

UNIVERSIDADE ESTADUAL DE PONTA GROSSA
SETOR DE CIÊNCIAS HUMANAS, LETRAS E ARTES
DEPARTAMENTO DE ESTUDOS DA LINGUAGEM
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS DA LINGUAGEM

TAYNÁ MARIA COELHO BUGAI

O ROTACISMO NA FALA DO NOROESTE PAULISTA:
UMA ANÁLISE SOCIOLINGUÍSTICA

PONTA GROSSA – PR

2024

TAYNÁ MARIA COELHO BUGAI

O ROTACISMO NA FALA DO NOROESTE PAULISTA:
UMA ANÁLISE SOCIOLINGUÍSTICA

Dissertação de Mestrado apresentada como requisito parcial para obtenção do título de Mestre em Estudos da Linguagem, pela Universidade Estadual de Ponta Grossa, junto ao Programa de Pós-Graduação em Estudos da Linguagem. Área de concentração: Linguagem, Identidade e Subjetividade.

Orientadora: Prof.^a Dr.^a Márcia Cristina do Carmo
Coorientadora: Prof.^a Dr.^a Luciane Trennephol da Costa

PONTA GROSSA – PR

2024

B931 Bugai, Tayná Maria Coelho
O rotacismo na fala do noroeste paulista: uma análise sociolinguística /
Tayná Maria Coelho Bugai. Ponta Grossa, 2024.
94 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.
Coorientadora: Profa. Dra. Luciane Trennephol da Costa.

1. Sociolinguística. 2. Variação fonético-fonológica. 3. Rotacismo. I. Carmo,
Márcia Cristina do. II. Costa, Luciane Trennephol da. III. Universidade Estadual
de Ponta Grossa. Linguagem, Identidade e Subjetividade. IV.T.

CDD: 808

TAYNÁ MARIA COELHO BUGAI

O ROTACISMO NA FALA DO NOROESTE PAULISTA:
UMA ANÁLISE SOCIOLINGUÍSTICA

Dissertação apresentada para obtenção do título de Mestre em Estudos da
Linguagem na Universidade Estadual de Ponta Grossa, Área de concentração
em Linguagem, Identidade e Subjetividade.

Ponta Grossa, 30 de julho de 2024.

Documento assinado digitalmente
 **MARCIA CRISTINA DO CARMO**
Data: 07/08/2024 19:26:36-0300
Verifique em <https://validar.iti.gov.br>

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

Documento assinado digitalmente
 **LUCIANE TRENNEPHOL DA COSTA**
Data: 05/08/2024 09:46:10-0300
Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Luciane Trennephol da Costa - Coorientadora
Doutora em Letras
Universidade Estadual do Centro-Oeste

Documento assinado digitalmente
 **PASCOALINA BAILON DE OLIVEIRA SALEH**
Data: 05/08/2024 18:34:08-0300
Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Pascoalina Bailon de Oliveira Saleh
Doutora em Linguística
Universidade Estadual de Ponta Grossa

Documento assinado digitalmente
 **SUSIELE MACHRY DA SILVA**
Data: 31/07/2024 14:55:30-0300
Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Susiele Machry da Silva
Doutora em Linguística Aplicada
Universidade Tecnológica Federal do Paraná

AGRADECIMENTOS

A Deus, pois sem Ele nada eu poderia fazer.

Aos meu pais, Luiz e Ângela e minha irmã Luziangela, por todo o suporte e apoio durante este período de pesquisa.

Às minhas amigas Jheniffer, Miliane, Thyelle, Djenifer, Alice, Suelen, Yohanna e Vitória Régia, que me ouviram, me aconselharam e me incentivaram nos momentos em que mais precisei.

À professora Márcia Cristina do Carmo, por todo o conhecimento compartilhado sobre sociolinguística, fonética e fonologia, enfim, por ter me guiado pelo caminho da pesquisa científica. Também agradeço pela orientação pacienciosa e compreensiva durante os momentos mais difíceis que enfrentei.

À professora Luciane Trennephol da Costa, pela coorientação e trocas de conhecimento sobre a variação linguística.

Às professoras Pascoalina Bailon de Oliveira Saleh e Susiele Machry da Silva, por terem aceitado fazer parte da banca de avaliação da minha dissertação e que, com seus comentários cuidadosos, contribuíram para a lapidação desta pesquisa.

Ao PPGEL/UEPG, incluindo todos os funcionários, que sempre foram compreensivos e estiveram dispostos a ajudar, o que foi fundamental para a finalização desta pesquisa.

E, por fim, à Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa concedida.

"A diversidade linguística é uma manifestação da riqueza cultural e humana que deve ser celebrada."

– Marcos Bagno

RESUMO

Este trabalho tem como objetivo investigar o fenômeno do rotacismo na fala de informantes domiciliados na região de São José do Rio Preto, noroeste do estado de São Paulo, Brasil. O rotacismo consiste na realização fonética de um som rótico em substituição a um som lateral, como exemplificado nos casos de *p[r]anta* e *sa[l]* (Silva, 2011). A pesquisa é fundamentada na Teoria da Variação e Mudança Linguística (Labov, 2008) e considera também descrições desse processo em outras variedades do Português Brasileiro, como São José do Norte (RS) (Costa, 2007), Rio de Janeiro (RJ) (Tem Tem, 2010) e oeste paranaense (Cristino, 2021). Para realizar a análise, foram utilizadas 16 entrevistas sociolinguísticas do banco de dados Iboruna (Gonçalves, 2024). Os informantes foram estratificados de acordo com dois sexos/gêneros (feminino; e masculino), quatro faixas etárias (16 a 25; 26 a 35; 36 a 55; e acima de 55 anos) e duas escolaridades (primeiro ciclo do Ensino Fundamental; e Ensino Superior), sendo essas as variáveis sociais consideradas para este trabalho. As variáveis linguísticas investigadas incluem: (i) posição do segmento na sílaba; (ii) contexto precedente ao segmento; (iii) sonoridade precedente ao segmento; e (iv) presença de outro segmento líquido na palavra. A análise quantitativa dos dados foi realizada por meio do programa estatístico *Goldvarb X* (Sankoff; Tagliamonte; Smith, 2005). Os resultados encontrados indicam uma taxa de aplicação do rotacismo de 10,2% nos 1.915 dados analisados. Destaca-se a posição do segmento na sílaba como um fator que favorece a ocorrência do rotacismo, especialmente em casos de ataque complexo e coda medial, como exemplificado por palavras como *sim.p[r]es* e *a[l].guém*, respectivamente.

Palavras-chave: Sociolinguística; Variação fonético-fonológica; Rotacismo.

ABSTRACT

This study aims to investigate the phenomenon of rhotacism in the speech of informants domiciled in the region of São José do Rio Preto, northwest of São Paulo state, Brazil. Rhotacism refers to the substitution of the lateral liquid sound with the rhotic sound, e.g. p[r]anta ('plant') and sa[l] ('salt') (Silva, 2011). The research is based on the Linguistic Variation and Change Theory (Labov, 2008) and also considers descriptions of this process in other varieties of Brazilian Portuguese, e.g. São José do Norte (RS) (Costa, 2007), Rio de Janeiro (RJ) (Tem Tem, 2010), and western Paraná (Cristino, 2021). To conduct the analysis, 16 samples of spontaneous speech were used from the Iboruna database (Gonçalves, 2024). The informants were stratified based on two genders (female and male), four age groups (16-25; 26-35; 36-55; and above 55 years old), and two levels of education (first cycle of Elementary School and Higher Education), which are the social variables considered for this paper. The linguistic variables investigated include: (i) position of the segment in the syllable; (ii) preceding context of the segment; (iii) preceding segment voicing; and (iv) presence of another liquid segment in the word. The data analysis is conducted quantitatively using the statistical software Goldvarb X (Sankoff; Tagliamonte; Smith, 2005). Preliminary results indicate a rhotacism application rate of 10.2% out of the 1,915 analyzed occurrences. The position of the segment in the syllable is highlighted as a factor that favors the occurrence of rhotacism, particularly in cases of complex onset and medial coda, e.g. sim.p[r]es ('simple') and a[l].guém ('somebody'), respectively.

Keywords: Sociolinguistics; Phonological and phonetic variation; Rhotacism.

LISTA DE FIGURAS

FIGURA 1 - Mapa de localização de São José do Rio Preto – SP	37
FIGURA 2 - Mapa de São José do Rio Preto e municípios limítrofes	41

LISTA DE GRÁFICOS

GRÁFICO 1 - Ocorrências gerais de rotacismo	48
GRÁFICO 2 - Percentual de rotacismo em relação à faixa etária	56

LISTA DE QUADROS

QUADRO 1 - Relação do português padrão com o latim	24
QUADRO 2 - Variáveis extralinguísticas	43

LISTA DE TABELAS

TABELA 1 - Frequências de aplicação para regras linguísticas.....	17
TABELA 2 - Rendimento mensal para domicílios particulares permanentes em São José do Rio Preto - SP.....	39
TABELA 3 - Frequências de aplicação para regras linguísticas.....	49
TABELA 4 - Rotacismo em relação à posição na sílaba	51
TABELA 5 - Rotacismo em relação ao sexo/gênero	53
TABELA 6 - Rotacismo em relação à escolaridade.....	54
TABELA 7 - Rotacismo em relação à faixa etária	55
TABELA 8 - Rotacismo em relação ao contexto precedente	57
TABELA 9 - Ocorrências de rotacismo com vogal posterior precedente.....	58

SUMÁRIO

1 INTRODUÇÃO	12
2 FUNDAMENTAÇÃO TEÓRICA	16
2.1 TEORIA DA VARIAÇÃO E MUDANÇA LINGUÍSTICA.....	16
2.2 ROTACISMO.....	23
2.3 O ROTACISMO EM OUTRAS VARIEDADES DO PB	27
3 MATERIAL E MÉTODOS	36
3.1 COMUNIDADE LINGUÍSTICA	36
3.2 CÓRPUS DA PESQUISA.....	40
3.3 VARIÁVEIS ESTUDADAS.....	42
3.4 PROCEDIMENTOS METODOLÓGICOS.....	46
4 ANÁLISE E DISCUSSÃO DOS DADOS	48
4.1 POSIÇÃO NA SÍLABA.....	50
4.2 SEXO/GÊNERO	52
4.3 ESCOLARIDADE	54
4.4 FAIXA ETÁRIA	55
4.5 CONTEXTO PRECEDENTE	57
5 CONSIDERAÇÕES FINAIS	60
REFERÊNCIAS.....	62
APÊNDICE A – LISTA DE VOCÁBULOS COM CONTEXTO PARA A REALIZAÇÃO DO ROTACISMO	67
ANEXO A – RODADAS DO GOLDVARB X PARA A REGRA DO ROTACISMO NA VARIEDADE DO NOROESTE PAULISTA.....	76

1 INTRODUÇÃO

A presente dissertação¹ investiga o processo fonético-fonológico denominado *rotacismo* na variedade de São José do Rio Preto, região noroeste do estado de São Paulo. O rotacismo é definido como um “fenômeno fonológico relacionado com a realização fonética de um som rótico em substituição a um som lateral ou vice-versa” (Silva, 2011, p. 197), ou seja, consiste na comutação entre as consoantes líquidas // e /r/.

Dessa forma, o rotacismo é associado ao *lambdacismo*, um processo inverso a ele, que afeta a pronúncia do fonema [l] (Cox; Assad, 1999), pois ocorre a substituição de uma líquida vibrante /r/ por uma líquida lateral //, por exemplo, a palavra *cé[r]ebro* é realizada como *ce[l]ebro*.

Em vista a isso, cabe definir que, como recorte desta pesquisa, é estudada especificamente a substituição da líquida lateral // por uma líquida vibrante /r/, como em *simp[r]es* (para *simples*) e *reso[l]vido* (para *resolvido*).

Segundo Câmara Jr. (1976), apenas as laterais e vibrantes anteriores se configuram como um segundo elemento em um grupo consonântico pré-vocálico, dessa forma, possibilitam a ocorrência de contrastes, ou seja, a substituição de um som pelo outro, como, por exemplo, em *bloco* e *broco*. O autor também aponta que “um e outro contraste são muito precários nos dialetos sociais inferiores e mesmo num registro muito familiar... Nos grupos de líquida como segundo elemento consonântico, há nos dialetos sociais populares o rotacismo do //, que o muda em /r/” (Câmara Jr., 1976, p. 40-41).

A partir dessa colocação de Câmara Jr., é possível observar que esse fenômeno é considerado estigmatizado no Português Brasileiro (doravante, PB), o que também se confere a partir de outros estudos, como o de Bagno (2007), em que o autor destaca o rotacismo com um dos fenômenos que mais sofrem preconceito linguístico no PB. Devido a isso, surgiu o interesse de se analisar o rotacismo, fenômeno que até então não havia sido investigado sistematicamente sob o

¹ O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

arcabouço teórico-metodológico da Sociolinguística Quantitativa (Labov, 2008) na variedade do noroeste paulista.

Sendo assim, no Trabalho de Conclusão de Curso² de Bugai (2021), foi realizada uma pesquisa inédita sobre o rotacismo em São José do Rio Preto, por meio do banco de dados Iboruna (Gonçalves, 2024). A pesquisa contou com oito entrevistas sociolinguísticas analisadas e tratadas estatisticamente à luz da sociolinguística quantitativa (Labov, 2008), e que trouxeram resultados sobre o comportamento do fenômeno da variedade investigada, como, por exemplo, a maior ocorrência do rotacismo na fala de informantes pertencentes à faixa etária mais velha (superior a 55 anos) e ao sexo/gênero masculino, o que, segundo a literatura sociolinguística (Labov, 2008), demonstra sinais de estigma do fenômeno na variedade considerada.

Além disso, a pesquisa também trouxe resultados em relação às variáveis linguísticas, pois a realização do rotacismo foi favorecida no contexto de ataque complexo, como em *plan.ta*, e em contexto de coda medial, como em *bal.cão*. Porém, algumas lacunas ficaram a ser preenchidas, devido ao tempo para a exequibilidade daquele trabalho, o que fomentou o interesse por aprofundar o assunto nesta dissertação.

Desse modo, esta pesquisa avança em relação ao trabalho anterior quanto ao número de entrevistas, que dobram em quantidade. Também são investigadas as faixas etárias intermediárias (26 a 35 anos e 36 a 55 anos), que anteriormente não haviam sido consideradas, conduzindo assim uma pesquisa em tempo aparente³ a fim de verificar se o processo se trata de um caso de variação estável ou de mudança em progresso.

Assim, a presente investigação é realizada a partir de um corpus formado por dezesseis entrevistas sociolinguísticas provindas do banco de dados Iboruna (Gonçalves, 2024), e estratificadas quanto aos sexos/gêneros (feminino e masculino), faixas etárias (16 a 25 anos; 26 a 35 anos; 36 a 55 anos; e acima de 55 anos) e

² Desdobramento da pesquisa de Iniciação Científica (PIBIC) intitulada *Análise do rotacismo no interior paulista: qual o problema em falar prob[r]jema?*, desenvolvida na Universidade Estadual de Ponta Grossa (UEPG) no período de agosto de 2019 a julho de 2020, com bolsa UEPG, e premiada com o terceiro lugar na área de Linguística, Letras e Artes no XXIX Encontro Anual de Iniciação Científica (EAIC/UEPG).

³ Na pesquisa de tempo aparente, não há um acompanhamento longitudinal da fala do mesmo informante, mas há uma comparação na “frequência de uso de uma determinada variável distribuída por faixas etárias [...]” (Costa, 2006, p. 78).

escolaridades (primeiro ciclo do ensino fundamental e ensino superior). Os dados coletados foram tratados pelo programa estatístico *Goldvarb X* (Sankoff; Tagliamonte; Smith, 2005).

Como fundamentação teórico-metodológica, é utilizada a Teoria da Variação e Mudança Linguística ou Sociolinguística Quantitativa (Labov, 2008), assim, a análise é procedida sob o entendimento da língua como um sistema heterogêneo, dinâmico e variável, cujas características e usos são influenciados pelo contexto social e características sociais dos informantes, numa correlação com aspectos linguísticos e comportamentos linguísticos dos falantes. Também foram utilizados estudos prévios sobre o rotacismo a fim de compreender como esse fenômeno vem sendo descrito em diferentes variedades do PB (Costa, 2006; Tem Tem, 2010; Cristino, 2021).

Esta pesquisa tem como o objetivo investigar o uso do rotacismo por falantes da comunidade de São José do Rio Preto, localizada na região noroeste de São Paulo (SP), e a atuação de variáveis linguísticas e extralinguísticas para a aplicação do processo, de modo a obter indícios do *status* do fenômeno em relação à variação e à mudança linguística na variedade considerada, e, também, discutir o papel das variáveis sociais e como elas influenciam o uso do rotacismo dentro do grupo social analisado. Sendo assim, as quatro variáveis linguísticas investigadas são: (i) posição do segmento na sílaba; (ii) contexto precedente ao segmento; (iii) sonoridade precedente ao segmento; e (iv) outro segmento líquido na palavra. Ademais, foram consideradas três variáveis extralinguísticas: (i) sexo/gênero;⁴ (ii) escolaridade; e (iii) faixa etária.

Como hipóteses a serem testadas, têm-se as de que o rotacismo na variedade do noroeste paulista: (i) seja influenciado por fatores sociais, assim como observado em outras variedades do PB (Costa, 2006; Tem Tem 2010), ou seja, espera-se que o rotacismo tenha mais probabilidade de ocorrer em falantes do sexo/gênero masculino e em informantes com menos escolaridade; e (ii) seja motivado por fatores de natureza linguística; e, mais especificamente, seja favorecido pelo contexto de ataque

⁴ Embora sexo e gênero não sejam tomados como sinônimos, é adotada nesta pesquisa a nomenclatura *sexo/gênero* pelo fato de seguir a mesma terminologia utilizada pelo banco de dados Iboruna (Gonçalves, 2024).

complexo,⁵ assim como observado em pesquisas anteriores (Costa, 2006; Bugai, 2021).

Sendo assim, a presente dissertação está estruturada da seguinte forma: no capítulo 2, é apresentada a fundamentação teórica, que se baseia na Teoria da Variação e Mudança Linguística (Labov, 2008) e que apresenta, também, estudos prévios sobre o fenômeno em outras variedades do PB. No capítulo 3, expõem-se os materiais e métodos empregados na análise do fenômeno variável, assim como informações sobre a comunidade de fala investigada e sobre o banco de dados utilizado. Em seguida, o capítulo 4 apresenta os resultados gerados pelo programa estatístico e os fatores de relevância para a ocorrência do fenômeno investigado, bem como a análise e discussão desses dados. Por último, no capítulo 5, são apresentadas as considerações finais, seguidas pelas referências bibliográficas, pelo apêndice e pelo anexo.

⁵ Do inglês *onset*, é o início da sílaba. Sílabas com mais de um segmento no ataque são consideradas um “ataque complexo”. Apenas as consoantes líquidas /l/ e /r/ podem ser um segundo segmento dentro de um ataque, por exemplo, a palavra blu.sa (Collischonn, 2006).

2 FUNDAMENTAÇÃO TEÓRICA

Conforme mencionado anteriormente, esta dissertação de Mestrado investiga o fenômeno fonético-fonológico denominado *rotacismo* na variedade do noroeste de São Paulo. Para esse fim, baseia-se no referencial da Teoria da Variação e Mudança Linguística (Labov, 2008).

Neste capítulo, são apresentados subtemas que contribuem para a compreensão de alguns conceitos essenciais para a condução desta pesquisa. São eles: (i) *Teoria da Variação e Mudança Linguística* (seção 2.1), que inclui estudos de Weinreich; Labov; Herzog, (2006), Labov (2008), Tarallo (2003), Costa (2006), Paiva e Duarte (2015) e outros; (ii) *rotacismo* (seção 2.2), que detalha as características do processo investigado nesta pesquisa; e (iii) *o rotacismo em outras variedades do PB* (seção 2.3), que apresenta alguns exemplos de estudos prévios do tema em outras variedades do PB.

2.1 TEORIA DA VARIAÇÃO E MUDANÇA LINGUÍSTICA

Esta pesquisa utiliza como arcabouço teórico-metodológico a Teoria da Variação e Mudança Linguística (Labov, 2008), tendo como um de seus representantes o linguista William Labov. Nos livros *Fundamentos empíricos para uma teoria da mudança linguística* (Weinreich; Labov; Herzog, 2006) e *Padrões Sociolinguísticos* (Labov, 2008), os autores apresentam a língua como um sistema social e heterogêneo, rompendo com a concepção de língua homogênea seguida pelos linguistas da época. Ao considerar a língua como um sistema social e heterogêneo, os defensores dessa teoria abordam a variabilidade da língua, tendo como objeto de estudos padrões de comportamento linguístico observáveis dentro de uma comunidade de fala.

Um dos principais fundamentos desse método é o conceito de *variável linguística*. Uma *variável* pode ser definida como um conjunto de formas linguísticas que se realizam de diferentes modos, conforme a variedade investigada. Como explica Labov, “dois enunciados que se referem ao mesmo estado de coisas e que têm o mesmo valor de verdade” (Labov, 1978, p. 7, tradução nossa),⁶ assim, cada

⁶ “Two utterances that refer to the same state of affairs have the same truth-value” (Labov, 1978, p. 7).

uma dessas formas possíveis é chamada de *variante*. Essa questão sobre o *valor de verdade*⁷ já foi amplamente debatida, inclusive por Lavandera (1977), então cabe ressaltar que a expressão “estado de coisas” utilizada aqui por Labov (1978) se refere ao significado representacional, relativizando assim o sentido de significado. Com base nesses conceitos, torna-se possível descrever a variação sistemática da língua, por meio dos condicionadores, que podem ser tanto internos (linguísticos) quanto externos (sociais).

Labov (2003), em seu livro *Some Sociolinguistic Principles*, apresenta o conceito de *variação estruturada*, que se refere a padrões sistemáticos de variação na língua(gem) falada. Ele contrasta isso com variações acidentais de performance, que são flutuações momentâneas (fadiga, distração, mudanças temporárias na atenção) e não sistemáticas na fala de um indivíduo. A partir disso, ele dispõe um modelo, com base em porcentagens para avaliar se determinada regra aplicada é *categórica*, *semi-categórica* ou *variável*, como ilustrado na Tabela 1:

TABELA 1 - Frequências de aplicação para regras linguísticas

Frequência de Aplicação	Tipo de Regra
100%	Categórica
95-99%	Semi-categórica
5-95%	Variável

Fonte: Elaboração própria, com base em Labov (2003, p. 243)

Assim, quando uma regra apresenta uma frequência de 100% de aplicação, ela é considerada categórica, o que significa que, caso haja uma “violação” dessa regra, será perceptível que ela não ocorre na fala natural e pode gerar confusão ou dificuldade de interpretação pelos falantes, pois “tais regras automáticas existem em todas as formas de comportamento social, mas são extremamente difíceis de detectar simplesmente porque nunca são violadas e nunca se pensa nelas” (Labov, 2003, p. 242, tradução nossa).⁸ Agora, quando uma regra tem uma porcentagem entre 95-99% de aplicação, ela é considerada semi-categórica, pois é de ocorrência rara e

⁷ As variantes linguísticas, teoricamente, deveriam ter o mesmo valor de verdade porque o significado comunicativo não está intrinsecamente ligado à forma linguística usada. No entanto, o preconceito linguístico pode distorcer essa percepção ao influenciar negativamente a credibilidade percebida de um falante com base na variedade linguística que ele ou ela utiliza.

⁸ “Such automatic rules exist in all forms of social behavior, but they are extremely hard to detect simply because they are never violated and one never thinks about them at all” (Labov, 2003, p. 242).

reportável, mas já é de conhecimento do falante que “as pessoas realmente cometem violações e que é possível interpretar tais violações. Existe um rótulo ou interpretação pronta que acompanha a quebra das regras” (Labov, 2003, p. 242, tradução nossa).⁹ Por fim, existem as regras variáveis, que são detectadas quando as suas porcentagens de aplicação são entre 5-95%. Essas regras são consideradas irreportáveis e que não existem por definição (norma padrão). Labov (2003, p. 241) aponta que, quando essa regra variável ocorre, podem acontecer casos de estigma em relação a essa forma, associando o falante a alguém que não conhece bem a sua língua, ou que não teve uma boa educação. Assim, essas regras variáveis, “apesar de não poderem ser violadas com qualquer pronúncia de uma palavra, [...] são uma parte importante da nossa competência linguística” (Labov, 2003, p. 242, tradução nossa).¹⁰

Visto como a frequência de uso de determinada regra pode definir como ela é percebida pelos falantes, é possível observar a importância da *Sociolinguística Variacionista* ou *Quantitativa*, essa corrente teórica, que, segundo Tarallo (2003), é assim denominada por apurar frequência de uso e conferir tratamento estatístico aos dados coletados e que surgiu com fortes oposições às concepções de língua do estruturalismo e do gerativismo.

Conforme Weinreich, Labov e Herzog (2006), a sociolinguística se opõe a duas dicotomias saussurianas: (i) *langue* x *parole*; e (ii) *diacronia* x *sincronia*. Nos estudos estruturalistas, o objeto de estudo é a *langue*, definida como “social em sua essência e independente do indivíduo; esse estudo é unicamente psíquico” (Saussure, 2006, p. 27). A partir disso, a *parole* é considerada secundária, sendo caracterizada como “a parte individual da linguagem, vale dizer, a fala, inclusive a fonação e é psico-física” (Saussure, 2006, p. 27). Essa dicotomia influenciou os estudos linguísticos, pois as pesquisas que tinham como base a visão estruturalista utilizavam amostras invariáveis da linguagem, uma vez que a variação era considerada da ordem da *parole*.

É importante salientar que, apesar de Saussure usar o termo “social”, este não é empregado com o fim de considerar os aspectos sociais no sentido tradicional de estudar a interação e a estrutura da sociedade. Em vez disso, Saussure utiliza “social”

⁹ “[...] people do indeed make violations, and that one can interpret such violations. There is a ready made label or interpretation which goes with the breaking of the rules” (Labov, 2003, p. 242).

¹⁰ “Despite the fact that they cannot be violated with any given pronunciation of a word, they are an important part of our linguistic competence” (Labov, 2003, p. 242).

para destacar a natureza coletiva e compartilhada da língua. Para ele, a língua é um produto social porque seu sistema de sinais e significados é coletivamente acordado e compreendido dentro de uma comunidade de falantes. Uma vez que tais estudiosos desconsideram questões externas, como explica Labov (2008, p. 217) em relação ao estruturalismo e gerativismo,

trabalham com um ou dois informantes em seus escritórios ou examinam seu próprio conhecimento da língua. Além disso, insistem em que as explicações dos fatos linguísticos sejam derivadas de outros fatos linguísticos, não de quaisquer dados “externos” sobre o comportamento social.

Sendo assim, a sociolinguística se opõe a essas teorias ao levar em conta a variação e os aspectos sociais da comunidade de fala.

A dicotomia *diacronia x sincronia* realiza uma divisão temporal na investigação linguística. Segundo Saussure (2006, p. 116),

a Linguística sincrônica se ocupará das relações lógicas e psicológicas que unem os termos coexistentes e que formam sistemas; tais como são percebidos pela consciência coletiva. A Linguística diacrônica estudará, ao contrário, as relações que unem termos sucessivos não percebidos por uma mesma consciência coletiva, e que se substituem uns aos outros sem formar sistema entre si.

Saussure (2006) explica que a sincronia é diferente da diacronia, porque a sincronia observa a relação de elementos simultâneos, e a diacronia, a substituição de uma forma por outra no tempo. Com essa divisão, só era possível observar a mudança linguística por meio de uma comparação entre duas sincronias, ou seja, quando a mudança já havia sido completada, no seu término. Entretanto, com o advento da sociolinguística (Labov, 2008), tornou-se possível observar a mudança em curso.

No que se refere à teoria gerativa, a oposição está relacionada à dicotomia *competência x desempenho*, sendo a competência o conhecimento abstrato das regras da língua, e o desempenho, a seleção e execução dessas regras (Weinreich; Labov; Herzog, 2006). No viés gerativo, o objeto de estudos é a competência linguística, ou seja, “comunidade de fala abstrata, homogênea, em que todo mundo fala igual e aprende a língua instantaneamente” (Labov, 2008, p. 218), abrangendo as intuições e julgamentos do falante acerca da língua, bem como a relação entre frases que apresentam o mesmo significado.

A partir disso, Labov (2008, p. 224-225) argumenta que linguistas adeptos da gramática gerativa trabalham com textos normativos editados, visto que é descartada uma grande quantidade de variações “problemáticas”, uma vez que não só deixam de lado o comportamento contextual e se baseiam, muitas vezes, em dados de apenas um informante, como também são influenciados pelas intuições do próprio pesquisador:

Pode se ver exemplos desse problema em qualquer encontro de linguistas, onde uma comunicação após a outra citará dados importantes como aceitáveis ou inaceitáveis sem que se obtenha o consenso da plateia. Isso não se deve à negligência ou à inabilidade linguística da parte dos autores: em sua sincera intenção de explorar a teoria linguística com base em suas intuições, eles inevitavelmente atingem o ponto em que os dados assumem tal forma.

Como consequência, a sociolinguística se apresentou como uma reação à ausência do componente social nas investigações assentadas em modelos formalistas. Segundo estudiosos e críticos da área de linguística, a obra *Padrões Sociolinguísticos* (*Sociolinguistic Patterns* – Labov, 2008) é considerada uma das obras mais influentes na área da sociolinguística. Nesse livro, o autor apresenta estudos sobre o inglês falado na ilha de Martha’s Vineyard, a estratificação social do inglês falado em Nova Iorque, dentre outras pesquisas relevantes para os estudos variacionistas.

Em *A Motivação Social de uma Mudança Sonora* (Labov, 2008), o autor apresenta uma análise de uma alteração fonética em ditongos /aj/ e /aw/, na ilha *Martha’s Vineyard*, localizada no estado norte-americano de Massachusetts. Essa ilha possuía quase seis mil habitantes em 1960 e foram feitas 69 entrevistas com falantes nativos, levando em consideração variáveis independentes como: região, etnia, faixa etária e profissão.

O autor estudou a centralização dos ditongos /aj/ e /aw/, pronunciados de formas diferentes do inglês padrão, por exemplo [ej], em *white* (‘branco’), ou [ew], como em *house* (‘casa’). Como resultado, Labov constatou que os fatores extralinguísticos tinham maior relevância para a ocorrência dos ditongos centralizados, corroborando a sua hipótese inicial, de que essa forma não-padrão de realização dos ditongos partiria dos falantes nativos como uma forma de tradição e de marcar a identidade dos moradores da ilha (Labov, 2008).

No estudo *A Estratificação do /r/ nas Lojas de Departamento de Nova York* (Labov, 2008), que investigou a presença ou ausência do /r/ em posição pós-vocálica, como nas palavras do inglês *car* ('carro'), *card* ('cartão'), *four* ('quatro'), *fourth* ('quarto(a)'), as técnicas usadas na pesquisa sobre centralização do ditongo foram "refinadas e aplicadas a uma situação muito mais complexa" (Labov, 2008, p. 62). Os informantes selecionados para essa pesquisa eram atendentes das lojas de departamento *Saks*, *Macy* e *S. Klein*. A escolha das lojas foi estratégica, representando três classes sociais: *Saks* – classe alta; *Macy's* – classe média; e *S. Klein* – classe baixa. A premissa era a de que um atendente de uma loja de classe social alta iria se adequar à linguagem utilizada por seus clientes.

De acordo com o autor, "a abordagem básica de isolamento das variáveis socialmente significativas e de correlação delas com os padrões das forças sociais gerais foi a mesma usada em *Martha's Vineyard*" (Labov, 2008, p. 62). Para a coleta dos dados, o pesquisador fazia perguntas cujas respostas eram *fourth floor* ('quarto andar') e pedia para que o informante repetisse a resposta com a finalidade de observar se havia um reforço ou uma correção da fala.

Como resultado, Labov (2008) confirmou a sua hipótese ao constatar que existia um fator social relacionado à pronúncia do /r/: *S. Klein*, loja voltada para um público de classe baixa, foi a que apresentou o maior índice de apagamento do /r/; na *Macy's*, voltada para um público médio, os resultados colocaram a loja em posição intermediária; e, por fim, na *Saks*, loja voltada para o público de classe alta, os funcionários foram os que menos apresentaram a supressão do /r/.

Após a apresentação desse estudo sobre a relação entre a realização do /r/ e as questões sociais, faz-se necessário abordar os conceitos de variantes *padrão/não-padrão*, *conservadora/inovadora* e *de prestígio/estigmatizada*. A variante padrão é considerada conservadora e prestigiada e tende a ser usada por uma comunidade de fala que apresenta um alto nível de escolaridade e prestígio social. As variantes classificadas como não-padrão podem ser consideradas inovadoras e estigmatizadas, visto que costumam estar associadas a um segmento social que vive à margem da sociedade e com pouco acesso à escolarização, como explica Tarallo (2003).

A noção de variantes padrão/não-padrão está correlacionada à concepção de comunidade de fala. Segundo Labov (2008, p. 188), "uma comunidade de fala não pode ser concebida como um grupo de falantes que usam todos as mesmas formas;

ela é mais bem definida como um grupo que compartilha as mesmas normas a respeito da língua”, ou seja, é uma comunidade que compartilha de atitudes e normas sociais diante de uma língua. Para o autor, os falantes da mesma comunidade não precisam falar da mesma maneira, pois eles apresentam a mesma série de avaliações sobre a fala.

Ao estudar a vida social da comunidade, consideram-se variáveis extralinguísticas, como *escolaridade*, *sexo/gênero* e *faixa etária*. Por meio da variável *escolaridade*, é possível investigar a influência do acesso ao ensino formal na fala dos informantes. O *sexo/gênero* é utilizado para verificar se há uma distinção no uso da língua entre homens e mulheres, pois, segundo Labov (2008), as mulheres, assim como indivíduos mais escolarizados em geral, tendem a evitar as formas estigmatizadas socialmente, aproximando-se de um falar mais prestigiado.

A variável *faixa etária* é produtiva para a análise do *status* da variação e mudança linguística. Para traçar um panorama de mudança linguística, são necessários dois tipos de estudos: em *tempo real* e *tempo aparente*, ambos são essenciais para observar uma mudança em curso ou uma variação estável. Conforme Paiva e Duarte (2015), em um estudo em *tempo aparente*, é realizado um recorte sincrônico de amostras de fala de diferentes gerações. Na pesquisa em tempo aparente, não há um acompanhamento longitudinal da fala do mesmo informante, mas há uma comparação na “frequência de uso de uma determinada variável distribuída por faixas etárias [...]” (Costa, 2006, p. 78). A presente pesquisa adota uma abordagem de estudo em tempo aparente.

Por sua vez, uma pesquisa em *tempo real* pode proceder de duas formas: (i) *longa duração*: em que é realizada uma investigação diacrônica de um fenômeno para verificar a sua trajetória até o momento atual; e (ii) *curta duração*, caracterizada por duas formas diferentes de análise: *painel*, em que são investigadas amostras de fala do mesmo informante em dois períodos de tempo; e *tendência*, em que há o confronto entre amostras distintas de fala da mesma comunidade, com perfis “estratificados com base nos mesmos parâmetros sociais, em dois momentos do tempo” (Paiva; Duarte, 2015, p. 188).

Para uma análise de mudança linguística, a maneira mais adequada, segundo Paiva e Duarte (2015), é a de conjugar as evidências obtidas por meio do estudo da mudança no tempo aparente com as evidências fornecidas pelo tempo real.

Agora avança-se para a próxima seção, na qual será abordado o processo fonético-fonológico do *rotacismo*, objeto central de estudo deste trabalho. Nessa seção, serão apresentadas informações detalhadas sobre o rotacismo, sua definição, características e relevância para a pesquisa em questão. Serão explorados aspectos fonéticos e fonológicos desse fenômeno, visando fornecer um embasamento teórico para a compreensão do seu funcionamento e das análises realizadas ao longo desta pesquisa.

2.2 ROTACISMO

O rotacismo, como já explicado inicialmente, pode ser definido como um processo fonético-fonológico “relacionado com a realização fonética de um som rótico em substituição a um som lateral” (Silva, 2011, p. 197). O processo inverso a ele é conhecido como *lambdacismo*, em que, segundo Cox e Assad (1999), ocorre um som lateral no lugar de um som rótico, como, por exemplo, *re.f[l]i.ge.rante* para *re.f[r]i.ge.ran.te*.

Câmara Jr. (1976, p. 40-41) aponta que “nos grupos de líquida como segundo elemento consonântico, há nos dialetos sociais populares o rotacismo do //, que o muda em /r/”, como *p//a.ca > p[r]a.ca*. Cabe ressaltar que o rotacismo no PB é propício de ocorrer em três contextos silábicos diferentes, sendo eles: coda final, como em *sa// > sa[ɹ]*¹¹; coda medial, como em *a//.fa.ce > a[ɹ].fa.ce*; e ataque complexo, como em *b//u.sa > b[r]u.sa* (Cristino; Busse, 2017).

Em relação às consoantes líquidas, o // e o /r/ se comutam e dão origem ao processo do rotacismo. Segundo Costa (2006, p. 12), os estudos desses segmentos são desafiadores às teorias linguísticas pelo fato de possuírem um “largo espectro fonético sem terem uma propriedade comum”. Porém, por partilharem um padrão silábico, e também serem suscetíveis a diversos fenômenos, como vocalização, apagamento, metátese e posteriorização de articulação, formam uma classe fonológica (Costa, 2006, p. 16-17).

Além disso, as consoantes líquidas apresentam uma alta variação alofônica nas línguas. Em Dickey (1997), foi realizado um estudo abrangendo um banco de

¹¹ O R em coda silábica é representado como um retroflexo nesta pesquisa, pelo motivo de que, por meio da análise de oitiva, foi observado que os informantes selecionados utilizaram esse rótico para esse contexto silábico.

dados com 78 línguas, que revelou que 56% dos segmentos róticos apresentam variação alofônica com outro tipo de rótico. Além disso, conforme Costa (2006), é comum encontrar casos em que um segmento rótico alterna com um segmento lateral.

Em algumas línguas, como o coreano, por exemplo, esses segmentos são alofones em distribuição complementar (Costa, 2006, p. 16). O som /r/ é utilizado no início ou no meio de uma sílaba, principalmente em posição de ataque silábico. Ele é produzido com um toque rápido da ponta da língua no alvéolo dental. Já o som /l/ é utilizado na posição de coda, ou seja, no final de uma sílaba. Ele é produzido com a elevação da ponta da língua em direção ao palato alveolar.

Já no PB, segundo Câmara Jr. (1976, p. 39-41), “as laterais e vibrantes só possuem ressonância oral, diferindo entre si pelo ruído de oclusão parcial (/l/, /l,/) e pelo de vibração (/r/ e /r’/),” e que só elas podem ocupar o segundo lugar em um grupo consonântico pós-vocálico, criando-se assim contrastes em palavras como *bloco* > *broco*, *atlas* > *atras*, *clave* > *crave*, *fluir* > *fruir* (Câmara Jr., 1976).

Freitag (2010) discorre que a “neutralização” distintiva dos segmentos líquidos está relacionada ao ponto de vista articulatorio, pois “ambos são realizados com a ponta da língua tocando os alvéolos dos dentes” (Freitag, 2010, p. 18). Todavia, no português, os sons /l/ e /r/ são fonemas distintos, o que significa que a troca de um por outro pode resultar em mudança de significado. Essa diferença fonêmica é evidenciada pelo princípio da comutação a pares mínimos idênticos ou análogos de palavras, como exemplificado pelas palavras *mal* e *mar* (Cox; Assad, 1999).

A tendência natural de mudança fonética pode ser observada na evolução das línguas românicas, que têm sua origem no latim vulgar, como ilustra o quadro abaixo, com base em Bagno (2007, p. 73):

QUADRO 1 - Relação do português padrão com o latim

Português padrão	Latim
<i>Brando</i>	<i>Blandu</i>
<i>Cravo</i>	<i>Clavu</i>
<i>Dobro</i>	<i>Duplu</i>
<i>Escravo</i>	<i>Sclavu</i>
<i>Fraco/Frouxo</i>	<i>Flaccu</i>
<i>Grude</i>	<i>Fluxugluten</i>
<i>Obrigar</i>	<i>Obligare</i>
<i>Praga</i>	<i>Plaga</i>

Fonte: Bagno (2007, p. 73)

Como pode se observar no quadro 1, nos dados em latim havia um // em contexto de ataque complexo e que se converteu em /r/ no português.

Além disso, também existem alguns registros da ocorrência do rotacismo no *Appendix Probi*, texto em que se compilam os mais frequentes vocábulos considerados inapropriados na fala latina vulgar, no século IV d.C., em comparação ao latim clássico, como nos seguintes exemplos: *flagellum non fragellum*, *terebra non telebra* e *glari non cracli*. Silva Neto (1956, p. 99) cita ocorrer uma “dissimilação de acordo com a conhecida fórmula $r - r > l - r$ ” e menciona exemplos como *albor > arbor* e *pelegrinus > peregrinus*.

Em vista a isso, é cabível afirmar que casos de rotacismo ocorrem desde as primeiras mudanças do latim para o português, levando em consideração exemplos de palavras como *blanco*, *clavo*, *flaccu*, atualmente realizadas categoricamente no PB contemporâneo como *branco*, *cravo* e *fraco*, respectivamente (Botelho; Leite, 2005, p. 6).

O rotacismo demonstra estar presente amplamente ao longo da história da língua portuguesa; outro exemplo de sua presença é na obra, escrita por Luís de Camões, *Os Lusíadas* (1572), em que são encontradas palavras como *frauta*, *frecha*, *ingrês*, *pranta*, *pruma*, *pubrica*, como demonstra o excerto citado em Bagno (2007, p. 217).

E não de agreste avena, ou frauta ruda” (canto I, verso 5) “Doenças, frechas, e trovões ardentes” (X, 46) “Era este inglês potente.” (VI, 47) “Nas ilhas de Maldiva nasce a pranta” (X,136) “Pruma no gorro, um pouco declinada” (II, 98) “Onde o profeta jaz, que a lei pubrica” (VIII, 34).

Sem a percepção diacrônica dos falantes da língua de que o fenômeno pode ser uma continuidade de uma tendência antiga no português, fenômenos como o rotacismo são avaliados de forma estigmatizada e tratados como um “erro”, pelo motivo da distância em relação à norma pré-estabelecida pelas gramáticas prescritivas (Bagno, 2007, p. 73-74). Desse modo, Freitag (2010, p. 19) explana que:

a neutralização do traço distintivo dos fonemas // e /r/ não se trata de um “erro”, mas sim efeito de um processo cíclico pelo qual passa a nossa língua: formas do passado (como por exemplo, do latim) podem se repetir em formas linguísticas atuais, como é o caso do rotacismo.

No livro *O dialeto Caipira* (1981), de Amadeu Amaral, o autor apresenta variáveis linguísticas consideradas pertencentes ao dialeto “caipira” no estado de São

Paulo. O rotacismo é ilustrado em coda silábica¹² quando a consoante “l em final de sílaba, muda-se em r” (Amaral, 1981, p. 52) e como exemplos encontram-se as palavras *quarquér, papér, mér, arma* para, respectivamente, *qualquer, papel, mel, alma*. Também é mencionado o rotacismo em ataque complexo (Amaral, 1981, p. 52), exemplificado nas palavras *craro, cumpreto, cramôr, frô(r)* em relação, respectivamente, a *claro, completo, clamor, flor*. Amaral (1981, p. 52) destaca que “esta troca é um dos vícios de pronúncia mais radicados no falar dos paulistas, sendo mesmo frequente entre muitos dos que se acham, por educação ou posição social, menos em contato com o povo rude”. Dessa forma, o rotacismo demonstra ser estigmatizado à época do autor para a variedade paulista, sendo associado ao falar de um povo “rude” e também considerado um “vício” de linguagem.

Ante isso, é demonstrada a importância de estudos sobre fenômenos fonético-fonológicos para a conscientização sobre o preconceito linguístico, pois o rotacismo é considerado um dos fenômenos mais estigmatizados no PB, “se um falante pronuncia “pranta” ao invés de “planta”, “framengo” em lugar de “flamengo”, é caracterizado como pertencente a camadas sociais desprestigiadas, marginalizadas, que não tem acesso à educação da escola (Freitag, 2010, p. 19).

Assim, como uma variante está associada a fatores socioeconômicos, sexos/gêneros, faixas etárias, escolaridade, o preconceito não se volta ao fenômeno do rotacismo, que demonstra ser inerente à língua e consistente desde as primeiras mudanças do latim, mas, sim, trata-se de um preconceito social, “pois o erro não está no que se fala e sim em quem fala” (Freitag, 2010, p. 19).

Por esses e outros motivos, já foram realizados estudos sobre o fenômeno do rotacismo, considerando diferentes variáveis independentes e variedades específicas do PB. Na próxima seção, serão apresentados exemplos desses estudos sobre o tema, de modo a enriquecer o contexto teórico-comparativo, permitindo uma compreensão mais ampla do fenômeno do rotacismo e suas características em diversas variedades do PB.

¹² A coda é o final de uma sílaba, por exemplo, *al.fa.ce* (coda medial) ou *pa.pel* (coda final). Seu nome provém do latim e tem o significado de *cauda* ou *rabo*. Os únicos segmentos permitidos em coda no PB são: os representados pelos arquifonemas /R/, /L/, /S/, /N/ e as semivogais (Collischonn, 2006).

2.3 O ROTACISMO EM OUTRAS VARIEDADES DO PB

Nesta seção, são apresentados alguns estudos¹³ sobre o fenômeno do rotacismo em diferentes variedades do PB, sendo eles: (i) Costa (2006), para a variedade de São José do Norte – RS; (ii) Tem Tem (2010), para a variedade do Rio de Janeiro – RJ; (iii) Cristino (2021), para a variedade do oeste paranaense que abrange os municípios de Cascavel, Marechal Cândido Rondon e Assis Chateaubriand - PR; e (iv) Espírito Santo (2019), para a variedade de São Miguel Arcanjo - SP. Esses estudos fornecem informações para a análise realizada neste trabalho de como o fenômeno do rotacismo vem sendo caracterizado em outras variedades do PB, e se essas características se assemelham ou não às encontradas para a variedade de São José do Rio Preto (SP).

Na variedade gaúcha de São José do Norte, Costa (2006) investigou o processo de alternância entre as consoantes líquidas vibrante simples ou tepe e lateral alveolar. O trabalho traça um histórico do rotacismo nas gramáticas e obras clássicas com o objetivo de confirmar a hipótese de que este é um fenômeno estável no PB, e que ocorre desde a formação dessa língua até os dias atuais. O estudo também busca justificar a organização representacional dos segmentos líquidos na Geometria de Traços (Clements; Hume, 1995) por meio de diferentes propostas da Teoria dos Traços Fonológicos Distintivos (Chomsky; Halle, 1968) para descrever o rotacismo no PB.

Como corpus de pesquisa, é utilizada uma amostra de fala originária da comunidade de São José do Norte – RS, integrante do banco de dados Variação Linguística Urbana na Região Sul (VARSUL). A amostra selecionada conta com 40 informantes estratificados socialmente em: (i) sexo (*feminino* e *masculino*); (ii) escolaridade (*até 4 anos de estudos* e *mais de 4 anos de estudos*); e (iii) faixa etária (10 informantes para a idade de *até 40 anos*, 15 informantes para a idade de *40 a 55 anos* e 15 informantes para a idade de *mais de 55 anos*).

A variável dependente definida para o estudo foi a ocorrência ou não do rotacismo. As variáveis independentes linguísticas escolhidas foram: (i) posição na sílaba (*ataque complexo* ou *coda silábica*); (ii) contexto precedente (*oclusiva*, *fricativa*,

¹³ A seleção desses estudos se deu com o intuito de exemplificar o rotacismo em cada região e por se tratarem de variedades mais próximas à investigada nesta pesquisa.

vogal anterior ou *vogal posterior*); (iii) presença de outro segmento líquido na palavra (*presença* ou *ausência*); e (iv) sonoridade do segmento precedente (*surdo* ou *sonoro*). Em seguida, os dados foram analisados a partir da Teoria Variacionista (Labov, 2008) e tratados matematicamente pelo programa *Goldvarb*.

Como resultados, para uma rodada geral, foram obtidos 1.316 dados, dos quais 11% aplicaram o rotacismo. A primeira variável selecionada como favorecedora para a realização do rotacismo foi a *posição na sílaba*, que destacou o ambiente de ataque complexo, como em *plan.ta* (peso relativo, daqui em diante PR, = 0.86), como mais propício a apresentar o fenômeno do que o contexto de coda silábica, como em *pul.so* (PR = 0.27). Esse resultado tem como uma das explicações o fato de existir menos vocábulos com a lateral como um segundo elemento de ataque, demonstrando uma preferência do PB em ter uma vibrante nesta posição, segundo a pesquisa feita pela autora por meio de dois dicionários eletrônicos, Aurélio e Houaiss (Costa, 2006, p. 97-98). A segunda variável selecionada como relevante foi a *faixa etária*, que apontou que os mais jovens, até 40 anos (PR = 0.56), e os mais velhos, mais de 55 anos (PR = 0.70), tendem a realizar mais o rotacismo do que a faixa intermediária, de 40 a 55 anos (PR = 0.21), demonstrando um padrão de variação estável nessa comunidade. A terceira variável selecionada como relevante foi a *escolaridade*, que revelou que os menos escolarizados, até 4 anos de estudo (PR = 0.57), favorecem a realização do rotacismo em relação aos mais escolarizados, com mais de 4 anos de estudo (PR = 0.41). Por fim, a última variável selecionada como relevante para essa rodada foi a *sonoridade do segmento precedente*, porém, houve um viés entre os PRs e porcentagens, pois os segmentos surdos tiveram uma aplicação de 21% e PR = 0.32, enquanto os segmentos sonoros tiveram uma aplicação de 6% e PR = 0.58. Devido a isso, foi realizado um cruzamento com a variável *segmento precedente*, que destacou que o rotacismo ocorreu mais em contexto de ataque complexo com um segmento precedente sonoro, ou seja, “[...] é mais provável uma estrutura formada por um segmento sonoro, por exemplo *bl* [...], sofrer um processo do que uma formada com um segmento surdo, por exemplo *pl* [...]” (Costa, 2006, p. 109).

Devido aos resultados relevantes para a posição silábica e com a hipótese de o rotacismo ter performances diferentes para a posição de coda silábica e ataque complexo, foram realizadas rodadas individuais para esses dois contextos. Desse modo, a rodada para Ataque Complexo obteve 462 dados, dos quais 23% tiveram a

aplicação do rotacismo. A primeira variável selecionada como relevante para esse contexto foi a *faixa etária*, que repetiu o padrão da rodada geral, em que tanto os mais novos, até 40 anos (PR = 0.64), quanto os mais velhos, mais de 55 anos (PR = 0.66) apresentaram maior probabilidade de rotacismo do que a faixa intermediária, de 40 a 50 anos (PR = 0.20), demonstrando um perfil de variação estável na comunidade. A segunda variável selecionada como relevante foi a *escolaridade*, que também repetiu o padrão da rodada geral, indicando que os menos escolarizados, até 4 anos de estudo (PR = 0.60), favorecem a realização do rotacismo em relação aos mais escolarizados, mais de 4 anos de estudo (PR = 0.37). A terceira variável selecionada como relevante para essa rodada foi a *sonoridade do contexto precedente*, que mostrou que os segmentos sonoros (PR = 0.72) precedentes favorecem mais o rotacismo do que os segmentos surdos (PR = 0.47). Foi realizado um cruzamento dessa variável com o *tipo de segmento*, que destacou que o rotacismo é mais favorecido quando o contexto precedente contém uma oclusiva sonora (40% de aplicação) do que uma fricativa surda (29%) (Costa, 2006, p. 113-114). A última variável destacada como relevante para o ambiente de ataque complexo foi o *sexo*, que indicou que o sexo/gênero feminino (PR = 0.56) favorece a aplicação do rotacismo em relação ao masculino (PR = 0.43), resultado que vai contra as teorias sociolinguísticas que apontam que as mulheres tendem a evitar as formas linguísticas estigmatizadas socialmente. A autora sugere que esse resultado pode estar relacionado ao encaixe dessa variável dentro da comunidade de fala, ou seja, a construção que o papel feminino apresenta neste lugar, assim, levando em conta que se trata de “uma comunidade rural e conservadora, podemos relativizar o papel das informantes mulheres, a maioria dona de casa, no favorecimento da regra” (Costa, 2006, p. 115).

Por fim, a rodada para coda silábica analisou 847 dados, dos quais 4% tiveram a aplicação do rotacismo. Para essa rodada, duas variáveis foram selecionadas como relevantes. A primeira foi a *faixa etária*, que retratou um resultado diferente dos encontrados para as demais rodadas, pois, nesta, apenas a faixa etária mais velha, mais de 55 anos (PR = 0.82) favoreceu a realização do rotacismo, enquanto a intermediária, 40 a 55 anos (PR = 0.32), e a mais jovem, até 40 anos (PR = 0.16), desfavoreceram o processo, o que pode indicar o declínio do rotacismo no ambiente de coda silábica. A segunda e última variável selecionada foi o *sexo*, que apresentou

o sexo/gênero masculino (PR = 0.66) como favorecedor do rotacismo em relação ao feminino (PR = 0.29), o que é coerente com as teorias variacionistas que afirmam que os homens tendem a fazer mais o uso de formas linguísticas estigmatizadas socialmente do que as mulheres. Para o ambiente de coda, foi destacado que o rotacismo entra em uma “relação de ordenamento de sangramento” (Costa, 2006, p. 119) com o processo de vocalização, ou seja, a aplicação de um processo impede a aplicação do outro. Além de também haver a preservação da lateral alveolar em coda para o falar sulista, justamente no falar de homens e dos mais velhos, o que afeta a regra do rotacismo nesse ambiente (Costa, 2006, p. 119-120).

Para a variedade carioca, Tem Tem (2010) analisou o rotacismo de // para /r/, a partir da Teoria da Variação e Mudança Linguística (Labov, 2008) e sob a visão da Geometria de Traços (Clements; Hume, 1995). Como cópula do trabalho, foram utilizados dados de fala de alunos de uma escola municipal pertencente à comunidade Jardim Moricaba, na zona oeste do Rio de Janeiro – RJ. Os dados foram coletados, pela primeira vez, por Tem Tem em 2006, por meio de um questionário contendo 31 palavras (Tem Tem, 2010, p. 61-62) e aplicados a 40 alunos: 20 pertencentes à educação infantil, com 5 anos de idade (10 de cada gênero/sexo masculino e feminino); e 20 pertencentes ao segundo ano do Ensino Fundamental, com 7 anos de idade (10 de cada gênero/sexo).

Um segundo estudo realizado quatro anos depois, em 2010, com o objetivo de analisar as mudanças ocorridas nas falas desses alunos em relação ao uso do rotacismo, recontatou esses mesmos informantes, porém somente 30 participaram do questionário: sendo 15 de cada gênero/sexo (masculino e feminino), e com idades de 9 e 11 anos.

A análise foi realizada a partir do programa estatístico *Goldvarb*, como variável dependente foi definida a realização variável do rotacismo. Como variáveis independentes linguísticas foram investigados: (i) o modo de articulação do segmento precedente à líquida (oclusivo ou fricativo); (ii) o vozeamento do segmento precedente à líquida (surdo ou sonoro); (iii) a tonicidade do grupo consonantal (tônica ou átona); (iv) a posição do grupo consonantal na palavra (inicial ou não inicial); e (v) a presença de outro segmento líquido na palavra (presença ou ausência). Já as variáveis sociais analisadas foram: (i) gênero/sexo (masculino e feminino); (ii) faixa etária (5 a 7 anos e 9 a 11 anos); (iii) escolaridade (menos de 4 anos de escolarização e mais de 4 anos

de escolarização); (iv) *escolaridade dos genitores* – pai e mãe (*ensino fundamental incompleto; ensino fundamental completo; ensino médio incompleto e ensino médio completo*).

Como resultado, foi encontrado que a porcentagem de ocorrência de rotacismo caiu depois de quatro anos, em 2006 houve 45,4% de aplicação do fenômeno, enquanto em 2010 houve 23,1% de aplicação. O estudo indica que esse resultado está relacionado ao aumento da escolarização dos alunos nesse período e ao maior acesso à norma padrão, que contribui para a diminuição do uso de processos não-padrões estigmatizados, assim como o rotacismo. Esse apontamento é confirmado pelos resultados referentes à escolaridade, em que alunos com menos de 4 anos de escolarização realizaram mais o rotacismo, com 45,4% de aplicação, enquanto os alunos com mais de 4 anos de escolarização aplicaram 24,8% de rotacismo em suas falas.

Em relação à *faixa etária*, também é apresentado um resultado que revela que os alunos, em 2006, com idades entre 5 e 7 anos, realizaram mais rotacismo, com 45,4% de aplicação, enquanto, em 2010, com idades entre 9 e 11 anos, tiveram uma porcentagem menor de aplicação: 23,1%. Em relação ao *gênero/sexo*, tanto em 2006 (frequência de 52,8%) quanto em 2010 (frequência de 29,2%), o feminino realizou mais rotacismo do que o masculino. Mesmo com uma porcentagem menor após os quatro anos, o fato de as meninas¹⁴ realizarem mais o rotacismo, indo de encontro às literaturas sociolinguísticas que indicam que estas tendem a evitar as formas estigmatizadas socialmente, deve-se ao perfil econômico e social da comunidade, em que muitas mulheres deixam a escola ainda na adolescência para cuidar dos filhos e do lar, enquanto os homens são os responsáveis por trabalhar fora e ser o provedor da família (Tem Tem, 2010, p. 89).

Sobre as variáveis linguísticas, as análises dos fatores isoladamente não demonstraram ser produtivas, o que gerou a necessidade de cruzamento entre as variáveis. Primeiramente, foi realizado o cruzamento entre as variáveis *modo de articulação* e *vozeamento*, que apresentou que as oclusivas sonoras precedentes, como em *gla.cê*, favorecem o rotacismo com 49% de aplicação do fenômeno. Depois,

¹⁴ Vale salientar que esse resultado, apesar de relevante para o mapeamento do rotacismo no PB, não pode ser comparado aos demais estudos realizados com adultos, por se tratar grupo bastante específico de crianças e pré-adolescentes.

houve o cruzamento entre as variáveis modo de articulação e tonicidade, que apontou que as fricativas átonas precedentes, como em *flo.res.ta*, apresentam uma baixa aplicação do rotacismo, com um percentual de 28%; o estudo mostra que esse contexto é favorecedor de outro processo: o apagamento da lateral, como ocorre em *floresta > foresta* (Tem Tem, 2010, p. 77). Por último, foi realizado o cruzamento entre as variáveis *modo de articulação* e *presença de outra líquida*, que trouxe o resultado de que tanto nas oclusivas, como em *glo.bo* (48%), quantos nas fricativas, como em *fla.men.go* (51%), houve um maior percentual de ocorrência de rotacismo quando há a ausência de outra líquida na palavra (Tem Tem, 2010, p. 81). Em relação à variável *posição do grupo consonantal*, a posição inicial, como em *plan.ta*, favoreceu a realização do rotacismo, tanto em 2006, com 46,6% de aplicação, quanto em 2010, com 24,4% de aplicação.

Por fim, a pesquisa ressaltou a importância do grupo de fatores *escolaridade*, pois quanto menor a escolarização, maior o número de ocorrências do rotacismo. Também comprovou que, pelo fato de o rotacismo ser estigmatizado socialmente, a implementação da variante padrão na fala desses alunos está ocorrendo de uma forma mais rápida do que para outros fenômenos, que, em estudos do tipo painel, normalmente só se verifica a mudança em um período de pelo menos uma geração.

Sobre a variedade do oeste paranaense, Cristino (2021) descreve os fenômenos do rotacismo e do retroflexo em três localidades do Paraná: Cascavel, Marechal Cândido Rondon e Assis Chateaubriand. A pesquisa tem como objetivo identificar se esses fenômenos são um processo de manutenção ou de inovação linguística na comunidade de fala investigada. Como aporte teórico-metodológico, o estudo é fundamentado na dialetologia e na geolinguística, e, também, na Sociolinguística quantitativa (Labov, 2008).

São utilizadas, como corpus de pesquisa, as respostas do inquérito fonético-fonológico produzido por Busse em 2010, no centro-oeste paranaense, que conta com 84 questões, porém, para o estudo de Cristino (2021), apenas 8 questões foram selecionadas para investigar o fenômeno do rotacismo, por resultarem em palavras com ambiente propício para a realização do fenômeno, e 17 para análise do [ɹ] retroflexo. Ao total, foram utilizadas respostas de 24 informantes, 8 de cada localidade investigada, e estratificados em três dimensões: (i) *diatrática (Analfabeto ou Ensino Fundamental Incompleto e Ensino Fundamental Completo ou Ensino Médio*

incompleto); (ii) diageracional (*de 18 a 35 anos* e *de 45 a 65 anos*); e (iii) diassexual (*masculino* e *feminino*). Também foi investigada a variável linguística *posição na sílaba* em três ambientes: *coda medial*, *coda final* e *encontro consonantal*.

A análise do rotacismo foi realizada a partir de 8 questões que resultavam nas seguintes palavras: *placa*, *planta* e *clara* (encontro consonantal); *pólvora* e *almoço* (coda medial); e *azul*, *sal* e *sol* (coda final). Foi observada ausência de rotacismo para a localidade de Marechal Cândido Rondon, o que é explicado no estudo como indício de um povoamento homogêneo dessa comunidade com bastante contato sulista e de descendência alemã, assim a forma realizada com a lateral alveolar é preservada (Cristino, 2021, p. 67, p. 94). Também não houve ocorrência de rotacismo em contexto de coda final, o que no estudo é justificado pela ocorrência de outros processos, como a velarização, como em *a.lu.gue[ʔ]*, a vocalização, como em *a.lu.gue[w]*, e o apagamento de R, como em *a.lu.gueØ*.

Sendo assim, para o contexto de encontro consonantal (Cristino, 2021, p. 68, p. 73), os homens realizaram relativamente mais rotacismo (*clara* 8%, *placa* 8%, *planta* 8%) do que as mulheres (*clara* 0%, *placa* 8%, *planta* 8%). Os menos escolarizados também realizaram mais rotacismo (*clara* 8%, *placa* 17%, *planta* 8%) do que os mais escolarizados (*clara* 0%, *placa* 0%, *planta* 8%). Por fim, os mais velhos realizaram mais rotacismo (*clara* 8%, *placa* 8%, *planta* 17%) do que os mais jovens (*clara* 0%, *placa* 8%, *planta* 0%).

Em relação ao contexto de coda medial (Cristino, 2021, p. 73, p. 76), foi obtido um resultado diferente, pelo motivo de as mulheres demonstrarem realizar mais o rotacismo (*almoço* 0% e *pólvora* 25%) do que os homens (*almoço* 8% e *pólvora* 8%). Já em relação aos demais fatores, os resultados pareceram se repetir, pois os menos escolarizados realizaram mais o rotacismo (*almoço* 0% e *pólvora* 33%) do que os mais escolarizados (*almoço* 0% e *pólvora* 0%) e, também, os mais velhos realizaram mais o rotacismo (*almoço* 8% e *pólvora* 25%) do que os mais jovens (*almoço* 0% e *pólvora* 8%).

A partir dos resultados, o estudo apontou que o rotacismo apresenta um padrão diferente de realização conforme o contexto silábico e que a aplicação do rotacismo em contexto de coda medial foi maior do que em encontro consonantal. Demonstra também que o contexto de coda final desfavorece a realização do fenômeno. A escolaridade também mostrou ser um fator relevante, pois os menos escolarizados

favoreceram a aplicação do rotacismo em relação ao mais escolarizados, o que pode dar indícios de estigma do fenômeno para a região. Em relação à dimensão diasssexual, foi possível perceber que, mesmo em proporções pequenas, o rotacismo teve maior ocorrência na fala das mulheres do que na dos homens, o que, segundo Cristino (2021, p. 94), demonstra que essa variante é concebida apenas como uma forma diferente de uso para essa comunidade de fala. Por fim, sobre a dimensão diageracional, o fato de os mais velhos aplicarem mais o fenômeno do que os mais jovens, se deve pela tendência de aqueles preservarem seus traços de origem, enquanto os mais jovens geralmente políam-se mais em entrevistas, por estarem mais expostos à norma padrão aprendida na escola.

A fim de contextualizar esta pesquisa, faz-se importante também trazer informações sobre o fenômeno do rotacismo em outra variedade paulista. Assim, tem-se a dissertação de Espírito Santo (2019) sobre o rotacismo para a variedade de São Miguel Arcanjo – SP. No trabalho, são estudados dois contextos silábicos para o rotacismo, ataque complexo e coda silábica. A investigação tem como principal objetivo investigar as noções sobre estigma e prestígios do fenômeno entre o rural vs o urbano e contou com um corpus de 24 entrevistas estratificadas em *residência* (campo ou cidade), grau de escolarização (Ensino Fundamental II e Ensino Médio), *faixa etária* (18-25 anos, 30-40 anos, 50 anos ou mais), sendo um informante para cada perfil social. Como variáveis linguísticas, foram levados em consideração a *extensão vocabular* (número de sílabas na palavra), *tonicidade*, *presença de /r/ na palavra*, *posição do // na palavra*, *contexto fônico precedente*, *contexto fônico seguinte*, *classe morfológica*, *estilo* e *item lexical*. Como aporte teórico-metodológico, foi seguida a Teoria da Variação e Mudança Linguística (Labov, 2008), e os resultados quantitativos e qualitativos demonstraram que o rotacismo em coda silábica é “fortemente evitado” na fala dos são-miguelenses, tratando-se de um estereótipo local por não chegar a 1% do total dos dados (Espírito Santo, 2019).

Já para o contexto de ataque complexo, confirmou-se a hipótese de o rotacismo ser mais realizado pelos informantes do campo, pois esse contexto foi evitado na fala dos informantes mais jovens, escolarizados e residentes da cidade. Em relação ao *status* da mudança, o rotacismo em ataque complexo demonstra estar em um processo de mudança linguística da forma rotacizada para a forma “padrão”. Por fim, a autora declara que, apesar de haver um certo grau de estigmatização local para o

rotacismo em ataque complexo, devido a essa forma ser evitada na fala dos mais escolarizados e das mulheres, não é possível afirmar haver estigma do rotacismo em São Miguel Arcanjo, pois, quando aplicado um pequeno estudo de avaliação com a frase “*O Dougras imprica com tudo*”, metade dos informantes não notaram a pronúncia rotacizada.

Após a introdução do referencial teórico relacionado à Teoria da Variação e Mudança Linguística (Labov, 2008), ao fenômeno do rotacismo e a alguns estudos que abordaram esse fenômeno em diferentes variedades do PB, o próximo capítulo será dedicado à descrição do material e dos métodos utilizados nesta pesquisa. Serão apresentadas informações sobre o *cópus* do trabalho, bem como os procedimentos metodológicos adotados para a realização do levantamento e da análise dos dados que embasaram os resultados e discussões a serem apresentados posteriormente.

3 MATERIAL E MÉTODOS

Este capítulo abrange informações referentes ao material e aos métodos utilizados nesta dissertação. Para isso, está dividido em quatro seções, que são: (i) *Comunidade linguística*, em que se realiza a descrição da comunidade de São José do Rio Preto (seção 3.1); (ii) *Cópus de pesquisa*, que descreve o banco de dados Iboruna (Gonçalves, 2024) (seção 3.2); (iii) *Variáveis estudadas*, que apresenta as variáveis dependente e independentes consideradas na investigação do fenômeno do rotacismo no interior de São Paulo (seção 3.3); e, por último, (iv) *Procedimentos metodológicos*, em que são descritos os passos seguidos durante o levantamento dos dados analisados nesta pesquisa (seção 3.4).

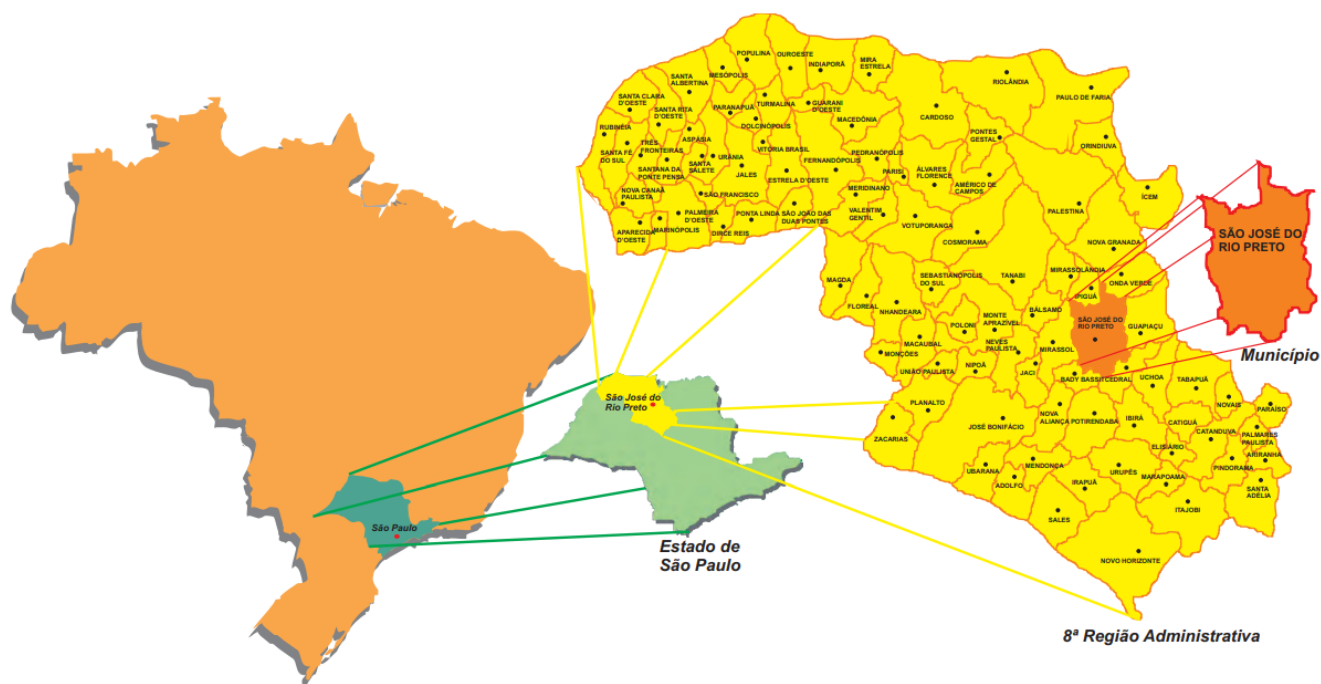
3.1 COMUNIDADE LINGUÍSTICA

Para compreender melhor os resultados encontrados por meio desta pesquisa, é necessário olhar para a comunidade linguística investigada, pois, como a sociolinguística (Labov, 2008) considera a língua como social e heterogênea, dados sociais como distribuição da comunidade entre sexos/gêneros, faixas etárias e escolaridades podem impactar na produtividade ou não de um fenômeno linguístico em determinada variedade. Sendo assim, serão apresentadas algumas informações referentes a esses dados sobre São José do Rio Preto.

Situada no noroeste do estado de São Paulo, São José do Rio Preto é localizada a uma distância de 442 km da capital São Paulo, e segundo o *site* da prefeitura do município, abrange um território de 431,963 km (Prefeitura Municipal de São José do Rio Preto, 2024). O município pertence à mesorregião e microrregião de mesmo nome. Conforme o censo de 2022¹⁵ do Instituto Brasileiro de Geografia e Estatística (IBGE, 2021), sua população é estimada em mais de 480 mil pessoas, tornando-o, assim, o 10º município mais populoso em seu estado e o 45º no Brasil. O mapa a seguir ilustra um panorama da localização da cidade de São José do Rio Preto em relação aos demais municípios da região administrativa homônima, depois em relação ao estado de São Paulo, e, de forma mais abrangente, em relação ao Brasil.

¹⁵ Optou-se por trazer os dados sobre a região mais atuais ao ano de produção da dissertação.

FIGURA 1 - Mapa de localização de São José do Rio Preto – SP



Fonte: São José do Rio Preto (2020, p. 15)

A formação histórica de São José do Rio Preto ocorreu no final do século XIX, quando pioneiros vindos de Minas Gerais chegaram à região e se estabeleceram, marcando o início da pecuária e da agricultura na área. Segundo Francisco (2011), durante esse período, pelo fato de a cidade estar localizada em um lugar estratégico ao longo da antiga Estrada Geral que conectava as províncias do Império, desde o Triângulo Mineiro até Cuiabá, o município serviu de apoio à colonização da região oeste ainda não desbravada. Apesar das más condições, a estrada geral era uma rota bastante conhecida, o que fez de São José do Rio Preto um ponto de parada frequente para os tropeiros que seguiam para o sertão (Francisco, 2011, p. 120).

O município começou a evoluir economicamente em 1912 com a construção da Estrada de Ferro Araraquarense, que tornou a cidade o principal eixo para o movimento do comércio agrícola da região centro-oeste do país (Francisco, 2011). Conforme o histórico disponível no *site* da Fundação Sistema Estadual de Análise de Dados (SEADE), no começo do século XX, a economia da cidade era influenciada pela cafeicultura e cultura de grãos (Fundação Seade, 2024).

A empresa *Swift*, indústria de extração de óleo comestível, retirado do algodão, amendoim, mamona e milho, foi uma das pioneiras a serem instaladas no município, no ano de 1942. A partir dessa época, conforme Francisco (2011), outros municípios foram sendo constituídos, fazendo com que a região urbana se expandisse e São José do Rio Preto se consolidasse como um polo regional.

Porém, foi a partir da década de 70, com o incentivo do governo, que visava ao desenvolvimento industrial do interior paulista, que São José do Rio Preto cresceu e se tornou um dos principais centros industriais, culturais e de serviços do interior de São Paulo (Francisco, 2011).

Nos dias atuais, o setor agrícola e industrial permanece forte para a economia de São José do Rio Preto. O *site* Desenvolve SP destaca a produção de laranjas, cana-de-açúcar e café, e indústrias nas áreas de alimentos, calçados e metalurgia como os principais geradores de renda do município.

São José do Rio Preto apresenta o Produto Interno Bruto (PIB) *per capita* de 40.204,08, de acordo com o último censo do IBGE em 2022. São José do Rio Preto também se encontra no Grupo 1 do Índice Paulista de Responsabilidade Social (IPRS), que, segundo o *site* SEADE, é onde se encontram os municípios caracterizados com bons níveis nos indicadores sociais e nível elevado de riquezas (Fundação Seade, 2024).

Apesar de a renda não ser uma variável em análise nesta pesquisa, é cabível apresentar alguns dados sobre esse fator para a variedade investigada. A tabela a seguir apresenta os índices de rendimento mensal domiciliar em São José do Rio Preto em 2010:

TABELA 2 - Rendimento mensal para domicílios particulares permanentes em São José do Rio Preto - SP

Município	Classes de rendimento nominal mensal domiciliar (salário-mínimo) (2)							Sem rendimento (3)
	Total (1)	Mais de 1/2 a 1	Mais de 1 a 2	Mais de 2 a 5	Mais de 5 a 10	Mais de 10 a 20	Mais de 20	
São José do Rio Preto	137.233	7.503	20.354	55.775	32.374	11.579	4.514	4.085

Nota: (1) Inclusive os domicílios sem declaração de rendimento nominal mensal domiciliar. (2) Salário-mínimo utilizado: R\$ 510,00. (3) Inclusive os domicílios com rendimento mensal domiciliar somente em benefícios.

Fonte: São José do Rio Preto (2020, p. 33)

Como descrito na tabela, com base em 137.233 domicílios particulares permanentes, a média de renda era de 2 a 5 salários-mínimos por domicílio, o que aponta que os 16 informantes determinados para esta investigação estão dentro da média de rendimento do município, pois todos são pertencentes à renda familiar de até 5 salários-mínimos (Gonçalves, 2019, p. 286).

No que tange à educação, segundo o censo de 2010 do IBGE, São José do Rio Preto apresenta uma taxa de escolarização de 98% de crianças entre 6 a 14 anos. Seu Índice de Desenvolvimento da Educação Básica (IDEB, 2021), responsável por indicar a qualidade do ensino das escolas, no ano de 2021, obteve notas relativamente acima da média brasileira, que foram de 5,8 para os anos iniciais do ensino fundamental e 5,1 para os anos finais do ensino fundamental, sendo em São José do Rio Preto a nota de 6,2 nos anos iniciais do ensino fundamental, e a nota de 5,4 para os anos finais do ensino fundamental, ambas em escolas públicas.

Ainda sobre educação, consta também no censo de 2010 do IBGE que essa região apresenta níveis adequados de escolarização, pois na faixa etária de 10 ou mais de idade, aproximadamente 137 mil não possuíam instrução ou apresentavam o ensino fundamental incompleto, 69 mil pessoas apresentavam o ensino fundamental completo, 97 mil pessoas concluíram o ensino médio, enquanto 55 mil pessoas, por fim, possuíam o ensino superior completo.

Finalmente, para a faixa etária de 25 anos ou mais de idade, aproximadamente 82 mil pessoas não possuíam instrução ou possuíam o ensino fundamental incompleto, 36 mil completaram o ensino fundamental, 55 mil concluíram o ensino médio, e, por fim, 40 mil completaram o ensino superior.

Posto isso, passa-se à seção 3.2, em que serão apresentados o corpus e informações sobre o banco de dados utilizado neste trabalho.

3.2 CÓRPUS DA PESQUISA

Como corpus desta dissertação, são utilizadas 16 entrevistas sociolinguísticas estratificadas em dois sexos/gêneros (feminino e masculino), duas escolaridades (1º ciclo do Ensino Fundamental e Ensino Superior) e quatro faixas etárias (16 a 25 anos; 26 a 35 anos; 36 a 55 anos; e mais de 55 anos), pertencentes ao banco de dados Iboruna (Gonçalves, 2024), o qual é constituído por amostras de fala de informantes da região de São José do Rio Preto, apresentada na seção anterior.

O Iboruna originou-se a partir do Projeto Amostra Linguística do Interior Paulista (ALIP), que foi idealizado inicialmente por pesquisadores da Universidade Estadual Paulista (Unesp) de São José do Rio Preto, com os objetivos de prover um meio metodológico essencial para pesquisadores e de possibilitar avanços nas pesquisas linguísticas ao dispor um banco de dados anotado com amostras que capturassem o dinamismo da fala do interior paulista, que, até então, havia sido pouco estudada cientificamente (Gonçalves, 2024).

A constituição do banco de dados foi subsidiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP/Processo 03/08058-6), que possibilitou as coletas de amostras, entre os anos de 2004 a 2007, em sete municípios da região noroeste do Estado de São Paulo – Bady Bassit, Cedral, Guapiaçu, Ipiguá, Mirassol, Onda Verde e São José do Rio Preto –, ilustrados no mapa abaixo.

FIGURA 2 - Mapa de São José do Rio Preto e municípios limítrofes



Fonte: Junior (2011).

Segundo Gonçalves (2019), um dos idealizadores do projeto, em 2019 o Iboruna contava com trezentos pesquisadores, tanto do Brasil quanto do exterior, cadastrados como usuários do banco de dados. Os trabalhos já realizados incluem inúmeros estudos com diferentes aportes teóricos, dentre eles, análises variacionistas que contemplam a perspectiva fonético-fonológica, como, por exemplo, sobre (i) apagamento e alçamento vocálico em contexto postônico medial, como em *a.bó.bra* ~ *a.bó.b[u].ra* (Ramos, 2009); (ii) apagamento de /d/ em morfema de gerúndio, como em *cantanØo* (Ferreira, 2010); (iii) alçamento vocálico de vogais médias pretônicas, como em *p[i]queno* e *c[u]sturando* (Silveira, 2008; Carmo, 2009, 2013, 2014, 2018, 2019; Carmo; Carlos, 2019; Carmo; Tenani, 2013); (iv) apagamento de /r/ em coda silábica (Carmo; Taborda, 2019), como em *ficáØ* e *governadØØ*; (v) metátese (Dias; Carmo, 2021), como em *acobracia* e *dromir*; (vi) monotongação do ditongo oral decrescente [ej], como em *deØxa* e *dinheØro* (Carmo; Machado, 2023); (vii) monotongação do ditongo oral decrescente [ow], como em *roØpa* e *trabalhØØ* (Rezende; Carmo, 2023); e (viii) ditongação diante de /S/, como em *arrol[j]z* e *mê[j]s* (Prodelik; Carmo, 2024).

A constituição do banco de dados, conforme seu *site* oficial (Gonçalves, 2024 [2007]), é feita por dois tipos de amostras: (i) Amostra Censo ou Amostra Comunidade

(AC), que apresenta 152 entrevistas sociolinguísticas coletadas a partir de perfis sociais previamente estabelecidos; e (ii) Amostra Interação (AI), que contempla 11 interações dialógicas coletadas secretamente em contextos de interação social livre, sem o controle de perfis sociais. Cada uma das amostras é composta por arquivos sonoros; arquivos de ficha social dos informantes; arquivos com transcrição ortográfica da fala; e arquivos de diário de campo, disponíveis na forma *online* pelo *site* oficial do projeto ALIP.

Como justificativa de escolha para o uso de entrevistas sociolinguísticas pertencentes à Amostra Censo, tem-se o fato de esta estratificar os informantes socialmente, assim, é a mais propícia para estudos sociolinguísticos que averiguem a influência de variáveis independentes extralinguísticas em relação à aplicação do processo variável estudado. Vale ressaltar que, para esta pesquisa, foi utilizada uma subamostra da AC, isto é, não são utilizadas todas as entrevistas disponíveis na amostra.

Após a explanação sobre o *cópus* da pesquisa, passa-se à próxima seção, em que são apresentadas, de forma detalhada, a variável dependente e as variáveis independentes linguísticas, bem como os grupos de fatores extralinguísticos que norteiam esta pesquisa.

3.3 VARIÁVEIS ESTUDADAS

A variável dependente é utilizada para descrever o fenômeno linguístico que está sendo estudado e que é influenciado por fatores linguísticos e sociais. Em outras palavras, é a característica linguística que é afetada ou determinada por outras variáveis na pesquisa sociolinguística (Coelho *et al.*, 2015). Para esta análise, a variável dependente é a realização variável do rotacismo, ou seja, a presença (*pra.ca*) ou não presença do rotacismo (*pla.ca*).

As variáveis independentes são compostas pelos grupos de fatores que podem ter influência sobre a variável dependente estudada. Elas são manipuladas ou selecionadas pelo pesquisador para entender como afetam ou se relacionam com a variável dependente, podendo ser linguísticas ou extralinguísticas - ou seja, de natureza social, como sexo/gênero, escolaridade, faixa etária, *etc.* Essas variáveis são consideradas independentes porque se defende que elas possam ter um efeito ou

correlação com a variável dependente, mas não são influenciadas por ela (Coelho *et al.*, 2015).

As 16 entrevistas analisadas na presente pesquisa são resultado da combinação das variáveis sociais, sendo elas quatro faixas etárias, dois sexos/gêneros e duas escolaridades ($4 \times 2 \times 2 = 16$), como mostra o quadro a seguir:

QUADRO 2 - Variáveis extralinguísticas

Faixa etária	Escolaridade	Sexo/gênero	
		Masculino	Feminino
16-25 anos	1º Ciclo EF	AC-031	AC-032
	Ensino Superior	AC-055	AC-056
26-35 anos	1º Ciclo EF	AC-063	AC-064
	Ensino Superior	AC-087	AC-088
36-55 anos	1º Ciclo EF	AC-095	AC-096
	Ensino Superior	AC-119	AC-120
Mais de 55 anos	1º Ciclo EF	AC-127	AC-128
	Ensino Superior	AC-151	AC-152

Fonte: Elaboração própria, com base em Gonçalves (2019, p. 286).

Nesta dissertação, como variáveis independentes, são considerados quatro grupos de fatores linguísticos e três extralinguísticos, sendo eles: (i) posição do segmento na sílaba; (ii) contexto precedente ao segmento; (iii) sonoridade precedente ao segmento; (iv) outro segmento líquido na palavra; (v) sexo/gênero; (vi) escolaridade; e (vii) faixa etária.

A seguir, são apresentadas as justificativas de escolha dessas variáveis bem como sua caracterização detalhada:

i) *Posição do segmento na sílaba*: essa variável é considerada com o objetivo de verificar se o contexto de ataque complexo (*clu.be*) é mais favorável à ocorrência do rotacismo em comparação aos contextos de coda silábica medial (*fal.tar*), e de coda silábica final (*a.tu.al*) na variedade do interior paulista. Essa análise é feita com base em estudos anteriores, como o de Costa (2006) e Cristino (2021), que constatarem uma preferência pelo rotacismo em contextos de ataque complexo. Portanto, a investigação dessa variável busca identificar se existem semelhanças ou

diferenças nos padrões de realização do rotacismo em relação à posição do segmento na sílaba na variedade estudada em comparação com outros estudos.

ii) *Contexto precedente ao segmento*: essa variável é utilizada com o objetivo de investigar se o contexto precedente, seja uma *oclusiva* (e.xem.**plo**), uma *fricativa* (**flor**), uma *vogal anterior*¹⁶ (pa.**pel**) ou uma *vogal posterior* (**vol.to**), pode favorecer a ocorrência do rotacismo. Essa variável é relevante, uma vez que estudos anteriores, como os de Costa (2006) e Tem Tem (2010), já demonstraram sua importância na análise desse fenômeno linguístico. Portanto, é interessante examinar se os resultados obtidos neste estudo se aproximam ou se afastam das conclusões dessas pesquisas anteriores no que diz respeito à variedade analisada.

(iii) *Sonoridade precedente ao segmento*: investiga-se o contexto precedente à consoante líquida em relação à sonoridade, mais especificamente se é *vozeado* (**al.gum**) ou *desvozeado* (**sim.ples**). O objetivo é analisar se essa variável tem relevância para a aplicação do rotacismo, conforme sugerido pelo estudo de Costa (2006), no qual o contexto precedente vozeado foi identificado como favorecedor da ocorrência do fenômeno para a variedade gaúcha. Dessa forma, busca-se verificar se os resultados obtidos neste estudo corroboram a conclusão de Costa (2006) ou se apresentam resultados diferentes em relação à influência da sonoridade na realização do fenômeno do rotacismo no que tange ao noroeste paulista.

(iv) *Outro segmento líquido na palavra*: a investigação dessa variável se baseia na análise da influência da presença (**nor.mal**) ou ausência (**i.gual**) de outro segmento líquido // ou /r/ na ocorrência do rotacismo, como constatado por Tem Tem (2010) na variedade carioca. Esse estudo revelou um leve favorecimento para a realização do rotacismo em palavras que contêm outra líquida, como em (**fla.gra**). Portanto, neste estudo, pretende-se examinar se os resultados encontrados por Tem Tem (2010) também se aplicam à variedade estudada, verificando se a presença de outra líquida influencia ou não a ocorrência do rotacismo.

(v) *Sexo/gênero*: é uma variável utilizada com base na literatura sociolinguística, pois permite identificar diferenças significativas na maneira como homens e mulheres usam a linguagem. Um exemplo é o clássico estudo realizado na cidade de Nova York, em que Labov (2008) identificou diferenças na pronúncia do /r/

¹⁶ Os termos *anterior* e *posterior* referem-se aos traços articulatórios da vogal precedente ao segmento líquido e não à posição delas em relação a esse segmento.

entre homens e mulheres. Esses estudos sociolinguísticos também sugerem que as mulheres têm uma tendência a evitar formas estigmatizadas socialmente em comparação aos homens: “[...] as mulheres usam as formas mais avançadas em sua própria fala informal e se corrigem mais nitidamente no outro extremo da fala monitorada” (Labov, 2008, p. 346). Isso pode ser atribuído ao papel social de subordinação que as mulheres historicamente enfrentam, resultando em uma insegurança linguística que as leva a evitar estilos estigmatizados e buscar prestígio e reconhecimento dentro de seu grupo social por meio da linguagem (Barrozo; Aguilera, 2014). A análise dessa variável pode fornecer *insights* sobre as atitudes e comportamentos linguísticos relacionados ao rotacismo, especialmente no que diz respeito às diferenças de gênero e às pressões sociais que influenciam o uso da variante prestigiada ou estigmatizada. Vale salientar que é adotada nesta pesquisa a nomenclatura *sexo/gênero*¹⁷ pelo fato de seguir a mesma terminologia utilizada pelo banco de dados Iboruna (Gonçalves, 2024), sem tomar *sexo* e *gênero* como sinônimos.

(vi) *Escolaridade*: é uma variável relevante para averiguar a atuação de diferentes níveis educacionais em relação à realização do rotacismo, a fim de identificar possíveis indícios de estigmas sociais relacionados ao fenômeno na fala do noroeste paulista. Este estudo optou por investigar dois extremos, sendo o *1º ciclo do ensino fundamental* e *ensino superior*, para verificar o contraste entre esses fatores. Conforme apresentado anteriormente, o rotacismo é considerado um dos fenômenos mais estigmatizados socialmente no PB, visto que, “[...] para os falantes urbanos escolarizados, pronúncias como broco, ingrês, chicrete, pranta, etc. são feias, erradas e toscas” (Bago, 2007, p. 73). Esse preconceito linguístico ocorre devido à associação desse fenômeno com a fala de indivíduos socialmente desprivilegiados, como analfabetos ou aqueles com acesso limitado à educação formal. Essa percepção negativa em relação ao rotacismo decorre do estigma atribuído a determinadas variantes linguísticas, que são consideradas inferiores ou inadequadas dentro de certos contextos sociais. É importante destacar que esse preconceito reflete uma visão linguística baseada em normas e valores sociais, e não em uma avaliação

¹⁷ Para melhor compreensão dos usos de nomenclaturas como *sexo/gênero* por bancos de dados linguísticos brasileiros, recomenda-se a leitura do capítulo *(Re)Discutindo Sexo/Gênero na Sociolinguística*, de Freitag (2015, p. 17-74).

objetiva das variantes linguísticas em si. Sendo assim, ao analisar a escolaridade, busca-se compreender como diferentes níveis educacionais podem influenciar o uso do rotacismo e, conseqüentemente, identificar a existência ou não de estigma social relacionado a essa variante linguística na variedade estudada.

(vii) *Faixa etária*: utilizada para investigar indícios que possam indicar se o fenômeno do rotacismo em São José do Rio Preto corresponde a um caso de variação estável ou mudança em progresso. Para realizar essa análise, é adotada uma abordagem de estudo em tempo aparente, em que é realizado um recorte sincrônico de amostras de fala de diferentes gerações (Paiva; Duarte, 2015). Na pesquisa em tempo aparente, não há um acompanhamento longitudinal da fala do mesmo informante, mas há uma comparação da “frequência de uso de uma determinada variável distribuída por faixas etárias [...]” (Costa, 2006, p. 78). Assim, uma vez que determinada variante linguística seja mais frequente na fala de pessoas mais jovens, pode tratar-se de duas hipóteses: a primeira sugere uma mudança linguística, ou seja, a instalação gradual de uma nova variante na língua; e a segunda sugere uma diferenciação linguística etária, que não gera uma mudança no sistema linguístico, os falantes apenas apresentam uma alteração no seu comportamento linguístico ao longo da vida (Paiva; Duarte, 2015, p. 179). A partir dessa metodologia, é possível examinar se o uso do rotacismo se mantém constante ao longo do tempo ou se está passando por mudanças evolutivas na variedade estudada. Essa abordagem ajuda a compreender a dinâmica e a trajetória do fenômeno do rotacismo na comunidade linguística do noroeste paulista.

Na seção a seguir, são apresentados os procedimentos metodológicos adotados para o levantamento e a análise dos dados desta pesquisa.

3.4 PROCEDIMENTOS METODOLÓGICOS

Para realizar o levantamento dos dados, foram utilizados os arquivos de áudio e de transcrição ortográfica das dezesseis entrevistas selecionadas do banco de dados Iboruna (Gonçalves, 2024). Para fins de análise, foi feita uma pré-seleção nos arquivos de transcrição ortográfica, identificando palavras que apresentavam contexto para a ocorrência do rotacismo, ou seja, aquelas que apresentavam consoante líquida lateral em posição de ataque complexo (*cli.ma*), coda medial (*al.gum*) ou coda final (*fá.cil*).

Em seguida, foi realizada uma análise de oitiva dos arquivos de áudio para observar a ocorrência ou não do processo fonético-fonológico, bem como seus contextos linguísticos e extralinguísticos, e assim poder codificar esses dados. A análise de oitiva se deve ao fato de os arquivos sonoros do banco de dados Iboruna não apresentarem qualidade o suficiente para uma análise acústica dos dados, tendo em vista o fato de as gravações não terem sido realizadas em cabines com isolamento acústico.

A partir da codificação dos dados levantados na análise de oitiva, foi conduzida uma análise quantitativa utilizando o programa estatístico *Goldvarb X* disponibilizado para *download* em seu site, que é um *software* estatístico usado principalmente para a análise de dados sociolinguísticos, especialmente no estudo da variação linguística. Esse programa foi inicialmente desenvolvido na década de 70, por David Sankoff e Sali Tagliamonte com contribuições de outros linguistas e estatísticos ao longo dos anos. Essa ferramenta tem como uma das suas principais funcionalidades a de realizar análises de regressão logística binária e multinomial, permitindo aos pesquisadores examinar, por meio de pesos relativos e sua interface gráfica, como diferentes variáveis linguísticas e sociais influenciam a variação e mudança linguística (Sankoff; Tagliamonte; Smith, 2005).

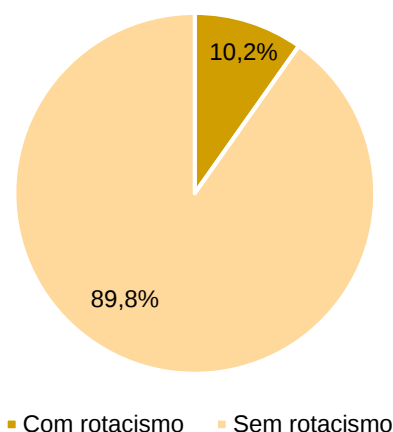
Por fim, é realizada uma análise qualitativa desses dados com base principalmente na Teoria da Variação e Mudança Linguística, de Labov (2008). Também é feita a discussão dos resultados com base em pesquisas anteriores sobre o tema, a fim de verificar possíveis relações quanto ao uso do rotacismo na variedade paulista de São José do Rio Preto. Esses resultados são descritos e discutidos no capítulo a seguir.

4 ANÁLISE E DISCUSSÃO DOS DADOS

Neste capítulo, são expostos e discutidos qualitativamente os resultados da análise estatística do fenômeno do rotacismo na variedade do noroeste paulista. Os dados quantitativos foram obtidos a partir do programa *Goldvarb X*, que identificou os contextos que favorecem ou não a ocorrência do fenômeno. Esses grupos de fatores são apresentados nas tabelas, juntamente com seus pesos relativos (PRs) e frequências de ocorrência.

A partir das dezesseis entrevistas sociolinguísticas, foi levantado o total de 1.915 dados, ou seja, palavras que tinham contexto para a realização do rotacismo. Desse total, 196 dados (10,2%) apresentaram a aplicação do rotacismo, como nas palavras *di.p[r]o.ma* e *a[l].to*. O gráfico a seguir ilustra as ocorrências gerais:

GRÁFICO 1 - Ocorrências gerais de rotacismo



Fonte: Elaboração própria

Para distinguir se esse fenômeno se trata de um processo de variação estruturada ou de um processo de flutuação típico da fala (acidental de performance), recorre-se ao texto *Some sociolinguistic principles* (Labov, 2003, p. 243), em que o autor propõe uma solução quantitativa para resolver o problema:

TABELA 3 - Frequências de aplicação para regras linguísticas

Frequência de aplicação	Tipo de regra
100%	Categórica
95-99%	Semi-categórica
5-95%	Variável

Fonte: Elaboração própria, com base em Labov (2003)

A partir disso, devido à sua frequência de aplicação de 10,2%, o rotacismo se comprova como uma regra variável para a variedade de São José do Rio Preto.

Agora passa-se à observação das variáveis descartadas e destacadas pelo programa estatístico *Goldvarb X*. As variáveis descartadas são aquelas que não demonstraram relevância estatística para a ocorrência do fenômeno do rotacismo. Neste estudo, essas variáveis foram duas: *outro segmento líquido na palavra* e *sonoridade precedente ao segmento*.

A variável *outro segmento líquido na palavra* investigava se a presença de outro /r/ ou /l/ na mesma palavra poderia influenciar a ocorrência do rotacismo, porém, o programa estatístico descartou essa variável durante o *stepping down*,¹⁸ devido aos dois fatores (presença e ausência de líquida) apresentarem porcentagens muito próximas. Para o fator *ausência de outra líquida na palavra*, foram levantados 133 dados, desses, 10,3% apresentaram o rotacismo, enquanto 89,7% não apresentaram a ocorrência do fenômeno; isso indica que, na ausência de outra líquida na palavra, a ocorrência de rotacismo é relativamente baixa. Já para o fator *presença de outra líquida na palavra*, foram levantados 63 dados, dos quais 10,1% resultaram em rotacismo, enquanto 89,9% não resultaram; isso indica que, na presença de outra líquida na palavra, a ocorrência de rotacismo também é baixa e quase idêntica ao caso sem outra líquida. Assim, as porcentagens muito próximas entre os dois fatores (10,3% e 10,1%) sugerem que a presença de outra líquida na palavra não tem um efeito significativo sobre a realização do rotacismo. Tanto na presença de outra líquida, como em *descartável*, quanto na ausência, como em *fácil*, a probabilidade de ocorrência do rotacismo permanece praticamente a mesma. Ou seja, a análise dos

¹⁸ O processo de *stepping down* no *GoldVarb X* é usado para identificar e eliminar fatores que não contribuem significativamente para a explicação da variação nos dados. Quando o programa estatístico descarta um fator durante esse processo, é porque, estatisticamente, esse fator não melhora a predição do modelo de maneira significativa (Sankoff; Tagliamonte; Smith, 2005).

dados indica que a variável *outro segmento líquido na palavra* não é determinante para a ocorrência do rotacismo.

A outra variável descartada pelo *Goldvarb X* foi a *sonoridade precedente ao segmento*, que investigava se o fato de o segmento que antecede a consoante líquida ser vozeado ou desvozeado poderia favorecer a realização do rotacismo. O fator *vozeado* obteve 120 dados, destes 7,4% tiveram a ocorrência do rotacismo, enquanto 92,6% não resultaram na ocorrência do fenômeno. Isso sugere que, quando o segmento precedente é vozeado, a realização do rotacismo é menos frequente. Para o fator *desvozeado*, 76 dados foram obtidos, desses 25,6% resultaram em rotacismo, enquanto 74,4% não resultaram. Isso sugere que, quando o segmento precedente é desvozeado, a realização do rotacismo é um pouco mais frequente. Embora os dados mostrem diferença nas taxas de rotacismo para os dois fatores, o *Goldvarb X* descartou essa variável por não apresentar uma contribuição significativa estatisticamente, isso pode significar que essa variável possa estar apresentando informações redundantes, que já foram capturadas por outros fatores que podem explicar a variação de forma mais eficaz.

Sendo assim, passa-se para as variáveis destacadas, a partir dos dados, como favorecedoras para a aplicação do rotacismo. O programa estatístico forneceu, em ordem decrescente de relevância, as variáveis que influenciam a aplicação do processo variável do rotacismo, sendo elas:

- (i) posição na sílaba;
- (ii) sexo/gênero;
- (iii) escolaridade;
- (iv) faixa etária; e
- (v) contexto precedente ao segmento.

A seguir, os resultados relativos a esses grupos de fatores serão descritos detalhadamente, junto de suas respectivas tabelas, em que são apresentados os PRs e as frequências de aplicação do rotacismo.

4.1 POSIÇÃO NA SÍLABA

A variável linguística *posição na sílaba* foi apontada em primeiro lugar em ordem de relevância para a aplicação do rotacismo. Esse resultado foi observado

também para a variedade gaúcha de São José do Norte (Costa, 2006). Nos resultados da presente pesquisa, o contexto de ataque complexo se destaca como favorecedor da realização do rotacismo, como se pode observar na tabela 4:

TABELA 4 - Rotacismo em relação à *posição na sílaba*

	Ataque Complexo			Coda Medial			Coda Final		
	Apl./Total	%	PR	Apl./Total	%	PR	Apl./Total	%	PR
Com rotacismo	90/374	24,1	0.789¹⁹	94/780	11,9	0.685	12/753	1,6	0.187
Sem rotacismo	284/374	75,9	0.211	686/780	88,1	0.315	741/753	98,4	0.813

Fonte: Elaboração própria

Input: 0.025

Significância: 0.000²⁰

Conforme evidenciado na Tabela 4, o ataque complexo, como em *p[r]an.ta.va* e *re.c[r]a.ma.va*, é o contexto que mais favorece a aplicação do rotacismo, com PR = 0.789, seguido pelo contexto de coda medial, com PR = 0.685, como em *a[l].quei.re* e *re.so[l].via*, e, por fim, desfavorece o rotacismo o contexto de coda final, como em *sa[w]* e *a.ni.ma[w]*, com PR = 0.187.

Uma das possibilidades de explicação para o rotacismo ser mais provável em contexto de ataque complexo e em coda medial pode estar relacionada ao fato de que, em ambiente de coda final, as consoantes líquidas estão mais suscetíveis a outros processos variáveis, como a vocalização: *pa.pe[w]*; velarização: *sa[h]*; e apagamento do /r/: *jor.naØ*. No estudo de Cristino (2021) sobre a variedade do oeste paranaense, foi encontrado resultado semelhante em relação ao desfavorecimento do rotacismo em contexto de coda final. Conforme destacado pela autora, esse desfavorecimento pode ser explicado pelo processo de vocalização, que é definido por Silva (2011, p. 220) como um “fenômeno fonológico de alteração de uma consoante para vogal. No português, a lateral pós-vocálica vocalizou-se na maioria dos dialetos e é manifestada como um glide posterior [w]”. Assim, as palavras *sa/* e *animal* seriam realizadas como *sa[w]* e *a.ni.ma[w]*. Cristino (2021) também identificou ocorrências de velarização de /l/ em contexto de coda final. Segundo Silva (2011, p. 219), esse processo ocorre quando a lateral pós-vocálica alveolar é coarticulada com o “levantamento da parte posterior da língua”, como em *sa[h]* e *me[h]*. Assim, cabe retornar à explicação de Costa (2006), a partir da teoria de Kenstowick (1994) sobre

¹⁹ Nesta tabela e nas seguintes, são destacados em negrito os PRs referentes à aplicação do rotacismo.

²⁰ Convergência na iteração 8 para todas as tabelas.

o ordenamento de sangramento que ocorre quando a aplicação de uma *regra A* impede a aplicação de uma *regra B*. Em outras palavras, a *regra A* "sangra" a *regra B*, fazendo com que o ambiente no qual a *regra B* seria aplicada não exista mais após a aplicação da regra:

A regra de vocalização da lateral na coda silábica aplica-se primeiro e desfaz o contexto para aplicação do rotacismo. Assim, o baixo índice de rotacismo na coda pode ser atribuído à interferência da vocalização que a lateral sofre como um processo de mudança em progresso, apontado em estudos variacionistas realizados sobre a lateral (Costa, 2006, p. 119-120).

Além disso, Carmo e Taborda (2019), em sua pesquisa sobre o apagamento de /r/ na variedade do noroeste paulista, observaram que “[...] quando o segmento se localiza em meio de vocábulo, há tendência de desfavorecimento do apagamento de /r/ em coda. Em contrapartida, quando a ocorrência é em fim de vocábulo, o apagamento é altamente favorecido” (Carmo; Taborda, 2019, p. 47). Assim, o contexto de coda final favorece a omissão desse som, como nos casos de *fa.láØ* e *jo.ga.dôØ*.

Sendo assim, os resultados encontrados nesta pesquisa estão de acordo com os encontrados nas demais variedades do PB, como de Costa (2006), Tem Tem (2010) e Cristino (2021), em que o contexto de ataque complexo favorece a aplicação da regra do rotacismo, confirmando a hipótese inicial em relação a esse fator linguístico. Além disso, os resultados demonstraram que a coda medial também é um ambiente favorecedor do rotacismo, pois está menos suscetível à aplicação de outros processos variáveis como a vocalização e velarização, em comparação com a coda final, que é um ambiente propício à realização de outras regras que, quando aplicadas, entram em ordenamento de sangramento com o rotacismo, impedindo a sua realização.

4.2 SEXO/GÊNERO

A variável extralinguística *sexo/gênero* foi apontada como a segunda mais relevante para a aplicação do rotacismo. O resultado demonstra que o sexo/gênero masculino favorece a aplicação do processo, como ilustra a tabela a seguir.

TABELA 5 - Rotacismo em relação ao sexo/gênero

	Feminino			Masculino		
	Apl./Total	%	PR	Apl./Total	%	PR
Com rotacismo	13/742	1,8	0.182	183/1173	15,6	0.721
Sem rotacismo	729/742	98,2	0.818	990/1173	84,2	0.279

Fonte: Elaboração própria.

Input: 0.025

Significância: 0.000

Conforme a tabela 5, o sexo/gênero feminino apresentou PR = 0.182, enquanto o sexo/gênero masculino favoreceu a aplicação do rotacismo com PR = 0.721, o que demonstra uma grande distância de probabilidade de aplicação do fenômeno entre os dois sexos/gêneros.

Em relação às outras variedades do PB, Costa (2006) - variedade de São José do Norte; Cristino (2021) - variedade do oeste paranaense; e Tem Tem (2010) - variedade carioca, foram obtidos resultados diferentes, em que o sexo/gênero feminino realizou mais rotacismo, mesmo que em pequenas proporções, do que o sexo/gênero masculino. Inclusive, a pesquisa de Cristino (2021) aponta que as mulheres lideraram a realização de rotacismo em contexto de coda medial. Como argumento, as autoras sugerem que esses resultados diferentes do esperado, quando levado em consideração o fato de o rotacismo ser um fenômeno estigmatizado socialmente no PB (Bagno, 2007), e que muitas vezes as mulheres evitam ou corrigem-se mais ao uso de formas estigmatizadas, segundo as teorias sociolinguísticas (Labov, 2008), se devem ao encaixamento e à tendência dessa variável no que concerne ao papel da mulher dentro dessa determinada comunidade de fala:

A análise da correlação entre gênero/sexo e a variação linguística tem de, necessariamente, fazer referência não só ao prestígio atribuído pela comunidade às variantes linguísticas como também à forma de organização social de uma dada comunidade de fala (Paiva, 2013, p. 35).

Em contrapartida, os resultados encontrados na variedade paulista de São Miguel Arcanjo, realizada por Espírito Santo (2019), aponta que os homens realizaram mais rotacismo em suas falas do que as mulheres. Esse resultado também condiz com o encontrado por Bugai (2021) para a variedade de São José do Rio Preto - SP.

Sendo assim, essa observação sugere uma influência social na adoção de determinadas variantes linguísticas e destaca o papel das normas e percepções sociais na configuração dos padrões linguísticos; todavia, para detectar estigma ou prestígio, é necessária a condução de pesquisa de avaliação subjetiva das formas alternantes (percepção), a qual não é o foco da presente pesquisa.

4.3 ESCOLARIDADE

A terceira variável destacada como relevante, a partir dos dados estatísticos, para a realização do rotacismo foi a *escolaridade*. Seus resultados podem ser visualizados na Tabela 6:

TABELA 6 - Rotacismo em relação à *escolaridade*

	1º Ciclo do Ensino Fundamental			Ensino Superior		
	Apl./Total	%	PR	Apl./Total	%	PR
Com rotacismo	129/717	18,0	0.682	67/1198	15,6	0.388
Sem rotacismo	588/717	82,0	0.318	1131/1198	84,4	0.612

Fonte: Elaboração própria

Input: 0.025
Significância: 0.000

Como pode-se observar acima, o 1º Ciclo do Ensino Fundamental favoreceu a realização do rotacismo com PR = 0.628, enquanto o Ensino Superior a desfavoreceu, com PR = 0.388.

Esse resultado se assemelha a pesquisas anteriores, em que a escolaridade foi considerada um fator relevante para a realização do rotacismo, como os estudos de Costa (2006), Tem Tem (2010) e Cristino (2021), que abordaram as variedades linguísticas de São José do Norte (RS), Rio de Janeiro (RJ) e região oeste do Paraná, respectivamente. Todos esses estudos chegaram à mesma conclusão: o fenômeno do rotacismo é mais comum na fala dos indivíduos com menor nível de escolaridade.

Portanto, mais uma vez, existe um indício de que o processo do rotacismo é socialmente estigmatizado na variedade estudada e em outras variedades do PB, visto que, segundo Bagno (2007, p. 73), “[...] para os falantes urbanos escolarizados, pronúncias como *broco*, *ingrês*, *chicrete*, *pranta*, etc. são feias, erradas e toscas”.

É de senso comum saber que os indivíduos com maior escolaridade tendem a ser mais expostos à norma padrão da língua, tanto no ambiente educacional quanto em contextos profissionais e sociais. Essa exposição contínua pode influenciar a adoção de formas linguísticas consideradas padrão, que muitas vezes são valorizadas pelas redes sociais que frequentam esses ambientes mais formais (como universidades e locais de trabalho). Outra questão, sobre a maior escolaridade desfavorecer o processo do rotacismo, pode estar associada ao acesso a recursos disponíveis nas escolas e no nível superior, como livros, materiais didáticos e programas educacionais que acabam reforçando o uso da norma padrão, e também podem ajudar a desenvolver nesses indivíduos uma consciência linguística sobre a forma como falam, tornando-os mais propensos à autocorreção. Assim, essas influências trabalham juntas para reduzir a frequência de rotacismo em contextos de maior escolaridade.

4.4 FAIXA ETÁRIA

A quarta variável selecionada como relevante para a realização do rotacismo foi a *faixa etária*, cujos resultados são expostos na tabela 7 a seguir:

TABELA 7 - Rotacismo em relação à *faixa etária*

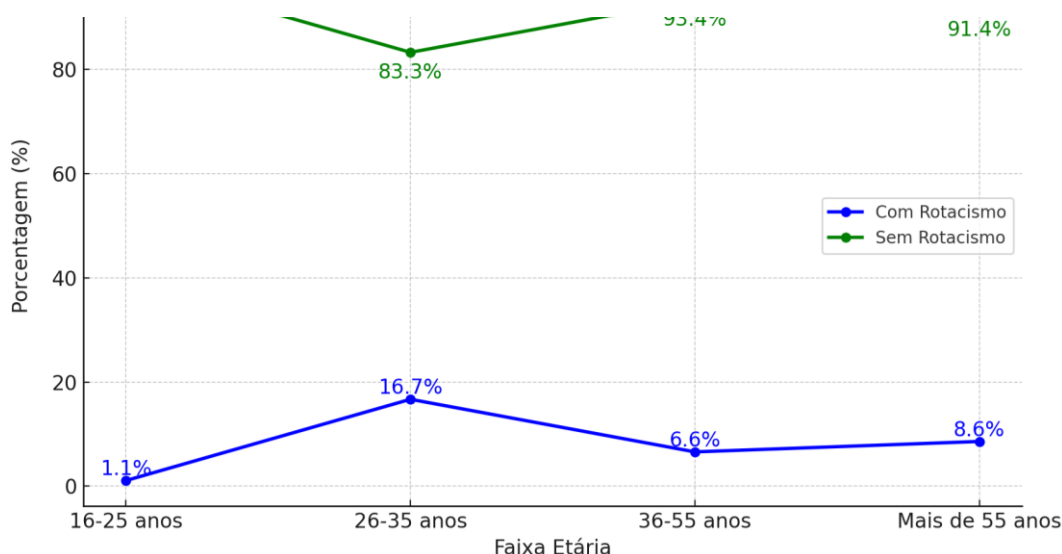
	16-25 anos			26-35 anos			36-55 anos			Mais de 55 anos		
	Apl./Total	%	PR	Apl./Total	%	PR	Apl./Total	%	PR	Apl./Total	%	PR
Com rotacismo	3/274	1,1	0.094	121/725	16,7	0.604	22/333	6,6	0.553	50/583	8,6	0.602
Sem rotacismo	271/274	98,9	0.906	604/725	83,3	0.396	311/333	93,4	0.447	533/583	91,4	0.398

Fonte: Elaboração própria

Input: 0.025
Significância: 0.000

Constata-se que o processo é relativamente favorecido, em ordem decrescente, pelas seguintes faixas etárias: 26 a 35 anos, com PR = 0.604, acima de 55 anos, com PR = 0.602, e 36 a 55 anos, com PR = 0.553. No entanto, os mais jovens (16 a 25 anos) demonstram desfavorecer a ocorrência do rotacismo, com PR = 0.094. Para entender melhor esses resultados, faz-se a observação de um gráfico curvilíneo com as porcentagens de aplicação ou não do rotacismo por faixa etária:

GRÁFICO 2 - Percentual de rotacismo em relação à faixa etária



Fonte: Elaboração própria.

Especificamente, no grupo de 16-25 anos, a frequência de rotacismo é de apenas 1,1%, indicando uma prática linguística menos comum entre os mais jovens. No entanto, observamos um aumento na faixa etária de 26-35 anos, em que a frequência de rotacismo sobe para 16,7%. Em seguida, há uma redução notável na faixa etária de 36-55 anos, com a taxa de rotacismo caindo para 6,6%. Finalmente, no grupo de mais de 55 anos, a frequência de rotacismo aumenta novamente para 8,6%.

Se considerado que os jovens de 16-25 anos têm a menor frequência de rotacismo, isso pode indicar o início de uma mudança linguística em direção à redução do rotacismo. O aumento abrupto na faixa etária de 26-35 anos pode refletir uma resistência a essa mudança ou uma falta de exposição às influências linguísticas que afetam os mais jovens.

Por outro lado, o padrão observado pode ser uma variação estável se considerarmos que as flutuações entre as faixas etárias não seguem um padrão linear de mudança, mas sim refletem uma coexistência de formas linguísticas. A variação na frequência de rotacismo pode ser atribuída a diferentes fatores sociais, econômicos e culturais que afetam cada grupo etário de maneira distinta.

Assim, para se chegar a uma conclusão, seria ideal realizar uma análise longitudinal, ou seja, acompanhar esses padrões ao longo do tempo para averiguar se o rotacismo poderia desaparecer ou não para as faixas etárias mais elevadas nessa variedade.

Por fim, cabe acrescentar, a respeito do favorecimento da aplicação do rotacismo pela faixa etária *superior a 55 anos*, que esse resultado se repete para a variedade de São José do Rio Preto, pois, anteriormente, no trabalho realizado por Bugai (2021), as faixas etárias intermediárias não foram investigadas (26-35 anos e 36-55 anos) e os informantes com a faixa etária de mais de 55 anos favoreceram a realização do rotacismo, com $PR = 0.671$, ao passo que a faixa-etária de 16 a 25 anos a desfavoreceu, com $PR = 0.180$ (Bugai, 2021). Ademais, em relação a outras variedades do PB, o estudo de Costa (2006), que investigou a variedade gaúcha de São José do Norte, também obteve o resultado que apontou o favorecimento do rotacismo na faixa etária mais elevada (mais de 55 anos de idade), com $PR = 0.70$.

4.5 CONTEXTO PRECEDENTE

Como último grupo de fatores destacado como favorecedor da aplicação do rotacismo, tem-se o *contexto precedente*. Os resultados são apresentados na tabela seguinte:

TABELA 8 - Rotacismo em relação ao *contexto precedente*

	Fricativas			Oclusivas			Vogal Anterior			Vogal Posterior		
	Apl./Total	%	PR	Apl./Total	%	PR	Apl./Total	%	PR	Apl./Total	%	PR
Com rotacismo	1/17	5,9	0.209	89/356	25,0	0.606	40/1136	3,5	0.389	66/406	16,3	0.720
Sem rotacismo	16/17	94,1	0.791	267/356	75,0	0.394	1096/1136	96,5	0.611	340/406	83,7	0.280

Fonte: Elaboração própria.

Input: 0.025
Significância: 0.000

A partir da análise da Tabela 8, pode-se observar que o rotacismo é favorecido quando a consoante-alvo /l/ é precedida uma vogal posterior, como em *cu[l].tu.ra*, com $PR = 0.720$, e também quando precedida por uma oclusiva, como em *di.p[r]o.ma*, com $PR = 0.606$. No entanto, deve-se notar que, das 66 ocorrências de rotacismo precedidas por uma vogal posterior, 40 (60,6%) estão presentes em apenas três paradigmas verbais: *re.so[l].ver*, *so[l].tar* e *vo[l].tar*, conforme exposto na tabela 9:

TABELA 9 - Ocorrências de rotacismo com vogal posterior precedente

Vocábulo	Ocorrências
<i>Bulldog</i>	1
<i>Culpa</i>	2
<i>Cultura</i>	1
<i>Dificuldade(s)</i>	5
<i>Dissolver</i>	1
<i>Empolgar</i>	1
<i>Golpe</i>	3
<i>Insultar</i>	3
<i>Mirassol</i>	2
<i>Multar</i>	1
<i>Pulsção</i>	1
<i>Resolver</i>	8
<i>Revoltar</i>	2
<i>Revólver</i>	2
<i>Soltar</i>	14
<i>Ultimamente</i>	1
<i>Voltar</i>	18

Fonte: Elaboração própria.

Assim, esse resultado pode estar atrelado ao fato de essas palavras serem recorrentes na amostra.

Ainda em relação ao contexto precedente, os resultados quantitativos revelam que o rotacismo é desfavorecido quando a consoante-alvo segue uma vogal anterior, como em *fi[l].mar*, com PR = 0.389, e especialmente quando segue uma consoante fricativa, como em *fi[l]or*, com PR = 0.209.

A maior ocorrência do rotacismo em oclusivas e a baixa ocorrência em fricativas como contexto precedente também foram identificadas nos trabalhos de Costa (2006), Bugai (2021) e Tem Tem (2010). Tem Tem (2010) explica o favorecimento de alguns contextos com base na Teoria dos Traços Distintivos,²¹ pela qual os traços oclusivos deveriam favorecer o contexto de rotacização: “as oclusivas favorecem a rotacização, pois têm pouca semelhança com o /l/ e são mais próximas do /r/” (Tem Tem, 2010, p. 72), ou seja, as oclusivas e o tepe têm uma semelhança no ponto de vista articulatorio,

²¹ Resumo do sistema de Jakobson, Fant e Halle (1963, *apud* Tem Tem, 2010, p. 72). Vocálico: vogais e líquidas; Não-vocálico: os *glides* e as obstruentes; Consonântico: obstruentes e líquidas; Não-consonântico: vogais e *glides*; Contínuo: fricativas e laterais; Não-contínuo: oclusivas, vibrantes e africadas.

pois ambos apresentam uma liberação abrupta do fluxo de ar, tornando-os sons não-contínuos. Por outro lado, o que as fricativas e o /l/ têm em comum, em termos de modo de articulação, é que ambos liberam o fluxo de ar sem interrupção, tornando-os sons contínuos: “as fricativas podem favorecer o uso da forma padrão, pela proximidade com o /l/, ou mesmo tenderem ao apagamento do /l/” (Tem Tem, 2010, p. 72-73).

Neste capítulo, foram analisados os resultados do fenômeno do rotacismo na variedade linguística do noroeste paulista. Os fatores destacados pelo *Goldvarb X* como favorecedores do rotacismo foram, em ordem de relevância: posição na sílaba, sexo/gênero, escolaridade, faixa etária e contexto precedente. A posição na sílaba se mostrou o fator mais relevante, com o ataque complexo sendo o ambiente que mais favorece o rotacismo. Em seguida, foi observado que o sexo/gênero masculino favorece mais o fenômeno do que o feminino. Quanto à escolaridade, indivíduos com menor nível de instrução apresentaram maior frequência de rotacismo, sugerindo um indício de estigma social associado ao processo.

No que tange à faixa etária, enquanto os mais jovens (16-25 anos) quase não realizam rotacismo, os falantes de 26-35 anos mostram um aumento significativo na aplicação, com uma redução nas faixas seguintes, exceto nos indivíduos acima de 55 anos, que voltam a favorecer o fenômeno. O contexto precedente, como fricativas e vogais, também desempenha um papel importante na realização do rotacismo.

Esses resultados apontam que o rotacismo na região investigada é influenciado tanto por fatores linguísticos quanto extralinguísticos, refletindo tendências observadas em outras variedades do PB, como as de São José do Norte (Costa, 2006) e o oeste do Paraná (Cristino, 2021). Além disso, os dados sugerem que o rotacismo, especialmente entre os mais jovens, possa estar em declínio, o que indica uma possível mudança linguística em andamento.

Feitas as análises quantitativa e qualitativa dos dados obtidos, passa-se agora para as considerações finais.

5 CONSIDERAÇÕES FINAIS

O presente estudo examinou a variação do rotacismo na região do noroeste paulista, com base em dezesseis entrevistas sociolinguísticas providas do banco de dados Iboruna (Gonçalves, 2024) e analisadas à luz da Sociolinguística quantitativa.

De maneira geral, o rotacismo demonstrou ser pouco frequente na variedade analisada, ocorrendo em apenas 10,2% dos casos, portanto confirmando-se como uma regra variável para essa variedade, mas com pouca aplicação na variedade em estudo.

O processo do rotacismo foi favorecido por um conjunto de fatores linguísticos e extralinguísticos, incluindo: (i) posição da sílaba em ataque complexo, como em *c[r]í.ni.ca*, e em coda medial, como em *di.fi.cu[ɹ].da.de*; (ii) sexo/gênero (masculino); (iii) menor nível de escolaridade (1º ciclo do Ensino Fundamental); (iv) faixas etárias mais avançadas (26 a 35 anos, 35 a 55 anos e acima de 55 anos); e (v) contexto precedente preenchido por uma consoante oclusiva, como em *in.g[r]ês*.

O fato de o rotacismo ser favorecido no ambiente de ataque complexo confirma a hipótese inicial de que esse contexto linguístico seria favorecedor da aplicação da regra variável, visto já haver resultados prévios semelhantes nos estudos de Bugai (2021) e Costa (2006).

O resultado relacionado ao fator extralinguístico *escolaridade* sugere um indício de estigma social associado ao processo do rotacismo na região do noroeste de São Paulo, uma vez que a menor escolaridade mostrou-se favorecer a ocorrência do fenômeno, o que está de acordo com o que é proposto nas literaturas sociolinguísticas (Labov, 2008). Esse resultado confirma a hipótese inicial de que o fenômeno do rotacismo poderia ser influenciado por fatores sociais, pelo fato de essa variedade ter maior ocorrência do fenômeno na fala do sexo/gênero masculino e do nível de escolaridade do 1º ciclo do ensino fundamental.

Em relação à aplicação do processo, a variedade analisada apresenta semelhanças com outras variedades do PB, especialmente em relação à influência do contexto silábico (Costa, 2006; Cristino, 2021) e ao favorecimento do rotacismo por parte de falantes com menor nível de escolaridade (Costa, 2006; Tem Tem, 2010; Cristino, 2021).

Assim, os objetivos propostos inicialmente foram alcançados, pois, por meio desta investigação, foi possível obter indícios de variação e mudança linguística na variedade de São José do Rio Preto – SP em relação ao rotacismo, o que foi averiguado por meio das faixas etárias. Porém, para uma confirmação conclusiva do *status* de variação do rotacismo nessa variedade, sugere-se um estudo longitudinal. Foi possível, também, observar como as variáveis sociais *sexo/gênero* e *escolaridade* foram relevantes e influenciaram a ocorrência do rotacismo para a variedade investigada.

Por fim, espera-se que os resultados desta pesquisa possam contribuir para o mapeamento sociolinguístico do Português falado no noroeste paulista e, de forma mais ampla, do PB.

REFERÊNCIAS

- AMARAL, Amadeu. **O dialeto caipira**: gramática, vocabulário. 4. ed. São Paulo: HUCITEC; Brasília: INL, 1981.
- BAGNO, Marcos. **Nada na língua é por acaso**: por uma pedagogia da variação linguística. São Paulo: Parábola, 2007.
- BARROZO, Thais Aranda; AGUILERA, Vanderci de Andrade. Sexo e linguagem: uma análise a partir das sabatinas dos ministros do Supremo Tribunal Federal Joaquim Barbosa e Rosa Weber. **Revista da ABRALIN**, v. 13, n. 1, 2014.
- BOTELHO, José Mario; LEITE, Isabelle Lins. Metaplasmos contemporâneos—um estudo acerca das atuais transformações fonéticas da Língua Portuguesa. *In: II Congresso de Letras da UERJ – São Gonçalo (II CLUERJ-SG)*. 2005. p. 1-12.
- BUGAI, Tayná Maria Coelho. **A aplicação 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.
- CÂMARA JR., Joaquim Mattoso. **Estrutura da língua portuguesa**. 7. ed., Petrópolis, RJ: Editora Vozes, 1976.
- CARMO, Márcia Cristina do. **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, Márcia Cristina do. **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, Márcia Cristina do. 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, Márcia Cristina do. Variação linguística das vogais médias pretônicas em contexto medial no noroeste paulista. **UniLetras**, v. 40, p. 221-239, 2018.
- CARMO, Márcia Cristina do. Alçamento vocálico das vogais médias pretônicas iniciais na variedade do noroeste paulista. **Estudos Linguísticos** (São Paulo. 1978), v. 48, p. 800-821, 2019.
- CARMO, Márcia Cristina do; CARLOS, Valeska Gracioso. 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, Márcia Cristina do; MACHADO, Talia Ferreira. Monotongação variável do ditongo oral decrescente [ej] no noroeste paulista. **Working Papers em Linguística**, v. 24, p. 160-184, 2023.

CARMO, Márcia Cristina do; TABORDA, Isabela Ribeiro. 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, Márcia Cristina do; TENANI, Luciani Ester. As vogais médias pretônicas na variedade do noroeste paulista: uma análise sociolinguística. **Alfa: Revista de Linguística** (UNESP. Online), v. 57, p. 607-637, 2013.

COELHO, Izete Lehmkuhl et al. **Para conhecer sociolinguística**. São Paulo: Contexto, 2015.

COLLISCHONN, Gisela. **Fonologia do português brasileiro, da sílaba à frase**. Instituto de Letras – UFRGS, 2006. p. 34-45. Universidade Federal do Rio Grande do Sul.

COSTA, Luciane Trennephol. **Estudo do Rotacismo**: variação entre as consoantes líquidas. 166 f. 2006. Dissertação (Mestrado em Letras). Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.

COSTA, Luciane Trennephol. Análise variacionista do rotacismo. **Revista Virtual de Estudos da Linguagem**. v. 5, n. 9, ago. 2007, p. 1-29.

COX, Maria Inês Pagliarini; ASSAD, Cader Faisal. O ele e o erre só trazem 'compricação'- um estudo das representações de /l/ e /r/ na escrita de crianças em processo de alfabetização. **Revista de Educação Pública**, Cuiabá, v. 8, n.13, p. 143-156, 1999.

CRISTINO, Tathiane; BUSSE, Sanimar. Rotacismo: estudo do fenômeno na fala do Oeste do Paraná. In: JORNADA DE ESTUDOS LINGUÍSTICOS E LITERÁRIOS – JELL, 19., 2017. **Anais [...]**. Marechal Cândido Rondon: JELL, 2017.

CRISTINO, Tathiane. **Descrição do rotacismo e do retroflexo em localidades do Oeste paranaense**. 105 f. 2021. Dissertação (Mestrado em Letras). Universidade Estadual do Oeste do Paraná, Cascavel, 2021.

DESENVOLVE SP. **Região administrativa**: São José do Rio Preto. Disponível em: <https://www.desenvolvesp.com.br/mapadaeconomiaapaulista/ra/sao-jose-do-rio-preto/>. Acesso em: 8 jul. 2024.

DIAS, Jheniffer Amanda; CARMO, Márcia Cristina do. A metátese na variedade do interior paulista. **Muitas Vozes**, v. 10, p. 1-23, 2021.

DICKEY, Laura Walsh. **The Phonology of Liquids**. Tese de Doutorado. Massachusetts: Amherst, 1997.

ESPÍRITO SANTO, Júlia Maria França. **Entre o campo e a cidade: rotacismo em São Miguel Arcanjo**. 116 f. 2019. Dissertação (Mestrado). Universidade de São Paulo, São Paulo, 2019.

FERREIRA, Jesuelem Salvani. **O apagamento do /d/ em morfema de gerúndio no dialeto de São José do Rio Preto**. 145 f. 2010. Dissertação (Mestrado em Estudos Linguísticos). UNESP, São José do Rio Preto, 2010.

FRANCISCO, Arlete Maria. Contribuição à história da urbanização de São José do Rio Preto-SP. **Revista Tópos**, v. 5, n. 1, p. 119-142, 2011.

FREITAG, Raquel Meister Ko et al. Vamos prantar froes no grobo da terra: estudando o rotacismo nas séries iniciais da rede municipal de ensino de Moita Bonita/SE. **Revista Virtual de Letras**, v. 2, n. 2, p. 17-31, 2010.

FREITAG, Raquel Meister Ko. (Re)Discutindo Sexo/Gênero na Sociolinguística. *In*: FREITAG, Raquel Meister Ko; SEVERO, Cristine Gorski (Org). **Mulheres, Linguagem e Poder: Estudos de Gênero na Sociolinguística Brasileira**. São Paulo: Blucher, 2015, p. 17-74.

FUNDAÇÃO SEADE. **Perfil dos Municípios Paulistas**. São Paulo: Fundação Sistema Estadual de Análise de Dados, Disponível em: <https://perfil.seade.gov.br/>. Acesso em: 15 mai. 2025.

FUNDAÇÃO SEADE. **Histórico**. Disponível em: http://www.perfil.seade.gov.br/historico/hist_498.pdf Acesso em: 8 jul. 2023.

FUNDAÇÃO SEADE. **IMP - Informações dos Municípios Paulistas**. Disponível em: <http://www.perfil.seade.gov.br/>. Acesso em: 8 jul. 2023.

GONÇALVES, Sebastião Carlos Leite. Projeto ALIP (Amostra Linguística do Interior Paulista) e banco de dados Iboruna: 10 anos de contribuição com a descrição do português brasileiro. **Estudos Linguísticos (São Paulo. 1978)**, v. 48, n. 1, p. 276-297, 2019.

GONÇALVES, Sebastião Carlos Leite. **Banco de dados Iboruna: amostras eletrônicas do português falado no interior paulista**. 2007. Disponível em: <http://www.alip.ibilce.unesp.br>. Acesso em: 8 jul. 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **São José do Rio Preto – Histórico**. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/dtbs/saopaulo/saojosedoriopreto.pdf> Acesso em: 6 jul. 2023.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **São José do Rio Preto – Panorama**. Disponível em: <https://cidades.ibge.gov.br/brasil/sp/sao-jose-do-rio-preto/panorama>. Acesso em: 8 jul. 2024.

JÚNIOR, Marcos Elias de Oliveira. **São José do Rio Preto e municípios limítrofes**. 2011. Mapa. Disponível em:

https://commons.wikimedia.org/wiki/File:São_José_do_Rio_Preto_e_municípios_limítrofes.svg. Acesso em: 15 maio 2025.

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

LABOV, William. Some sociolinguistic principles. *In*: PAULSTON, C. B.; TUCKER, G. R. (Eds.). **Sociolinguistics: the essential readings**. Oxford: Blackwell, 2003. p. 235-250.

LABOV, William. Where does the linguistic variable stop? **Sociolinguistic Working Papers**. Austin, Texas, n. 44, abril, 1978.

LAVANDERA, Beatriz Rosario. Where does the sociolinguistic variable stop? **Language and society**, v. 7, p. 171-182, 1977.

PAIVA, Maria da Conceição de. A variável gênero/sexo. *In*: MOLLICA, Maria Cecilia e BRAGA, Maria Luiza (orgs.). 4. ed. **Introdução à Sociolinguística: o tratamento da variação**. São Paulo: Contexto, 2013, p. 33-42.

PAIVA, Maria da Conceição de; DUARTE, Maria Eugênia Lamóglia. Mudança Linguística: observações no tempo real. *In*: MOLLICA, Cecília; BRAGA, Maria Luiza (Org.). **Introdução à sociolinguística: o tratamento da variação**. 4. ed. São Paulo: Contexto. p. 179-190. 2015.

PREFEITURA DE SÃO JOSÉ DO RIO PRETO. **Sobre Rio Preto**. Disponível em: <https://www.riopreto.sp.gov.br/sobre/> Acesso em: 8 jul. 2023.

PREFEITURA DE SÃO JOSÉ DO RIO PRETO. Secretaria Municipal de Planejamento Estratégico, Ciência, Tecnologia e Inovação. **Conjuntura Econômica de São José do Rio Preto** / coordenada por Gisele Madi de Freitas. 35. ed. São José do Rio Preto: Secretaria Municipal de Planejamento Estratégico, Ciência, Tecnologia e Inovação, 2020. 107p.: il., [11] fots., mapas color., gráfs.; 21 x 29 cm. Disponível em: <http://www.riopreto.sp.gov.br> Acesso em: 8 jul. 2023. Bibliografia: p. 105-107.

PRODELIK, André Izaias Garus; CARMO, Márcia Cristina do. Ditongação diante de /S/ no interior paulista: análise sociolinguística. *In*: SANTOS, Giuliano Lellis Ito; OLIVEIRA, Jane Kelly de; LEGROSKI, Marina Chiara. (Org.). **Ensino, pesquisa e formação: a linguística no curso de Letras da UEPG**. 1. ed. Curitiba: PEDEX, 2024, p. 85-115.

RAMOS, Adriana Perpétua. **Descrição das vogais postônicas não-finais na variedade do noroeste Paulista**. 177 f. 2009. Dissertação (Mestrado em Estudos Linguísticos) UNESP, São José do Rio Preto, 2009.

REZENDE, Brenda Soares; CARMO, Márcia Cristina do. Análise sociolinguística da monotongação de [ow] na variedade do interior paulista. *In*: SANTOS, Giuliano Lellis Ito; LEGROSKI, Marina. (Org.). **Letras graduadas: trabalhos dos alunos do curso de Letras da Universidade Estadual de Ponta Grossa**. 1. ed. Ponta Grossa: 2023, p. 98-135.

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

SAUSSURE, Ferdinand de. **Curso de linguística geral**. 27. ed. São Paulo: Cultrix, 2006 [1916].

SILVA, Thaís Cristófar. **Dicionário de fonética e fonologia**. São Paulo: Contexto, 2011.

SILVA NETO, Serafim da. **Ensaio de filologia portuguesa**. São Paulo: Nacional, 1956.

TAGLIAMONTE, Sali A. **GoldVarb**. Disponível em: <http://individual.utoronto.ca/tagliamonte/goldvarb.html>. Acesso em: 15 maio 2025.

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

TEM TEM, Luiza Fernandes. **Rotacização das líquidas nos grupos consonantais: representação fonológica e variação**, 2010. 156 f. Dissertação (Mestrado em Letras Vernáculas). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

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

APÊNDICE A – LISTA DE VOCÁBULOS COM CONTEXTO PARA A REALIZAÇÃO DO ROTACISMO

Neste apêndice, são listados os vocábulos levantados e analisados nesta pesquisa, seguidos pelo número de ocorrências de rotacismo e do total de dados.

Vocábulo	Ocorrências com rotacismo	Total de ocorrências
<i>Aconselhável</i>	0	2
<i>Adail</i>	0	1
<i>Adulto</i>	0	1
<i>Afinal</i>	0	3
<i>Agradável</i>	0	1
<i>Agricultura</i>	0	2
<i>Al dente</i>	0	3
<i>Alberto</i>	0	2
<i>ALCA</i>	1	11
<i>Álcool</i>	0	4
<i>Álcool</i>	0	4
<i>Alcólica</i>	0	1
<i>Alfabetização</i>	0	10
<i>Alfabetizada</i>	0	1
<i>Alfabetizados</i>	0	1
<i>Alfabetizar</i>	0	1
<i>Alfabetizaram</i>	0	2
<i>Alfabeto</i>	0	3
<i>Alfabetos</i>	0	2
<i>Alface</i>	0	2
<i>Alfaataria</i>	0	10
<i>Alfaate</i>	0	14
<i>Alfaates</i>	0	3
<i>Algemo</i>	0	1
<i>Algo</i>	0	3
<i>Algodão</i>	5	5
<i>Alguém</i>	3	20
<i>Algum</i>	4	21
<i>Alguma</i>	1	31
<i>Algumas</i>	0	7
<i>Alguns</i>	0	20
<i>Almeirão</i>	0	1
<i>Almoça</i>	0	1
<i>Almoçamos</i>	0	1
<i>Almoçar</i>	0	4
<i>Almocei</i>	0	1
<i>Almoço</i>	0	7
<i>Almoçou</i>	0	1
<i>Almofadas</i>	0	3

<i>Almoxarifado</i>	0	5
<i>Alqueire</i>	3	4
<i>Alta</i>	0	5
<i>Altas</i>	0	1
<i>Alto</i>	1	5
<i>Altura</i>	0	5
<i>Aluguel</i>	0	3
<i>Alvorçados</i>	0	1
<i>Ampla</i>	0	3
<i>Animal</i>	0	8
<i>Anual</i>	0	2
<i>Aplica</i>	1	1
<i>Aplicação</i>	1	1
<i>Aplicava</i>	0	2
<i>Araldite</i>	0	1
<i>Artesanal</i>	0	1
<i>Asfaltado</i>	0	2
<i>Asfalto</i>	0	1
<i>Assaltada</i>	0	1
<i>Assaltado</i>	0	1
<i>Assalto</i>	0	2
<i>Assembleia</i>	0	1
<i>Atual</i>	0	1
<i>Atualmente</i>	0	1
<i>Automóvel</i>	0	4
<i>Azul</i>	0	6
<i>Balcão</i>	1	1
<i>Bíblia</i>	0	3
<i>Biblioteca</i>	0	6
<i>Bicicleta</i>	1	3
<i>Blazer</i>	0	2
<i>Bloco</i>	0	5
<i>Blusa</i>	0	3
<i>Bolsa</i>	0	2
<i>Bolsas</i>	0	1
<i>Bolsinha</i>	0	1
<i>Bolso</i>	0	1
<i>Bombril</i>	0	1
<i>Brasil</i>	0	29
<i>Bulldog</i>	1	1
<i>Cafeicultores</i>	0	1
<i>Cafezal</i>	0	1
<i>Calça</i>	0	9
<i>Calçadas</i>	0	2
<i>Calças</i>	0	1
<i>Cálcio</i>	0	1
<i>Calda</i>	0	1
<i>Caldo</i>	0	17

<i>Calminho</i>	0	1
<i>Calmo</i>	0	1
<i>Carnaval</i>	0	6
<i>Casal</i>	0	8
<i>Casalzinho</i>	0	1
<i>Cascavel</i>	0	5
<i>Celso</i>	0	1
<i>Central</i>	0	1
<i>Ciclano</i>	1	1
<i>Ciclo</i>	0	3
<i>Ciclos</i>	0	1
<i>Clara</i>	0	13
<i>Clareando</i>	0	1
<i>Clarear</i>	0	2
<i>Clareia</i>	0	2
<i>Claro</i>	0	5
<i>Classe</i>	0	14
<i>Clepto</i>	0	1
<i>Clima</i>	0	1
<i>Clínica</i>	2	2
<i>Clube</i>	0	3
<i>Cobalto</i>	0	3
<i>Colcha</i>	0	1
<i>Colchão</i>	0	3
<i>Colegial</i>	0	3
<i>Colesterol</i>	0	1
<i>Comercial</i>	1	1
<i>Completa</i>	0	1
<i>Completamente</i>	1	4
<i>Completei</i>	0	1
<i>Completo</i>	1	2
<i>Complexas</i>	0	1
<i>Complexo</i>	0	2
<i>Complicada</i>	2	2
<i>Complicado</i>	2	5
<i>Conflito</i>	0	1
<i>Consulta</i>	0	1
<i>Consultora</i>	0	2
<i>Consultou</i>	0	1
<i>Coral</i>	0	1
<i>Coronel</i>	6	6
<i>Culpa</i>	2	4
<i>Culpado</i>	0	4
<i>Culpados</i>	0	1
<i>Cultura</i>	1	2
<i>Cultural</i>	0	2
<i>Cultural</i>	0	2
<i>Desagradável</i>	0	1

<i>Descalça</i>	0	2
<i>Descalço</i>	1	2
<i>Descartável</i>	0	1
<i>Desculpa</i>	0	4
<i>Desenvolve</i>	0	1
<i>Desenvolvendo</i>	0	5
<i>Desenvolveu</i>	0	1
<i>Desenvolviam</i>	0	1
<i>Desenvolvimento</i>	0	3
<i>Desfalque</i>	0	1
<i>Diagonal</i>	0	1
<i>Diesel</i>	0	2
<i>Difícil</i>	0	37
<i>Difícilmente</i>	0	1
<i>Dificuldade</i>	4	10
<i>Dificuldades</i>	1	4
<i>Diploma</i>	3	3
<i>Dissolve</i>	0	6
<i>Dissolvendo</i>	0	1
<i>Dissolver</i>	0	1
<i>Dissolveu</i>	1	2
<i>Dissolvi</i>	0	1
<i>Dissolvo</i>	0	3
<i>Divulgado</i>	0	1
<i>Dócil</i>	0	1
<i>Dribles</i>	0	1
<i>Duplicata</i>	0	1
<i>Duplo</i>	0	1
<i>Educacional</i>	0	1
<i>Elmaz</i>	0	1
<i>Empolgação</i>	0	1
<i>Empolgada</i>	0	2
<i>Empolgaram</i>	1	1
<i>Engloba</i>	0	1
<i>Envolve</i>	0	2
<i>Envolveu</i>	0	1
<i>Enxovalzinho</i>	0	1
<i>Especial</i>	0	2
<i>Espiritual</i>	0	1
<i>Essencial</i>	0	5
<i>Estadual</i>	0	13
<i>Exclusiva</i>	0	1
<i>Exclusivo</i>	0	2
<i>Exemplo</i>	9	43
<i>Explicação</i>	1	2
<i>Explicando</i>	0	1
<i>Explicar</i>	0	3
<i>Explicava</i>	0	1

<i>Explicou</i>	0	3
<i>Fácil</i>	0	23
<i>Faculdade</i>	0	27
<i>Faculdades</i>	0	2
<i>Falsidade</i>	0	1
<i>Falta</i>	0	24
<i>Faltando</i>	0	1
<i>Faltar</i>	1	3
<i>Faltava</i>	0	3
<i>Farol</i>	0	2
<i>Federal</i>	1	2
<i>Filmar</i>	0	2
<i>Filme</i>	0	4
<i>Filminho</i>	0	5
<i>Filmou</i>	0	1
<i>Final</i>	0	11
<i>Flanela</i>	0	1
<i>Flor</i>	0	3
<i>Flores</i>	0	4
<i>Floresta</i>	0	1
<i>Florido</i>	0	1
<i>Fundamental</i>	0	6
<i>Funil</i>	0	1
<i>Futebol</i>	0	6
<i>Garibaldi</i>	0	1
<i>General</i>	0	1
<i>Geral</i>	0	3
<i>Geralmente</i>	1	36
<i>Ginasial</i>	0	2
<i>Glicério</i>	0	1
<i>Glicose</i>	0	1
<i>Globo</i>	1	4
<i>Gol</i>	0	4
<i>Golfinho</i>	0	1
<i>Golpe</i>	3	4
<i>Homossexual</i>	0	1
<i>Horrível</i>	0	1
<i>Hospital</i>	1	14
<i>Humilde</i>	0	2
<i>Ideal</i>	0	1
<i>Igual</i>	0	17
<i>Igualmente</i>	0	1
<i>Imóvel</i>	0	1
<i>Implica</i>	2	2
<i>Implicam</i>	1	1
<i>Impulsão</i>	0	2
<i>Impulsividade</i>	0	1
<i>Impulso</i>	0	4

<i>Inclinada</i>	0	3
<i>Inclinadona</i>	0	1
<i>Inclusive</i>	0	22
<i>Indisciplinado</i>	0	1
<i>Industrial</i>	0	1
<i>Infiltração</i>	0	1
<i>Influência</i>	0	2
<i>Influir</i>	0	1
<i>Inglês</i>	2	2
<i>Inglesas</i>	1	1
<i>Insultando</i>	1	1
<i>Insultou</i>	2	2
<i>Intercontinental</i>	0	1
<i>Internacional</i>	0	2
<i>Jornal</i>	0	10
<i>Julgar</i>	0	1
<i>Lateral</i>	0	1
<i>Legal</i>	0	23
<i>Lençol</i>	0	2
<i>Litoral</i>	0	1
<i>Local</i>	0	6
<i>Mal</i>	0	23
<i>Malcriada</i>	0	3
<i>Malcriados</i>	0	1
<i>Maltratava</i>	0	1
<i>Manual</i>	0	2
<i>Maraldi</i>	0	1
<i>Marechal</i>	0	6
<i>Marginal</i>	0	1
<i>Material</i>	0	7
<i>Mel</i>	0	2
<i>Mental</i>	0	2
<i>Mentalmente</i>	1	1
<i>Mil</i>	0	47
<i>Mirassol</i>	2	73
<i>Moldes</i>	0	1
<i>Motel</i>	0	1
<i>Multar</i>	1	1
<i>Mundial</i>	0	1
<i>Municipal</i>	0	1
<i>Nacional</i>	0	2
<i>Natal</i>	0	14
<i>Nestlé</i>	0	1
<i>Nível</i>	0	3
<i>Normal</i>	0	31
<i>Normalmente</i>	0	5
<i>Oficial</i>	0	9
<i>Painel</i>	0	8

<i>Palco</i>	0	4
<i>Palma</i>	0	3
<i>Palpite</i>	0	1
<i>Panfleteo</i>	0	1
<i>Panfletos</i>	0	1
<i>Papel</i>	0	7
<i>Paroquial</i>	0	1
<i>Pênalti</i>	0	15
<i>Perfil</i>	0	1
<i>Pernil</i>	0	1
<i>Pessoal</i>	0	33
<i>Placa</i>	2	3
<i>Plana</i>	0	2
<i>Planejado</i>	0	1
<i>Planejou</i>	1	1
<i>Planeta</i>	0	3
<i>Planificada</i>	0	1
<i>Planificou</i>	0	1
<i>Plano</i>	0	3
<i>Planta</i>	12	16
<i>Plantação</i>	4	7
<i>Plantado</i>	1	1
<i>Plantam</i>	0	2
<i>Plantando</i>	2	2
<i>Plantão</i>	0	1
<i>Plantar</i>	2	14
<i>Plantas</i>	0	1
<i>Plantasse</i>	0	1
<i>Plantava</i>	4	5
<i>Plantei</i>	3	3
<i>Plantio</i>	1	1
<i>Plantou</i>	6	7
<i>Plástica</i>	0	1
<i>Plástico</i>	0	2
<i>Plásticos</i>	0	1
<i>Plastiquinho</i>	0	1
<i>Policia</i>	0	2
<i>Portugal</i>	0	5
<i>Possível</i>	0	3
<i>Principal</i>	0	6
<i>Principalmente</i>	0	10
<i>Problema</i>	8	33
<i>Problemas</i>	1	7
<i>Profissional</i>	0	1
<i>Prol</i>	0	1
<i>Provavelmente</i>	0	2
<i>Pública</i>	0	1
<i>Publicação</i>	0	1

<i>Publicado</i>	0	1
<i>Público</i>	0	1
<i>Pulsção</i>	1	3
<i>Qual</i>	0	10
<i>Qualquer</i>	5	26
<i>Quintal</i>	0	8
<i>Razoavelmente</i>	0	1
<i>Real</i>	0	12
<i>Realmente</i>	0	21
<i>Reclamar</i>	3	4
<i>Reclamava</i>	3	3
<i>Reclamei</i>	0	1
<i>Reflete</i>	1	1
<i>Regional</i>	0	3
<i>Replanta</i>	0	1
<i>Republicano</i>	1	1
<i>Repulsa</i>	0	1
<i>Repulsão</i>	0	1
<i>Resolve</i>	3	4
<i>Resolvemos</i>	0	1
<i>Resolvendo</i>	0	1
<i>Resolver</i>	0	4
<i>Resolveu</i>	2	9
<i>Resolvi</i>	1	1
<i>Resolvia</i>	1	1
<i>Resolvido</i>	1	1
<i>Responsável</i>	0	6
<i>Revoltado</i>	1	1
<i>Revoltar</i>	1	1
<i>Revólver</i>	2	2
<i>Rival</i>	0	1
<i>Royal</i>	0	4
<i>Sal</i>	0	8
<i>Salgada</i>	0	1
<i>Salgadinho</i>	0	1
<i>Salgadinhos</i>	0	1
<i>Salgado</i>	0	1
<i>Salgados</i>	0	1
<i>Salsão</i>	0	3
<i>Salsaretti</i>	0	1
<i>Salsicha</i>	0	2
<i>Salvação</i>	0	3
<i>Salvo</i>	0	1
<i>Salvou</i>	0	1
<i>Sentimental</i>	0	1
<i>Sexualmente</i>	0	1
<i>Simples</i>	2	14
<i>Simplesmente</i>	1	7

<i>Sinal</i>	0	2
<i>Social</i>	0	1
<i>Socialmente</i>	0	1
<i>Sol</i>	0	9
<i>Solda</i>	0	1
<i>Soldar</i>	0	1
<i>Solta</i>	4	4
<i>Soltar</i>	4	5
<i>Soltaram</i>	2	2
<i>Soltava</i>	1	4
<i>Solteira</i>	0	2
<i>Solteiro</i>	0	2
<i>Soltinho</i>	0	4
<i>Solto</i>	0	1
<i>Soltou</i>	3	3
<i>Sul</i>	0	4
<i>Tabletinho</i>	0	1
<i>Tal</i>	1	71
<i>Talvez</i>	1	8
<i>Terminal</i>	0	2
<i>Terrível</i>	0	2
<i>Total</i>	0	2
<i>Totalmente</i>	0	8
<i>Tradicional</i>	0	1
<i>Transmissível</i>	0	1
<i>Tumulto</i>	0	1
<i>Úlcera</i>	0	3
<i>Última</i>	0	3
<i>Ultimamente</i>	1	4
<i>Último</i>	0	2
<i>Varal</i>	0	3
<i>Voleibol</i>	0	1
<i>Volta</i>	4	37
<i>Voltados</i>	0	1
<i>Voltamos</i>	0	2
<i>Voltando</i>	2	5
<i>Voltar</i>	3	23
<i>Voltaram</i>	0	1
<i>Voltas</i>	0	1
<i>Voltasse</i>	0	1
<i>Voltava</i>	4	5
<i>Voltei</i>	3	7
<i>Volto</i>	0	1
<i>Voltou</i>	3	13

ANEXO A – RODADAS DO GOLDVARB X PARA A REGRA DO ROTACISMO NA VARIEDADE DO NOROESTE PAULISTA

• CELL CREATION • 10/10/2022 15:36:33 •.....

Name of token file: rot.tkn

Name of condition file: Untitled.cnd

(

; Identity recode: All groups included as is.

(1)

(2)

(3)

(4)

(5)

(6)

(7)

(8)

)

Number of cells: 166

Application value(s): 10

Total no. of factors: 19

Group	1	0	Total	%
-------	---	---	-------	---

1 (2)	1	0		
-------	---	---	--	--

M N	183	990	1173	61.3
-----	-----	-----	------	------

%	15.6	84.4		
---	------	------	--	--

F N	13	729	742	38.7
-----	----	-----	-----	------

%	1.8	98.2		
---	-----	------	--	--

Total N	196	1719	1915	
---------	-----	------	------	--

%	10.2	89.8		
---	------	------	--	--

2 (3) 1 0
 1 N 3 271 274 14.3
 % 1.1 98.9

2 N 121 604 725 37.9
 % 16.7 83.3

3 N 22 311 333 17.4
 % 6.6 93.4

4 N 50 533 583 30.4
 % 8.6 91.4

Total N 196 1719 1915
 % 10.2 89.8

3 (4) 1 0
 F N 129 588 717 37.4
 % 18.0 82.0

S N 67 1131 1198 62.6
 % 5.6 94.4

Total N 196 1719 1915
 % 10.2 89.8

4 (5) 1 0
 F N 12 741 753 39.3
 % 1.6 98.4

A N 90 284 374 19.5

% 24.1 75.9

M N 94 694 788 41.1

% 11.9 88.1

Total N 196 1719 1915

% 10.2 89.8

5 (6) 1 0

A N 40 1096 1136 59.3

% 3.5 96.5

O N 89 267 356 18.6

% 25.0 75.0

P N 66 340 406 21.2

% 16.3 83.7

F N 1 16 17 0.9

% 5.9 94.1

Total N 196 1719 1915

% 10.2 89.8

6 (7) 1 0

N N 133 1156 1289 67.3

% 10.3 89.7

S N 63 563 626 32.7

% 10.1 89.9

Total N 196 1719 1915

% 10.2 89.8

7 (8) 1 0

V N 120 1498 1618 84.5

% 7.4 92.6

D N 76 221 297 15.5

% 25.6 74.4

Total N 196 1719 1915

% 10.2 89.8

TOTAL N 196 1719 1915

% 10.2 89.8

Name of new cell file: .cel

• BINOMIAL VARBRUL • 10/10/2022 15:36:48 •.....

Name of cell file: .cel

Averaging by weighting factors.

Threshold, step-up/down: 0.050001

Stepping up...

----- Level # 0 -----

Run # 1, 1 cells:

Convergence at Iteration 2

Input 0.102

Log likelihood = -632.363

----- Level # 1 -----

Run # 2, 2 cells:

Convergence at Iteration 6

Input 0.070

Group # 1 -- M: 0.712, F: 0.193

Log likelihood = -573.365 Significance = 0.000

Run # 3, 4 cells:

Convergence at Iteration 6

Input 0.081

Group # 2 -- 1: 0.112, 2: 0.696, 3: 0.447, 4: 0.517

Log likelihood = -595.085 Significance = 0.000

Run # 4, 2 cells:

Convergence at Iteration 5

Input 0.088

Group # 3 -- F: 0.694, S: 0.380

Log likelihood = -596.198 Significance = 0.000

Run # 5, 3 cells:

Convergence at Iteration 6

Input 0.065

Group # 4 -- F: 0.189, A: 0.820, M: 0.661

Log likelihood = -555.973 Significance = 0.000

Run # 6, 4 cells:

Convergence at Iteration 5

Input 0.073

Group # 5 -- A: 0.317, O: 0.808, P: 0.711, F: 0.442

Log likelihood = -557.358 Significance = 0.000

Run # 7, 2 cells:

Convergence at Iteration 2

Input 0.102

Group # 6 -- N: 0.502, S: 0.496

Log likelihood = -632.348 Significance = 0.872

Run # 8, 2 cells:

Convergence at Iteration 5

Input 0.091

Group # 7 -- V: 0.444, D: 0.774

Log likelihood = -596.519 Significance = 0.000

Add Group # 4 with factors FAM

----- Level # 2 -----

Run # 9, 6 cells:

Convergence at Iteration 6

Input 0.041

Group # 1 -- M: 0.729, F: 0.173

Group # 4 -- F: 0.171, A: 0.835, M: 0.677

Log likelihood = -486.919 Significance = 0.000

Run # 10, 12 cells:

Convergence at Iteration 6

Input 0.051

Group # 2 -- 1: 0.105, 2: 0.679, 3: 0.442, 4: 0.552

Group # 4 -- F: 0.191, A: 0.819, M: 0.660

Log likelihood = -522.574 Significance = 0.000

Run # 11, 6 cells:

Convergence at Iteration 6

Input 0.054

Group # 3 -- F: 0.708, S: 0.371

Group # 4 -- F: 0.178, A: 0.820, M: 0.677

Log likelihood = -517.235 Significance = 0.000

Run # 12, 7 cells:

Convergence at Iteration 5

Input 0.058

Group # 4 -- F: 0.225, A: 0.784, M: 0.638

Group # 5 -- A: 0.395, O: 0.600, P: 0.709, F: 0.220

Log likelihood = -534.458 Significance = 0.000

Run # 13, 6 cells:

Convergence at Iteration 6

Input 0.065

Group # 4 -- F: 0.188, A: 0.821, M: 0.662

Group # 6 -- N: 0.488, S: 0.525

Log likelihood = -555.588 Significance = 0.398

Run # 14, 4 cells:

Convergence at Iteration 14

Input 0.065

Group # 4 -- F: 0.200, A: 0.777, M: 0.676

Group # 7 -- V: 0.484, D: 0.587

Log likelihood = -555.013 Significance = 0.174

Add Group # 1 with factors MF

----- Level # 3 -----

Run # 15, 24 cells:

Convergence at Iteration 6

Input 0.032

Group # 1 -- M: 0.731, F: 0.170

Group # 2 -- 1: 0.099, 2: 0.680, 3: 0.442, 4: 0.559

Group # 4 -- F: 0.166, A: 0.830, M: 0.688

Log likelihood = -454.155 Significance = 0.000

Run # 16, 12 cells:

Convergence at Iteration 6

Input 0.036

Group # 1 -- M: 0.722, F: 0.181

Group # 3 -- F: 0.693, S: 0.381

Group # 4 -- F: 0.166, A: 0.827, M: 0.689

Log likelihood = -457.203 Significance = 0.000

Run # 17, 13 cells:

Convergence at Iteration 6

Input 0.036

Group # 1 -- M: 0.728, F: 0.174

Group # 4 -- F: 0.198, A: 0.802, M: 0.662

Group # 5 -- A: 0.400, O: 0.604, P: 0.696, F: 0.190

Log likelihood = -468.314 Significance = 0.000

Run # 18, 12 cells:

Convergence at Iteration 6

Input 0.041

Group # 1 -- M: 0.729, F: 0.173

Group # 4 -- F: 0.171, A: 0.835, M: 0.677

Group # 6 -- N: 0.497, S: 0.505

Log likelihood = -486.902 Significance = 0.863

Run # 19, 8 cells:

Convergence at Iteration 17

Input 0.041

Group # 1 -- M: 0.729, F: 0.173

Group # 4 -- F: 0.172, A: 0.832, M: 0.678

Group # 7 -- V: 0.499, D: 0.506

*** Warning, negative change in likelihood (-0.00356859) replaced by 0.0.

Log likelihood = -486.920 Significance = 1.000

Add Group # 3 with factors FS

----- Level # 4 -----

Run # 20, 48 cells:

Convergence at Iteration 8

Input 0.030

Group # 1 -- M: 0.721, F: 0.182

Group # 2 -- 1: 0.106, 2: 0.608, 3: 0.518, 4: 0.602

Group # 3 -- F: 0.670, S: 0.396

Group # 4 -- F: 0.159, A: 0.825, M: 0.701

Log likelihood = -437.527 Significance = 0.000

Run # 21, 25 cells:

Convergence at Iteration 6

Input 0.032

Group # 1 -- M: 0.720, F: 0.183

Group # 3 -- F: 0.697, S: 0.378

Group # 4 -- F: 0.199, A: 0.790, M: 0.669

Group # 5 -- A: 0.394, O: 0.603, P: 0.708, F: 0.223

Log likelihood = -438.380 Significance = 0.000

Run # 22, 24 cells:

Convergence at Iteration 6

Input 0.036

Group # 1 -- M: 0.722, F: 0.181

Group # 3 -- F: 0.693, S: 0.380

Group # 4 -- F: 0.166, A: 0.827, M: 0.690

Group # 6 -- N: 0.504, S: 0.492

Log likelihood = -457.165 Significance = 0.784

Run # 23, 16 cells:

Convergence at Iteration 19

Input 0.036

Group # 1 -- M: 0.722, F: 0.181

Group # 3 -- F: 0.693, S: 0.380

Group # 4 -- F: 0.164, A: 0.837, M: 0.686

Group # 7 -- V: 0.504, D: 0.477

Log likelihood = -457.145 Significance = 0.738

Add Group # 2 with factors 1234

----- Level # 5 -----

Run # 24, 84 cells:

Convergence at Iteration 8

Input 0.025

Group # 1 -- M: 0.721, F: 0.182

Group # 2 -- 1: 0.094, 2: 0.604, 3: 0.553, 4: 0.602

Group # 3 -- F: 0.682, S: 0.388

Group # 4 -- F: 0.187, A: 0.789, M: 0.685

Group # 5 -- A: 0.389, O: 0.606, P: 0.720, F: 0.209

Log likelihood = -417.099 Significance = 0.000

Run # 25, 95 cells:

Convergence at Iteration 8

Input 0.030

Group # 1 -- M: 0.722, F: 0.181

Group # 2 -- 1: 0.106, 2: 0.610, 3: 0.517, 4: 0.601

Group # 3 -- F: 0.671, S: 0.395

Group # 4 -- F: 0.159, A: 0.824, M: 0.702

Group # 6 -- N: 0.511, S: 0.478

Log likelihood = -437.280 Significance = 0.487

Run # 26, 61 cells:

No Convergence at Iteration 20

Input 0.030

Group # 1 -- M: 0.724, F: 0.179

Group # 2 -- 1: 0.104, 2: 0.609, 3: 0.519, 4: 0.603

Group # 3 -- F: 0.671, S: 0.395

Group # 4 -- F: 0.151, A: 0.858, M: 0.689

Group # 7 -- V: 0.514, D: 0.424

Log likelihood = -437.042 Significance = 0.333

Add Group # 5 with factors AOPF

----- Level # 6 -----

Run # 27, 148 cells:

Convergence at Iteration 8

Input 0.025

Group # 1 -- M: 0.722, F: 0.181

Group # 2 -- 1: 0.093, 2: 0.605, 3: 0.554, 4: 0.602

Group # 3 -- F: 0.683, S: 0.387

Group # 4 -- F: 0.189, A: 0.788, M: 0.684

Group # 5 -- A: 0.389, O: 0.605, P: 0.721, F: 0.219

Group # 6 -- N: 0.509, S: 0.482

Log likelihood = -416.941 Significance = 0.591

Run # 28, 97 cells:

Convergence at Iteration 18

Input 0.025

Group # 1 -- M: 0.723, F: 0.180

Group # 2 -- 1: 0.092, 2: 0.605, 3: 0.554, 4: 0.604

Group # 3 -- F: 0.683, S: 0.387

Group # 4 -- F: 0.181, A: 0.817, M: 0.676

Group # 5 -- A: 0.386, O: 0.615, P: 0.718, F: 0.224

Group # 7 -- V: 0.512, D: 0.434

Log likelihood = -416.751 Significance = 0.422

No remaining groups significant

Groups selected while stepping up: 4 1 3 2 5

Best stepping up run: #24

Stepping down...

----- Level # 7 -----

Run # 29, 166 cells:

No Convergence at Iteration 20

Input 0.025

Group # 1 -- M: 0.725, F: 0.178

Group # 2 -- 1: 0.090, 2: 0.606, 3: 0.555, 4: 0.604

Group # 3 -- F: 0.685, S: 0.386

Group # 4 -- F: 0.181, A: 0.820, M: 0.673

Group # 5 -- A: 0.385, O: 0.617, P: 0.719, F: 0.242

Group # 6 -- N: 0.512, S: 0.476

Group # 7 -- V: 0.515, D: 0.421

Log likelihood = -416.465

----- Level # 6 -----

Run # 30, 98 cells:

No Convergence at Iteration 20

Input 0.039

Group # 2 -- 1: 0.105, 2: 0.649, 3: 0.495, 4: 0.563

Group # 3 -- F: 0.688, S: 0.384

Group # 4 -- F: 0.224, A: 0.750, M: 0.661

Group # 5 -- A: 0.383, O: 0.611, P: 0.729, F: 0.208

Group # 6 -- N: 0.493, S: 0.515

Group # 7 -- V: 0.492, D: 0.546

Log likelihood = -470.452 Significance = 0.000

Run # 31, 55 cells:

Convergence at Iteration 15

Input 0.032

Group # 1 -- M: 0.721, F: 0.183

Group # 3 -- F: 0.698, S: 0.377

Group # 4 -- F: 0.198, A: 0.794, M: 0.668

Group # 5 -- A: 0.393, O: 0.607, P: 0.708, F: 0.227

Group # 6 -- N: 0.500, S: 0.500

Group # 7 -- V: 0.502, D: 0.488

Log likelihood = -438.361 Significance = 0.000

Run # 32, 97 cells:

Convergence at Iteration 19

Input 0.028

Group # 1 -- M: 0.730, F: 0.172

Group # 2 -- 1: 0.089, 2: 0.681, 3: 0.466, 4: 0.557

Group # 4 -- F: 0.188, A: 0.817, M: 0.665

Group # 5 -- A: 0.395, O: 0.611, P: 0.701, F: 0.205

Group # 6 -- N: 0.507, S: 0.486

Group # 7 -- V: 0.510, D: 0.445

Log likelihood = -435.337 Significance = 0.000

Run # 33, 120 cells:

No Convergence at Iteration 20

Input 0.035

Group # 1 -- M: 0.709, F: 0.197

Group # 2 -- 1: 0.108, 2: 0.632, 3: 0.578, 4: 0.536

Group # 3 -- F: 0.672, S: 0.394

Group # 5 -- A: 0.291, O: 0.849, P: 0.724, F: 0.549

Group # 6 -- N: 0.523, S: 0.452

Group # 7 -- V: 0.515, D: 0.419

Log likelihood = -448.465 Significance = 0.000

Run # 34, 115 cells:

No Convergence at Iteration 20

Input 0.029

Group # 1 -- M: 0.725, F: 0.178

Group # 2 -- 1: 0.102, 2: 0.611, 3: 0.518, 4: 0.603

Group # 3 -- F: 0.673, S: 0.394

Group # 4 -- F: 0.150, A: 0.862, M: 0.688

Group # 6 -- N: 0.514, S: 0.472

Group # 7 -- V: 0.516, D: 0.411

Log likelihood = -436.635 Significance = 0.000

Run # 35, 97 cells:

Convergence at Iteration 18

Input 0.025

Group # 1 -- M: 0.723, F: 0.180

Group # 2 -- 1: 0.092, 2: 0.605, 3: 0.554, 4: 0.604

Group # 3 -- F: 0.683, S: 0.387

Group # 4 -- F: 0.181, A: 0.817, M: 0.676

Group # 5 -- A: 0.386, O: 0.615, P: 0.718, F: 0.224

Group # 7 -- V: 0.512, D: 0.434

Log likelihood = -416.751 Significance = 0.463

Run # 36, 148 cells:

Convergence at Iteration 8

Input 0.025

Group # 1 -- M: 0.722, F: 0.181

Group # 2 -- 1: 0.093, 2: 0.605, 3: 0.554, 4: 0.602

Group # 3 -- F: 0.683, S: 0.387

Group # 4 -- F: 0.189, A: 0.788, M: 0.684

Group # 5 -- A: 0.389, O: 0.605, P: 0.721, F: 0.219

Group # 6 -- N: 0.509, S: 0.482

Log likelihood = -416.941 Significance = 0.339

Cut Group # 6 with factors NS

----- Level # 5 -----

Run # 37, 54 cells:

No Convergence at Iteration 20

Input 0.039

Group # 2 -- 1: 0.104, 2: 0.650, 3: 0.495, 4: 0.563

Group # 3 -- F: 0.688, S: 0.384

Group # 4 -- F: 0.225, A: 0.752, M: 0.659

Group # 5 -- A: 0.383, O: 0.612, P: 0.729, F: 0.217

Group # 7 -- V: 0.493, D: 0.538

Log likelihood = -470.570 Significance = 0.000

Run # 38, 29 cells:

Convergence at Iteration 12

Input 0.032

Group # 1 -- M: 0.721, F: 0.183

Group # 3 -- F: 0.698, S: 0.377

Group # 4 -- F: 0.198, A: 0.794, M: 0.668

Group # 5 -- A: 0.394, O: 0.606, P: 0.708, F: 0.226

Group # 7 -- V: 0.502, D: 0.488

Log likelihood = -438.362 Significance = 0.000

Run # 39, 55 cells:

Convergence at Iteration 17

Input 0.028

Group # 1 -- M: 0.729, F: 0.172

Group # 2 -- 1: 0.090, 2: 0.680, 3: 0.466, 4: 0.557

Group # 4 -- F: 0.187, A: 0.816, M: 0.667

Group # 5 -- A: 0.396, O: 0.610, P: 0.701, F: 0.196

Group # 7 -- V: 0.509, D: 0.453

Log likelihood = -435.433 Significance = 0.000

Run # 40, 70 cells:

No Convergence at Iteration 20

Input 0.035

Group # 1 -- M: 0.705, F: 0.201

Group # 2 -- 1: 0.113, 2: 0.631, 3: 0.578, 4: 0.530

Group # 3 -- F: 0.672, S: 0.394

Group # 5 -- A: 0.297, O: 0.841, P: 0.721, F: 0.487

Group # 7 -- V: 0.510, D: 0.444

Log likelihood = -449.611 Significance = 0.000

Run # 41, 61 cells:

No Convergence at Iteration 20

Input 0.030

Group # 1 -- M: 0.724, F: 0.179

Group # 2 -- 1: 0.104, 2: 0.609, 3: 0.519, 4: 0.603

Group # 3 -- F: 0.671, S: 0.395

Group # 4 -- F: 0.151, A: 0.858, M: 0.689

Group # 7 -- V: 0.514, D: 0.424

Log likelihood = -437.042 Significance = 0.000

Run # 42, 84 cells:

Convergence at Iteration 8

Input 0.025

Group # 1 -- M: 0.721, F: 0.182

Group # 2 -- 1: 0.094, 2: 0.604, 3: 0.553, 4: 0.602

Group # 3 -- F: 0.682, S: 0.388

Group # 4 -- F: 0.187, A: 0.789, M: 0.685

Group # 5 -- A: 0.389, O: 0.606, P: 0.720, F: 0.209

Log likelihood = -417.099 Significance = 0.422

Cut Group # 7 with factors VD

----- Level # 4 -----

Run # 43, 46 cells:

No Convergence at Iteration 20

Input 0.039

Group # 2 -- 1: 0.104, 2: 0.650, 3: 0.496, 4: 0.563

Group # 3 -- F: 0.690, S: 0.383

Group # 4 -- F: 0.220, A: 0.772, M: 0.653

Group # 5 -- A: 0.383, O: 0.613, P: 0.728, F: 0.224

Log likelihood = -470.742 Significance = 0.000

Run # 44, 25 cells:

Convergence at Iteration 6

Input 0.032

Group # 1 -- M: 0.720, F: 0.183

Group # 3 -- F: 0.697, S: 0.378

Group # 4 -- F: 0.199, A: 0.790, M: 0.669

Group # 5 -- A: 0.394, O: 0.603, P: 0.708, F: 0.223

Log likelihood = -438.380 Significance = 0.000

Run # 45, 47 cells:

Convergence at Iteration 7

Input 0.028

Group # 1 -- M: 0.727, F: 0.175

Group # 2 -- 1: 0.091, 2: 0.679, 3: 0.466, 4: 0.557

Group # 4 -- F: 0.192, A: 0.797, M: 0.674

Group # 5 -- A: 0.397, O: 0.603, P: 0.702, F: 0.187

Log likelihood = -435.637 Significance = 0.000

Run # 46, 56 cells:

Convergence at Iteration 8

Input 0.036

Group # 1 -- M: 0.703, F: 0.204

Group # 2 -- 1: 0.115, 2: 0.630, 3: 0.577, 4: 0.529

Group # 3 -- F: 0.671, S: 0.395

Group # 5 -- A: 0.307, O: 0.816, P: 0.729, F: 0.432

Log likelihood = -449.891 Significance = 0.000

Run # 47, 48 cells:

Convergence at Iteration 8

Input 0.030

Group # 1 -- M: 0.721, F: 0.182

Group # 2 -- 1: 0.106, 2: 0.608, 3: 0.518, 4: 0.602

Group # 3 -- F: 0.670, S: 0.396

Group # 4 -- F: 0.159, A: 0.825, M: 0.701

Log likelihood = -437.527 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down: 6 7

Best stepping up run: #24

Best stepping down run: #42