

na DNP /mo(s)/; e (v) *assertividade da sentença*. Em decorrência também da quantidade de dados, não foi possível realizar as rodadas informativas de PR para os fatores dessas variáveis.

Por consequência, este trabalho deixa algumas lacunas quanto à descrição do fenômeno em relação a essas variáveis, principalmente em se tratando das variáveis linguísticas, pois estas são compostas de dois fatores, o que inviabiliza o uso da estratégia de realizar amálgamas para solucionar os nocautes e dar continuidade às rodadas estatísticas. Em contrapartida, com a execução desta pesquisa, foi possível atingir o objetivo de descrever as variáveis linguísticas e extralinguísticas selecionadas para a investigação e comparar o fenômeno nesta variedade com outras, principalmente com as faladas na região Sul do país.

Ainda por conta do número reduzido de dados, não foi possível fazer afirmações categóricas acerca das hipóteses iniciais, apenas é possível observar algumas tendências sobre o comportamento do fenômeno variável dentro da variedade de São José do Rio Preto. Algumas dessas tendências, inclusive, também puderam ser observadas ao serem comparados os resultados com os de outras variedades do PB.

Em relação à maioria das hipóteses iniciais, foi possível identificar indícios de sua confirmação. Os resultados obtidos para a variável *sexo/gênero*, que demonstraram maior porcentagem de uso da variante [ê] no repertório linguístico dos homens; e os resultados obtidos para a variável *escolaridade*, que indicaram um maior uso dessa mesma variante na fala dos indivíduos do 1º Ciclo do EF, podem ser um indicativo de confirmação da hipótese de se tratar de um fenômeno com indícios de estigma social. Esse indicativo é ainda corroborado pela descrição que as pesquisas em outras variedades fazem desse fenômeno, definindo-o como característico de comunidades rurais. No entanto, a baixa produtividade dos dados e sua distribuição entre os fatores da variável *faixa etária* não possibilitaram averiguar a hipótese inicial de se tratar de uma variação estável.

Já no que se refere às hipóteses iniciais relacionadas às variáveis linguísticas, era esperado que a variável *assertividade da sentença* não seria de muita relevância para a produção do fenômeno, assim como não foi para a variedade de Florianópolis (Pereira, 2014), porém a distribuição das ocorrências gerais dentre os fatores, a saber, o número maior de sentenças positivas do que negativas, não possibilita fazer essa afirmação. Ainda sobre as hipóteses iniciais acerca das variáveis linguísticas, também era esperado que os fatores *-s final da DNP /mo(s)/ ausente*, *tempo verbal pretérito* e *ausência de*

elemento temporal pudessem exercer algum tipo de força para a produção da variante não canônica [ẽ]. Ao final, com os resultados, verificou-se que o tempo verbal pretérito e o apagamento do -s final da DNP podem realmente estar influenciando a produção da variante [ẽ] em São José do Rio Preto.

Além da semelhança de baixa produtividade de dados, a variedade do noroeste do estado de São Paulo também se assemelha à variedade de Florianópolis (Pereira, 2014) ao identificar as mesmas tendências no quesito da forma não canônica estar sendo propiciada por níveis de escolaridade mais baixos, pelo repertório masculino, pelo pretérito perfeito e pelo apagamento do -s final da DNP /mo(s)/. Outra semelhança encontrada entre ambas as variedades é o aparente uso preferencial da forma *a gente* pelos indivíduos dessas comunidades.

Comparando este trabalho com outros, realizados em outras variedades do PB (Pereira; Margotti, 2018; Pereira; Lehmkuhl-Coelho; Loregian-Penkal, 2016; Zilles; Maya; Silva, 2000), pode-se perceber semelhanças, visto que apresentam baixa produtividade dos dados e os resultados dos fatores linguísticos aparentam ter similaridades para a regra variável. Dessa forma, a variação morfofonológica objeto de estudo desta pesquisa parece ter características semelhantes às de outras variedades do PB já investigadas, aproximando a variedade do noroeste paulista a essas variedades no que se refere aos padrões dessa variação. A partir dessas comparações, a variação morfofonológica da vogal temática, mais especificamente o uso da variante não canônica, pode ser definida como característica do falar de comunidades rurais e motivada por escolaridade mais baixa, por uso no tempo verbal pretérito e pelo apagamento do -s final da DNP /mo(s)/.

Ao final desta pesquisa, é possível refletir que a variação de [ẽ] ~ [ẽ] realmente parece estar relacionada a uma fala mais rural, e o uso da forma não canônica pode estar ligado ao repertório e perfil individual do informante. Além disso, como demonstrado nos resultados do estudo dessa variação em outras variedades do PB, a aplicação da regra variável investigada parece estar atrelada a escolaridades mais baixas.

REFERÊNCIAS

- AMARAL, A. **O dialeto caipira**. 1. ed. São Paulo: Parábola, 2020.
- BATTISTI, E. Fonologia. *In*: SCHWINDT, L. C. (Org.). **Manual de Linguística**: Fonologia, Morfologia e Sintaxe. Petrópolis, RJ: Vozes, 2014, p. 27-108.
- BORTONI-RICARDO, S. M. **Do campo para a cidade**: estudo sociolinguístico de migração e redes sociais. São Paulo: Parábola, 2011.
- BRESCANCINI, C. R. A análise de regra variável e o programa Varbrul 2S. *In*: BISOL, L.; BRESCANCINI, C. (Org.). **Fonologia e variação**: recortes do português brasileiro. 1. ed. Porto Alegre: EDIPUCRS, 2002, p. 13-49.
- BUGAI, T. M. C. **A apricação da regra variável de rotacismo**: um exercício de análise na fala do interior paulista. 2021. 56 f. Trabalho de Conclusão de Curso (Licenciatura em Letras) – Universidade Estadual de Ponta Grossa, Ponta Grossa, 2021.
- BUGAI, T. M. C. **O rotacismo na fala do noroeste paulista**: uma análise sociolinguística. 2024. 95f. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Ponta Grossa, Ponta Grossa, 2024.
- CÂMARA JÚNIOR, J. M. **Estrutura da língua portuguesa**. 47. ed. Petrópolis, RJ: Vozes, 2015.
- CARMO, M. C. **As vogais médias pretônicas dos verbos na fala culta do interior paulista**. 2009. 119 f. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista, São José do Rio Preto, 2009.
- CARMO, M. C. **As vogais médias pretônicas na variedade do interior paulista**. 2013. 249 f. Tese (Doutorado em Estudos Linguísticos) – Universidade Estadual Paulista, São José do Rio Preto, 2013.
- CARMO, M. C. As vogais médias pretônicas no noroeste paulista: comparação com outras variedades do Português Brasileiro. **Estudos linguísticos** (São Paulo. 1978), v. 43, p. 33-47, 2014.
- CARMO, M. C. Variação linguística das vogais médias pretônicas em contexto medial no noroeste paulista. **UniLetras**, v. 40, p. 221-239, 2018.
- CARMO, M. C. Alçamento vocálico das vogais médias pretônicas iniciais na variedade do noroeste paulista. **Estudos linguísticos** (São Paulo. 1978), v. 48, p. 800-821, 2019.
- CARMO, M. C.; CARLOS, V. G. Alçamento vocálico sem motivação aparente: as vogais médias pretônicas no noroeste do estado de São Paulo. **Signum** [Londrina]: Estudos de Linguagem, v. 22, p. 114-144, 2019.
- CARMO, M. C.; MACHADO, T. F. Monotongação variável do ditongo oral decrescente [ej] no noroeste paulista. **Working Papers em Linguística**, v. 24, p. 160-184, 2023.

CARMO, M. C.; TABORDA, I. R. Apagamento de /R/ em coda silábica na variedade do interior paulista. **Letras Escreve**. Macapá, v. 9, n. 3, 2. sem. p. 39-51, 2019.

CARMO, M. C.; TENANI, L. E. As vogais médias pretônicas na variedade do noroeste paulista: uma análise sociolinguística. **Alfa: Revista de Linguística** (UNESP. Online), v. 57, p. 607-637, 2013.

COSTA, I. B. **O verbo na fala de camponeses**: um estudo de variação. 1990. 223f. Tese (Doutorado em Linguística) – Departamento de Linguística do Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1990.

CRISTÓFARO SILVA, T. **Dicionário de Fonética e Fonologia**. São Paulo: Contexto, 2011.

DIAS, J. A.; CARMO, M. C. A metátese na variedade do interior paulista. **Muitas Vozes**, v. 10, p. 1-23, 2021.

DUARTE, M. E. L. **A perda do princípio "Evite Pronome" no Português Brasileiro**. 1995. 151f. Tese (Doutorado em Linguística) - Universidade Estadual De Campinas, Instituto de Estudos da Linguagem, Campinas, SP.

FERREIRA, J. S. **O apagamento do /d/ em morfema de gerúndio no dialeto de São José do Rio Preto**. 145 f. 2010. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista, São José do Rio Preto, 2010.

FOEGER, C. C. **A primeira pessoa do plural no português falado em Santa Leopoldina/ES**. 2014. p. 166. Dissertação (Mestrado em Estudos Linguísticos) - Universidade Federal do Espírito Santo, Vitória, 2014.

FREITAG, R. M. K. Sociolinguística no/do Brasil. **Cadernos de Estudos Linguísticos**, Campinas, SP, v. 58, n. 3, p. 445–460, 2016.

GONÇALVES, C. A. **Morfologia**. São Paulo: Parábola, 2019. 165p.

GONÇALVES, S. C. L. Projeto ALIP (Amostra Linguística do Interior Paulista) e banco de dados Iboruna: 10 anos de contribuição do português brasileiro. **Revista Estudos linguísticos**, 2019, p. 276-297.

GONÇALVES, S. C. L. **O português falado na região de São José do Rio Preto**: constituição de um banco de dados anotado para o seu estudo. Relatório científico final apresentado à FAPESP. 2005. Disponível em: <http://www.alip.ibilce.unesp.br>. Acesso em: 29 abr. 2025.

GONÇALVES, S. C. L. **Banco de dados Iboruna**: amostras eletrônicas do português falado no interior paulista. 2007. Disponível em: <https://www.alip.ibilce.unesp.br/>. Acesso em: 29 abr. 2025 [2007].

GONÇALVES, S. C. L. **O português falado na região de São José do Rio Preto**: constituição de um banco de dados anotado para o seu estudo. Relatório científico parcial III apresentado à FAPESP. 2007. Disponível em: <http://www.alip.ibilce.unesp.br>. Acesso

em: 29 abr. 2025.

GUY, G.; ZILLES, A. **Sociolinguística Quantitativa**: instrumental de análise. São Paulo: Parábola, 2007. 239 p.

HEINEN, A. P.; LOREGIAN-PENKAL, L. **Políticas Linguísticas E A Língua Talian**. Uniletras, [S.L.], v. 46, p. 1-19, 2024. Universidade Estadual de Ponta Grossa (UEPG). <http://dx.doi.org/10.5212/uniletras.v.46.23116.2024>.

HORA, D.; MATZENAUER, C. (Org.). **Linguagem**: Variação e estrutura da língua. Campinas: Pontes, 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Censo Demográfico 2022**: população e domicílios - primeiros resultados. Rio de Janeiro: IBGE, 2023.

KOCH, W.; ALTENHOFEN, C. V.; KLASSMANN, M. (Orgs.). **Atlas Lingüístico-Etnográfico da Região Sul do Brasil (ALERS)**: Introdução, Cartas fonéticas e morfossintáticas. 2. ed. Porto Alegre: Editora da UFRGS; Florianópolis: Editora da UFSC, 2011. 512p.

LABOV, W. **Padrões sociolinguísticos**. Tradução de Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008.

LOREGIAN-PENKAL, L.; BALTHAZAR, L. L. Contato Linguístico Português Brasileiro - Talian Em Santa Felicidade (Curitiba) E Colombo, Paraná. Web Revista Sociodiaeto, [S.L.], v. 11, n. 33, p. 1-39, 20 maio 2021. State University of Mato Grosso do Sul. <http://dx.doi.org/10.48211/sociodiaeto.v11i33.337>.

MARTELOTTA, M. E. **Manual de linguística**. Linguística-Conceitos básicos. São Paulo: Contexto, 2008.

MONTEIRO, J. L. **Morfologia Portuguesa**. 4. ed. Campinas: Pontes, 2002.

NARO, A. J. O Dinamismo das línguas. In: MOLLICA, M. C. *et al.* (org.). **Introdução à Sociolinguística**: o tratamento da variação. 4. ed. São Paulo: Contexto, 2013, p. 43-50.

ORTALE, F. L. **A formação de uma professora de italiano como língua de herança: o Pós-Método como caminho para uma prática docente de autoria**. 2016. Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2016. Tese de Livre-Docência.

OUSHIRO, L. Tratamento de dados com o R para análises sociolinguísticas. In: FREITAG, R. M. K. (Org.). **Metodologia de Coleta e Manipulação de Dados em Sociolinguística**. São Paulo: Editora Edgard Blücher, 2014, p. 134-177.

PAIVA, M. C.; DUARTE, M. E. L. Mudança linguística: observações no tempo real. In: MOLLICA, M. C. (org.). **Introdução à sociolinguística**: o tratamento da variação. 4. ed. São Paulo: Contexto, 2013. Cap. 16. p. 179-191.

PAVEZI, V. C. **A haplogia na variedade paulista**. 2006. 119 f. Dissertação (Mestrado em Estudos Linguísticos). Universidade Estadual Paulista, São José do Rio Preto, 2006.

PEREIRA, I. Cuidamo(s) e cuidemo(s): a variação morfêmica na P4 em verbos regulares de 1ª conjugação. **Work. Pap. Linguíst.**, 15(2): 49-71, Florianópolis, ago./dez., 2014.

PEREIRA, I. **O pirogue, nós aprendimo da mãe” e “agora nós mudemo o borscht”**: variação morfofonológica em comunidades rurais eslavo-brasileiras no sudeste do paraná. 2021. 466 f. Tese (Doutorado em Linguística). Universidade Federal de Santa Catarina, Florianópolis, 2021.

PEREIRA, I.; AGOSTINHO, A. L. "Analisemo" E "Descrevimo": um caso de alçamento vocálico de caráter morfofonológico. **Organon**, Porto Alegre, v. 39, n. 78, p. 1-26, jul/dez. 2024.

PEREIRA, I.; LEHMKUHL-COELHO, I.; LOREGIAN-PENKAL, L. Variação na concordância verbal de nós no presente e pretérito perfeito em verbos regulares de 1ª e 2ª conjugação: produtiva no sudeste paranaense? **Signótica**, Goiânia, 2016.

PEREIRA, I.; MARGOTTI, F. W. Sobre onde nós fiquemo: mapeamento diatópico de um traço linguístico rural brasileiro. **Web-Revista SOCIODIALETO**, v. 8, n. 24, p. 221-250, mar. 2018. Disponível em <https://periodicosonline.uems.br/index.php/sociodialeto/article/view/7817>. Acesso em: 09. Set. 2025

PRODELIK, A. I. G.; CARMO, M. C. Ditongação diante de /S/ no interior paulista: análise sociolinguística. In: SANTOS, G. L. I.; OLIVEIRA, J. K.; LEGROSKI, M. C. (Org.). **Ensino, pesquisa e formação: a linguística no curso de Letras da UEPG**. 1. ed. Curitiba: PEDEX, 2024, p. 85-115.

RAMOS, A. P. **Descrição das vogais postônicas não-finais na variedade do noroeste Paulista**. 177 f. 2009. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista, São José do Rio Preto, 2009.

REZENDE, B. S. **“Cheguemos” até aqui, mas e agora? Análise linguística da alternância morfêmica em primeira pessoa do plural**. Relatório científico final apresentado à Fundação Araucária. Ponta Grossa. 2021. 21 p.

REZENDE, B. S.; CARMO, M. C. Análise sociolinguística da monotongação de [ow] na variedade do interior paulista. In: SANTOS, G. L. I.; LEGROSKI, M. (Org.). **Letras graduadas: trabalhos dos alunos do curso de Letras da Universidade Estadual de Ponta Grossa**. 1. ed. Ponta Grossa: 2023, p. 98-135.

REZENDE, B. S.; CARMO, M. C.; FUJIHARA, A. K. Alternância morfofonológica em P4 de verbos regulares de primeira conjugação no noroeste paulista. **Revista Estudios Lingüísticos y Literarios**. 2024.

RUBIO, C. F.; GONÇALVES, S. C. L. A fala do interior paulista no cenário da sociolinguística brasileira: panorama da concordância verbal e da alternância pronominal. **Alfa: Revista de linguística** (São José do Rio Preto), [s.l.], v. 56, n. 3, p. 1003- 1034,

2012.

SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. **Goldvarb X**: a Variable Rule Application for Macintosh and Windows. 2005. Department of Linguistics, University of Toronto. Disponível em: individual.utoronto.ca/tagliamonte/goldvarb.html. Acesso em: 29 abr. 2025.

SAUSSURE, F. **Curso de linguística geral**. Apresentação de Carlos Faraco. Tradução, notas e posfácio de Marcos Bagno. São Paulo: Parábola, 2021.

SCHWINDT, L. C. S.; *et al.* A influência da variável escolaridade em fenômenos fonológicos variáveis: efeitos retroalimentadores da escrita. **Revista Virtual de Estudos da Linguagem – ReVEL**. v. 5, n. 9, ago. 2007. ISSN 1678-8931 [www.revel.inf.br].

SILVA, L. M. T. **“Pronto, professora, demoremo, mas chegemo”**: Um estudo Sociolinguístico sobre a variação morfofonológica em verbos da 1ª e 2ª conjugação na fala de Tefé- Amazonas. 2023. 138 f. Dissertação (Mestrado) - Curso de Mestrado em Letras, Universidade Federal do Amazona, Manaus, 2023.

SILVEIRA, A. A. M. **As vogais pretônicas na fala culta do noroeste paulista**. 143f. 2008. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Estadual Paulista, São José do Rio Preto, 2008.

SVOBODOVÁ, P. **Variação diatópica na flexão verbal da 1.ª pessoa plural**. Études romanes de Brno. 38, 2017, p. 185-196.

TAGLIAMONTE, S. A. **Analysing Sociolinguistic Variation**. New York: Cambridge University Press, 2006.

TARALLO, F. **A pesquisa sociolinguística**. 7. ed. São Paulo, Ática, 2003.

TRUDGILL, P. **Sociolinguistics**: an introduction to language and society. 4. ed. London, Penguin Books, 2000.

VASCONCELOS, J. L. de. **Esquisse d’une dialectologie portugaise**. Lisboa: Centro de Estudos Filológicos, 1970.

WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**. São Paulo: Parábola, 2006.

ZILLES, A. M. S.; MAYA, L. Z.; DA SILVA, K. Q. **A concordância verbal com a p4 em Panambi e Porto Alegre**, RS. *Organon*, v. 14, n. 28-29, 2000. Disponível em: <https://seer.ufrgs.br/index.php/organon/article/view/30205>. Acesso em: 9 set. 2025.

ZILLES, A. M.; BATISTA, H. H. A concordância verbal na primeira pessoa do plural na fala culta de Porto Alegre. *In: Variação, mudança e contato linguístico no português da região sul*. Pelotas: EDUCAT, p. 100-124, 2006.

APÊNDICE A – LISTAGEM DOS VOCÁBULOS LEVANTADOS

Neste apêndice, apresentam-se os vocábulos com ocorrências gerais relativas aos verbos regulares de primeira conjugação conjugados em primeira pessoa do plural no presente e no pretérito perfeito do modo indicativo coletados no corpus de pesquisa. Também constam as ocorrências com aplicação.

Vocábulo	Ocorrências gerais	Ocorrências com aplicação
Acabamos	9	0
Acampamos	1	0
Achamos	5	0
Acrescentamos	1	0
Adentramos	2	0
Adoramos	1	0
Aguardamos	1	0
Ajudamos	1	0
Ajustamos	1	0
Almoçamos	4	0
Alugamos	3	0
Analizamos	1	0
Andamos	9	1
Apanhamos	1	0
Apresentamos	1	0
Aproveitamos	2	0
Arrancamos	2	0
Arranjamos	1	0
Arrumamos	1	0
Atravessamos	7	0
Aumentamos	4	0
Batizamos	1	0
Brigamos	4	0
Brincamos	4	0
Buscamos	2	0
Cantamos	1	0
Carregamos	1	0
Casamos	28	0
Chamamos	3	0
Chegamos	70	2
Cimentamos	1	0
Colocamos	8	0
Combinamos	1	0
Começamos	46	0
Completamos	1	0
Compramos	15	0
Contamos	2	0

Continuamos	3	0
Conversamos	6	0
Cortamos	1	0
Criamos	2	0
Debitamos	1	0
Deixamos	1	0
Descarregamos	1	0
Desenhamos	1	0
Desocupamos	1	0
Dobramos	1	0
Embarcamos	1	0
Encontramos	9	0
Enfiamos	1	0
Engrenamos	1	0
Entramos	10	0
Escutamos	2	0
Esperamos	3	0
Estucamos	1	0
Falamos	10	0
Fechamos	2	0
Ficamos	88	1
Ganhamos	8	0
Gostamos	2	0
Guardamos	1	0
Imaginamos	1	0
Importamos	1	0
Iniciamos	1	0
Jantamos	2	0
Jogamos	3	0
Largamos	11	0
Levantamos	1	1
Ligamos	3	0
Mandamos	8	0
Marcamos	5	0
Matamos	1	0
Misturamos	1	0
Montamos	5	0
Moramos	7	0
Mudamos	10	0
Muramos	1	0
Nadamos	2	0
Namoramos	10	0
Olhamos	2	0
Organizamos	1	0
Pagamos	3	0
Paramos	6	0
Passamos	29	0
Passeamos	1	0

Pegamos	28	1
Perguntamos	1	0
Pintamos	2	0
Plantamos	4	0
Precisamos	6	0
Procuramos	2	0
Pulamos	2	0
Quebramos	2	0
Queimamos	1	0
Rasgamos	1	0
Reformamos	2	0
Separamos	4	0
Tentamos	3	0
Terminamos	7	0
Tiramos	9	0
Tocamos	2	0
Tomamos	4	0
Tornamos	1	0
Trabalhamos	6	0
Tratamos	1	0
Trocamos	5	0
Usamos	2	0
Viajamos	2	0
Voltamos	27	0

Fonte: Elaboração própria