



UNIVERSIDADE FEDERAL DO RIO GRANDE
INSTITUTO DE LETRAS E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ÁREA DE CONCENTRAÇÃO – ESTUDOS DA LINGUAGEM

ANDRESSA COELHO FRANCO

**A PREPOSIÇÃO *PARA* E SUA VARIANTE *PRA* NA ESCRITA DE
ESCOLARES**

RIO GRANDE – RS

2020



**UNIVERSIDADE FEDERAL DO RIO GRANDE
INSTITUTO DE LETRAS E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ÁREA DE CONCENTRAÇÃO – ESTUDOS DA LINGUAGEM**

ANDRESSA COELHO FRANCO

**A PREPOSIÇÃO *PARA* E SUA VARIANTE *PRA* NA ESCRITA DE
ESCOLARES**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande – FURG como requisito parcial para a obtenção do título de Mestre em Letras.

Orientadora: Prof^ª. Dr^ª. Marisa Porto do Amaral

RIO GRANDE – RS

2020

Ficha Catalográfica

F825p Franco, Andressa Coelho.
A preposição *para* e sua variante *pra* na escrita de
escolares / Andressa Coelho Franco. – 2020.
96 f.

Dissertação (mestrado) – Universidade Federal do Rio Grande –
FURG, Programa de Pós-Graduação em Letras, Rio Grande/RS,
2020.

Orientadora: Dra. Marisa Porto do Amaral.

1. Variação Linguística 2. Texto 3. Preposição *para* I.
Amaral, Marisa Porto do II. Título.

CDU 81

Aos meus pais: Mario Luiz Alves
Franco e Clenir Terezinha Coelho
Franco
Ao meu “Chéri”: Mario Augusto da
Cunha Feijó
Aos meus “filhos peludos”: Miu e Frito

AGRADECIMENTOS

Em primeiro lugar, ao nosso Senhor Jesus Cristo por mostrar-me a verdade e conduzir-me ao caminho da luz por meio do sentimento amoroso, da compaixão, da aceitação e do respeito para com todas as pessoas, sem distinção.

Agradeço também, ao Cristo, por ensinar-me a difícil tarefa de perdoar.

Assim, agradeço ao Filho, por mostrar-me a face de Deus, nosso Pai.

Em segundo, e não menor lugar, ao Espiritismo Kardecista que me mostrou um caminho de vida, menos doloroso e mais compreensível, por meio de uma doutrina simples, com ensinamentos de amor, compreensão e caridade.

À minha mãe, Clenir, que mesmo diante da dificuldade em conceber um filho, não desistiu até que eu nascesse. Agradeço a ela, também, por ter despertado em mim, o gosto pela leitura, através das muitas histórias da “Turma da Mônica” antes de dormir.

Ao meu pai, Mario, pela dedicação e cuidados para que eu crescesse bem e também por ensinar-me que a vida pode ser cruel, por vezes, mas que cabe a mim ser forte e enfrentar as adversidades de maneira honrosa e honesta.

Ao meu irmão, Marcel, que compartilhou comigo a infância pura e feliz.

Ao meu querido companheiro, Mario, meu “Chéri”, pelo amor e respeito, e também, por ter ficado ao meu lado, amparando e apoiando durante todo esse tempo de Mestrado.

Aos meus enteados, Sibelle e Gustavo, por proporcionarem-me amizade em um sentido diferente de família.

Aos meus “filhos de patas”, Miu e Frito por me darem todos os dias, amor e amizade incondicionais.

Aos meus primos, que me ensinaram o sentido alegre e divertido de família.

Aos meus padrinhos, Ronaldo (Vera) e Marilaine, por me darem um segundo apoio depois de meus pais.

Aos meus alunos, por me ensinarem a cada dia e por me lembrarem da minha adolescência, “tempo bom que não volta nunca mais...”.

Aos meus amigos, por estarem comigo ao longo dessa trajetória chamada vida.

Aos meus alunos colaboradores, que doaram suas produções textuais para que eu pudesse realizar a pesquisa presente nesta dissertação.

À E.E.E.M. Marechal Mascarenhas de Moraes por ser meu ambiente de trabalho e aperfeiçoamento profissional.

À leitura, que me fez e sempre me fará crescer como ser humano através dos conhecimentos infinitos que ela carrega consigo.

Ao ensino público, pois foi e é através dele que pude e posso estudar em busca de uma vida mais digna.

À minha orientadora, Marisa, pela dedicação para comigo e ajuda na realização deste trabalho.

À FURG, por me proporcionar, desde a primeira graduação, ambiente de ensino, sociabilidade e crescimento pessoal.

Poeminha do Contra

Todos esses que aí estão
Atravancando meu caminho,
Eles passarão...
Eu passarinho!

Mario Quintana

RESUMO

A presente pesquisa busca analisar a preposição *PARA* e sua variante *PRA* na escrita escolar de alunos, dos primeiros e o terceiro ano do Ensino Médio; e se embasa nos pressupostos teóricos da Sociolinguística Quantitativa de LABOV (2008[1972]). A pesquisa busca também, aperfeiçoar o estudo em nível de *Especialização* de Franco (2017). A coleta de dados foi realizada em uma escola pública localizada em um bairro próximo ao centro da cidade de Rio Grande – RS e teve como *corpus* os textos produzidos em *verso* e *prosa* por 152 escolares. A amostra, composta por 1116 dados, foi analisada pelos grupos de fatores linguísticos internos (*Contexto morfológico seguinte*, *Contexto fonológico seguinte*, *Tonicidade da sílaba seguinte*, *Número de sílabas do item seguinte* e *Tipo textual*) e os de fatores extralinguísticos (*Sexo*, *Faixa etária* e *Nível de escolaridade*), através do programa estatístico GOLDVARB X. A pesquisa teve como hipóteses as seguintes questões: a) as variantes *PARA* e *PRA* vão se comportar de forma distinta em cada modalidade (*verso* e *prosa*) e b) os alunos vão considerar a escolha das variantes a partir do grau de formalidade dos textos propostos. Através dos resultados oriundos da modalidade *Tipo textual* (*verso* e *prosa*), confirmamos as hipóteses de que as variantes *PARA* e *PRA* se comportam de forma distinta nos diferentes tipos de textos e também que os alunos consideram a escolha das variantes a partir do grau de formalidade dos textos propostos. O contexto *verso* é o que favorece mais a variante *PRA* (0.83), em oposição ao contexto *prosa* (0.45), que inibe o surgimento de tal variante.

Palavras-chave: **Variação Linguística; Texto; Preposição *PARA***

RÉSUMÉ

La présente recherche cherche à analyser la préposition *POUR* et sa variante *PRA* dans la rédaction des étudiants qui fréquentent les premières et troisièmes années du lycée; et s'appuie sur les hypothèses théoriques de la sociolinguistique quantitative de LABOV (2008[1972]). La recherche vise également à améliorer l'étude au niveau de spécialisation de Franco (2017). La collecte de données a été réalisée dans une école publique, située dans un quartier proche du centre-ville de Rio Grande - RS, et avait pour corpus les textes produits en vers et en prose pour 152 étudiantes. L'échantillon, composé de 1116 données, a été analysé par les groupes de facteurs linguistiques internes (*Contexte morphologique suivant, Contexte phonologique suivant, Tonicité des syllabes suivantes, Nombre de syllabes dans l'élément suivant* et *Typologie textuelle*), et ceux de facteurs extralinguistiques (*Sexe, Groupe d'âge* et *Niveau d'éducation*) à travers le programme statistique GOLDVARB X. La recherche avait pour hypothèse les faits suivantes: a) les variantes *POUR* et *PRA* se comporteront différemment selon chaque modalité et b) les étudiants choisiront des variantes en fonction du degré de formalité des textes proposés. D'après les résultats de la Typologie textuelle (*vers* et *prose*), nous confirmons l'hypothèse selon laquelle les variantes *POUR* et *PRA* se comporteront différemment dans les différents types de textes, et les étudiants considèrent le choix des variantes en fonction du degré de formalité des textes proposés. Le contexte en vers privilégie davantage la variante *PRA* (0,83) par opposition au contexte en prose (0,45), qui réfère l'émergence d'une telle variante.

Mots-clés : **Variation linguistique ; Texte ; Préposition *POUR***

LISTA DE GRÁFICOS

| | |
|--|----|
| Gráfico 1 – Distribuição geral dos dados | 42 |
| Gráfico 2 – Comparação dos dados entre pesquisas da variável <i>PARA</i> | 44 |
| Gráfico 3 – Dados referentes ao cruzamento dos fatores <i>Sexo</i> e <i>Nível de escolaridade</i> | 56 |
| Gráfico 4 – Dados referentes ao cruzamento dos fatores <i>Contexto fonológico seguinte</i> e <i>Contexto morfológico seguinte</i> | 58 |
| Gráfico 5 – Dados referentes ao cruzamento dos fatores <i>Tipo textual</i> e <i>Contexto</i> <i>morfológico seguinte</i> | 59 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1 – Distribuição dos dados por variante em cada pesquisa | 43 |
| Tabela 2 – Variante <i>PRA</i> considerando o <i>Tipo textual</i> | 45 |
| Tabela 3 – Variante <i>PRA</i> considerando o <i>Nível de escolaridade</i> | 47 |
| Tabela 4 – Variante <i>PRA</i> considerando o <i>Sexo</i> do informante..... | 48 |
| Tabela 5 – Variante <i>PRA</i> considerando o <i>Contexto fonológico Seguinte</i> (com amalgamação) | 50 |
| Tabela 6– Cruzamento entre os grupos de fatores <i>Faixa etária</i> e <i>Nível de escolaridade</i> | 52 |
| Tabela 7 – Variante <i>PRA</i> considerando a <i>Faixa etária</i> | 53 |
| Tabela 8 – Cruzamento entre os grupos de fatores <i>Sexo</i> e <i>Nível de escolaridade</i> | 55 |
| Tabela 9 – Cruzamento entre os grupos de fatores <i>Contexto fonológico seguinte</i> e <i>Contexto morfológico seguinte</i> | 57 |
| Tabela 10 – Cruzamento entre os grupos de fatores <i>Tipo textual</i> e <i>Contexto morfológico seguinte</i> | 58 |
| Tabela 11 – <i>Contexto morfológico seguinte</i> | 62 |
| Tabela 12 – <i>Tonicidade da sílaba seguinte</i> | 63 |
| Tabela 13 – Cruzamento entre os grupos de fatores <i>Número de sílabas do item seguinte</i> e <i>Contexto morfológico seguinte</i> | 64 |

SUMÁRIO

| | |
|---|-----------|
| 1 INTRODUÇÃO | 13 |
| 2 OBJETO DE ESTUDO: A PREPOSIÇÃO <i>PARA</i> | 15 |
| 3 REFERENCIAL TEÓRICO | 19 |
| 3.1 A SOCIOLINGUÍSTICA | 19 |
| 3.2 A FONOLOGIA | 23 |
| 3.2.1 A redução silábica de <i>PARA</i> | 23 |
| 3.3 DA FALA PARA A ESCRITA | 26 |
| 3.4 ESTUDOS SOCIOLINGUÍSTICOS | 28 |
| 3.4.1 O estudo de Franco (2017) | 28 |
| 3.4.2 O estudo de Nascimento & Vieira | 29 |
| 3.4.3 O estudo de Silva (2010) | 30 |
| 3.4.4 O estudo de Ferreira (2014) | 31 |
| 3.4.5 O estudo de Ferreira (2018) | 33 |
| 4. METODOLOGIA | 37 |
| 4.1 COMPOSIÇÃO DA AMOSTRA | 37 |
| 4.2 VARIÁVEIS CONTROLADAS | 39 |
| 4.2.1 Variável Dependente | 39 |
| 4.2.2 Variáveis Independentes | 39 |
| 4.2.2.1 Grupos de Fatores Linguísticos | 39 |
| 4.2.2.2 Grupos de Fatores Extralinguísticos | 40 |
| 4.3 OBJETIVO DO ESTUDO | 41 |
| 4.3.1 Objetivo geral | 41 |
| 4.3.2 Objetivos específicos | 41 |
| 5. RESULTADOS DA ANÁLISE ESTATÍSTICA | 42 |

| | |
|---|-----------|
| 5.1 DISTRIBUIÇÃO DA VARIANTE <i>PRA</i> REFERENTE À PREPOSIÇÃO <i>PARA</i> | 42 |
| 5.2 VARIÁVEIS LINGUÍSTICAS E EXTRALINGUÍSTICAS SIGNIFICATIVAS PARA O USO DA VARIANTE <i>PRA</i> NA MODALIDADE ESCRITA | 44 |
| 5.2.1 <i>Tipo textual</i> | 45 |
| 5.2.2 <i>Nível de escolaridade</i> | 46 |
| 5.2.3 <i>Sexo</i> | 48 |
| 5.2.4 <i>Contexto fonológico seguinte</i> | 49 |
| 5.3 VARIÁVEIS LINGUÍSTICAS DESCARTADAS | 51 |
| 5.3.1 <i>Número de sílabas do item seguinte</i> | 51 |
| 5.3.2 <i>Tonicidade da sílaba seguinte</i> | 51 |
| 5.3.3 <i>Contexto morfológico seguinte</i> | 51 |
| 5.4 FAIXA ETÁRIA (uma variável sobreposta) | 52 |
| 5.5 CRUZAMENTO ENTRE FATORES LINGUÍSTICOS OU EXTRALINGUÍSTICOS..... | 54 |
| 5.5.1 <i>Sexo x Nível de escolaridade</i> | 55 |
| 5.5.2 <i>Contexto Fonológico seguinte x Contexto morfológico seguinte</i> | 57 |
| 5.5.3 <i>Tipo textual x Nível de escolaridade</i> | 58 |
| | |
| 6. RODADA DE DADOS SEM OS GRUPOS DE FATORES TIPO TEXTUAL E FAIXA ETÁRIA | 60 |
| | |
| 6.1 GRUPOS DE FATORES SELECIONADOS PELO PROGRAMA GOLDVARB X | 60 |
| 6.1.1 <i>Contexto morfológico seguinte</i> | 61 |
| 6.1.2 <i>Tonicidade da sílaba seguinte</i> | 63 |
| 6.2 CRUZAMENTO ENTRE OS GRUPOS DE FATORES CONTEXTO MORFOLÓGICO SEGUINTE E NÚMERO DE SÍLABAS DO ITEM SEGUINTE | 63 |
| 6.2.1 <i>Número de sílabas do item seguinte x Contexto morfológico seguinte</i> | 64 |
| | |
| 7. CONCLUSÃO | 65 |

8. REFERÊNCIAS BIBLIOGRÁFICAS 68

1 INTRODUÇÃO

A presente pesquisa busca analisar a preposição *PARA*, e sua variante *PRA*, na escrita escolar de estudantes cursando os primeiros e o terceiro ano do Ensino Médio e tem como área de concentração os Estudos da Linguagem.

A motivação para realizar a pesquisa originou-se a partir da observação de produções textuais realizadas em algumas turmas do Ensino Médio, há alguns anos atrás, enquanto ministrava a disciplina de Língua Portuguesa nas respectivas turmas. Diante da análise das produções dos escolares constatou-se a dúvida dos estudantes em usar a forma *PARA* ou a variante *PRA* em textos escritos. Percebe-se que eles têm dificuldades em fazer a diferenciação entre *PARA* e *PRA* em textos mais formais, como redação escolar.

Os estudantes optam, muitas vezes, pela variante *PRA* ao invés da forma *PARA* em suas produções, o que pode ser devido à escrita da *internet*, veículo no qual eles se conectam muitas horas por dia, ou também de acordo com os registros da fala. Sabe-se, que o registro de *PRA* está disseminado em todos os níveis de escolaridades e faixas etárias, o que contribui para a incerteza dos escolares na produção de textos que exigem linguagem mais formal.

Nos mecanismos televisivos que são de rede aberta para toda a população brasileira (“rede em massa”), por tratarem de uma linguagem mais coloquial, frequentemente observa-se a aparição da variante *PRA* ao invés da forma *PARA*. Temos como exemplo, o quadro do programa humorístico do apresentador Danilo Gentili, que tem como nome a seguinte frase: “Pra quem você veste o Chapéu?”. Tal quadro é oriundo do programa de outro apresentador chamado Raul Gil, com o respectivo nome: “Pra quem você tira o chapéu?”.

Percebe-se que os livros infanto-juvenis também utilizam a modalidade *PRA* para o discurso direto, usando a modalidade *PARA* nas demais colocações preposicionais, o que gera ainda mais dúvida entre os alunos em relação a qual modalidade utilizar em suas produções textuais. Além disso, a discussão entre o uso de *PARA* e *PRA* estende-se pelos caminhos da *internet* e em um dos sites pesquisados por Franco (2017), um internauta posicionou-se comentando que o fato de *PARA* ser a forma “correta”, não elimina a possibilidade de usarmos *PRA*. Tudo, para ele “é apenas

uma questão de adequação.” Por esse motivo, resolvemos observar o uso das formas *PARA* x *PRA*, na escrita averiguando a ocorrência desta variação, mesmo quando o tipo textual implica um uso da modalidade mais formal.

A pesquisa tem como hipóteses: a) que as variantes *PARA* e *PRA* vão se comportar de formas distintas nos diferentes tipos de textos e b) que os alunos vão considerar a escolha das variantes a partir do grau de formalidade dos textos propostos. A questão ampla que norteia este trabalho é: Como se comportarão as variantes *PARA* e *PRA* na escrita dos alunos dos primeiros e terceiro ano do Ensino Médio?

O trabalho justifica-se pela escassez do estudo da preposição *PARA* e de sua variante *PRA* na escrita, além de possibilitar futuras pesquisas em torno do tema abordado. A pesquisa também se justifica pelo retorno ao meio escolar, já que a sala de aula é o local onde se realiza a coleta de dados. O trabalho fundamenta-se também pela necessidade de um estudo entre as formas *PARA* e *PRA* com vistas a uma discussão mais ampla entre os registros formais e informais, ampliando o interesse pela pesquisa referente a esse assunto. Por partir de uma prática escolar, permite, também, questionamentos pertinentes do professor frente às questões de mudança na língua, o que é tradicional e o que é inovador.

2 OBJETO DE ESTUDO: A PREPOSIÇÃO PARA

Comumente aprendemos na escola que a preposição é uma palavra invariável que liga dois termos da oração, ou seja, liga um termo subordinando esse ao outro. É-nos dito, também, que a preposição estabelece, ainda uma certa relação de dependência entre esses dois termos da oração.

Acredita-se que, sintaticamente, as preposições não exercem propriamente uma função, ou seja, são consideradas conectivos. Os conectivos são elementos de ligação entre termos oracionais. As preposições podem introduzir complementos verbais, complementos nominais, locuções adjetivas, locuções adverbiais e orações reduzidas.

Conforme Luft (1987), dependendo das funções que exercem as preposições, elas são classificadas em *acidentais* (palavras de outras classes gramaticais que, cumulativamente, podem apresentar-se como preposições), as quais podem ser provadas pela regência dos pronomes pessoais, ou *essenciais* (palavras que funcionam somente como preposições).

As preposições essenciais sempre funcionam como preposição e as preposições acidentais são palavras que, além de preposições, podem assumir outras funções morfológicas. Algumas preposições podem unir-se a outras palavras. Temos combinação, quando na junção da preposição com outra palavra não há perda de elemento fonético e temos contração, quando na junção da preposição com outra palavra há perda fonética.

Mattoso Câmara Jr (1979) em sua obra *História e estrutura da Língua Portuguesa*, relata que grande parte das preposições latinas se perderam, subsistindo apenas ao exercer função de prefixos na composição lexical. Assim, a preposição assumiu a marca de subordinação ao verbo de forma exclusiva, dando fim à redundância oriunda do latim, passando a funcionar também junto aos complementos verbais.

O autor nos diz que apenas algumas partículas foram favorecidas em um processo de economia; assim, o sistema preposicional funciona em dois planos gramaticais de significação: um primeiro plano mais concreto que é o da localização no

espaço e no tempo, e desse primeiro plano decorre um segundo, com empregos modais, conceituados, metaforicamente deduzidos, referentes ao estado, à origem, à posse, à finalidade, o meio, à causa, o objetivo e assim, por diante.

Câmara Jr (1979) cria também o conceito de forma *dependente* para os vocábulos formais. Para o autor, as formas *dependentes* são as partículas átonas (como o artigo e as preposições), elas são um vocábulo formal que faz parte de um vocábulo fonológico que tem como característica possuir acento proeminente.

Segundo Câmara Jr (1979), é do plano locativo que surge o objeto deste trabalho, a preposição *PARA*:

De uma aglutinação de per e ad, processada no latim vulgar imperial, surgiu a preposição para (port. arc. pera), que, de início, marcava um percurso com direção definida e, em português, torna a indicação de direção mais complexa, inclusive com as noções complementares de “chegada” e “permanência”; daí, a oposição entre – ir a Paris, com a significação geral de direção, e – ir para Paris, com a significação a mais de ali se estabelecer. (CÂMARA JR, 1979, p. 177 – 178).

A variante *PRA* não apareceu em Câmara Jr, o que foi percebido também ao refazer-se a pesquisa nas gramáticas tradicionais examinadas – Luft (1987), Cunha e Cintra (2001), Sacconi (2001), Almeida (2003), Ferreira (2003), Bechara (2004) e Cegalla (2005). Nessas, a referência do ideal é a forma *PARA*, apenas Luft (1987) e Bechara (2004) apontam a variante *PRA* em sua normatização.

Luft (1987) na *Gramática resumida* aborda a preposição *PARA* e sua variante *PRA*. Segundo o autor, sua obra foi criada para “um estudo e ensino da língua mais arejado e mais eficiente” (LUFT, 1987, p. 15) e opõe-se a análises ou *decorebas* de regras ilusórias, afirmando que conhecer a língua é saber usá-la com adequação e justiça.

Luft (1987, p. 112-113) acredita que o que caracteriza os substantivos e advérbios como complementos e adjuntos são as preposições, só dispensando a elas, os objetos diretos. O autor afirma que as preposições são palavras puramente gramaticais que regem objetos indiretos, e que são “vazias de sentido”.

Para o autor, a preposição *PARA* pertence à classificação *essencial*, podendo se

combinar com algumas outras palavras gramaticais, sobretudo na linguagem coloquial, resultando em combinações como *PRA*:

A preposição *para* também se pode combinar com algumas dessas palavras gramaticais, principalmente no linguajar familiar e popular, apresentando combinações como: *pro(s)*, *pra(s)*, *praquele*, *praquela*, *praquilo*, *prele*, *prai*, etc. (LUFT, 1987, p. 113).

Em sua *Moderna Gramática Portuguesa* (2004), Bechara acredita trazer descrições do nosso idioma, atualizadas e enriquecidas, provenientes da produção acadêmica universitária, das críticas e sugestões de estudiosos da língua, de leituras de teóricos da linguagem, e da leitura dos melhores escritores do nosso país.

Segundo Bechara, a preposição é uma unidade linguística desprovida de independência, em geral átona, juntando-se com substantivos, adjetivos, verbos e advérbios para marcar relações gramaticais que desempenham no discurso, nos grupos unitários nominais ou nas orações.

Para o autor, a preposição exerce unicamente o papel de ser índice da função gramatical do termo que introduz; pois na ligação com outra palavra, sofre redução, havendo uma contração (p. 300). O autor apresenta uma lista das preposições que se contraem, dentre elas, o exemplo de *PRA*: “**Para** (pra) – com o artigo definido: para (pra) + o = pro; para (pra) + os = pros; para (pra) + a = pra; para (pra) + as = pras” (BECHARA, 2004, p. 574).

Na pesquisa realizada em “gramáticas de uso”, Marcos Bagno em sua *Gramática Pedagógica Do Português Brasileiro* (2012), propõe descrever o funcionamento do português brasileiro contemporâneo; e trata das preposições como sendo palavras muito importantes no funcionamento da língua.

Bagno (2012) afirma que as preposições usadas com maior frequência têm sua gramaticalização comprovada por diversos aspectos, e que as preposições podem contrair-se com outras palavras formando amálgamas gramaticais.

O autor apresenta em sua lista a preposição *PARA*: *PRA* (s): “pralguém; pralgum (s); pralguma (s); praquela (s); praquela (s); praquile (s); praqui/prali/prai; praquilo; prela (s); prele

(s); pressa (s); presta (s); presse (s); preste (s); prisso; pristo; pro (s); procê (s); pronde/praonde; proutra (s); proutro (s); prum (s); pruma (s)” (BAGNO, 2012, p. 861).

Destaca-se, portanto, na pesquisa realizada nas gramáticas tradicionais que apenas Luft (1987) e Bechara (2004) tratam a questão da variação da preposição *PARA* e do surgimento da sua variante *PRA*, e que a maioria das gramáticas normativas seguem mostrando-se resistentes às mudanças linguísticas, apontando apenas a preposição *PARA* em sua forma padrão. Bagno (2012), em sua “gramática de uso”, também trata da variante *PRA*, abordando e detalhando o assunto e também apresentando exemplos.

3 REFERENCIAL TEÓRICO

3.1 A SOCIOLINGUÍSTICA

A Sociolinguística é uma ciência que possibilita observar a língua no interior das diversas comunidades de falantes, possibilitando a análise das diversas variações da língua por meio de aspectos sociais e culturais. Assim, essa ciência permite a reflexão da língua a partir da realidade sólida e heterogênea das produções dos falantes.

Mollica (2004) afirma a respeito da Sociolinguística:

A Sociolinguística é uma das subáreas da Linguística e estuda a língua em uso no seio das comunidades de fala, voltando à atenção para um tipo de investigação que correlaciona aspectos linguísticos e sociais. Esta ciência se faz presente num espaço interdisciplinar, na fronteira entre língua e sociedade, focalizando precipuamente os empregos linguísticos concretos, em especial os de caráter heterogêneo. (MOLLICA, 2004, p. 9).

Se partirmos do fato de que nenhum grupo linguístico apresenta uma língua homogênea, entendemos a importância dos estudos sociolinguísticos, pois tais estudos acontecem diante de um olhar voltado para a diversidade do uso da língua, e é fato que a mesma apresenta variações, o que pode ser pontuado nas escritas de diversos meios de circulação.

Tarallo (1990) nos diz que pela ótica da teoria da variação linguística, que é um modelo teórico- metodológico que assume o “caos” linguístico como objeto de estudo, não há caos algum na heterogeneidade da língua, pois os indivíduos se apropriam das mais diversas possibilidades significativas para fazer a facilitação desses processos comunicativos entre seus constituintes e agilizar a interação entre si.

Assim, ao tratar da variação linguística, que é um fenômeno universal, a Sociolinguística presume a realização de formas de alternâncias linguísticas, chamadas de variantes. As variantes caracterizam um fenômeno variável, ou seja, há a possibilidade de duas ou mais variantes para uma variável. Temos como exemplo o objeto de estudo deste trabalho: a preposição *PARA* e sua variante *PRA*, ou como é apresentado no estudo de Silva (2010): as variantes *PRA* e *PA*, ambas aceitas como variação do padrão linguístico da preposição *PARA*.

Cabe salientar que as variantes linguísticas podem permanecer perenes no sistema por longos períodos de tempo, ou podem sofrer mudança. A mudança constitui a existência de uma só variante, ocasionando o desaparecimento das outras, ou seja, uma forma substitui a outra ao invés de ambas existirem em concorrência dentro do sistema linguístico.

Mollica (2004) nos diz que:

Cabe à Sociolinguística investigar o grau de estabilidade ou de mutabilidade da variação, diagnosticar as variáveis que têm efeito positivo ou negativo sobre a emergência dos usos linguísticos alternativos e prever seu comportamento regular e sistemático. Assim, compreende-se que a variação e a mudança são contextualizadas, constituindo o conjunto de parâmetros um complexo estruturado de origens e níveis diversos. Vale dizer, os condicionamentos que concorrem para o emprego de formas variantes são em grande número, agem simultaneamente e emergem de dentro ou de fora dos sistemas linguísticos”. (MOLLICA, 2004, p. 11).

A Sociolinguística tem como foco maior a relação social da língua, abrangendo desde os pequenos grupos sócio-culturais como também as maiores comunidades. Seu estudo configura-se de total importância por abrir espaço para a variação, aceitando-a e colocando-a como fenômeno que atua em qualquer nível de linguagem, sem discriminar o falante que a produz.

Em nossa pesquisa, trabalharemos com a Sociolinguística Variacionista que pode ser tratada com outros termos. Segundo Coelho *et al* (2015):

(i) Sociolinguística Laboviana, porque seu principal expoente é o linguista norte-americano William Labov; (ii) Sociolinguística Quantitativa, porque, a princípio, os pesquisadores dessa área costumam lidar com uma grande quantidade de dados de uso da língua, o que requer normalmente uma análise estatística; e (iii) Teoria da Variação e Mudança Linguística, por conta de suas principais preocupações: a variação e a mudança na língua. (COELHO *et al*, 2015, p. 14).

Ao falar da Sociolinguística, não podemos esquecer-nos de mencionar quatro termos muito referidos nesse estudo, que são: variedade, variação, variável e variante. Em relação à variedade temos a culta (oriunda de um contexto mais formal de linguagem) e a coloquial (usada em contextos mais descontraídos, como a fala do dia a dia). Já em relação à variação, temos a “norma” (oriunda das gramáticas normativas) e sua variação, ou seja, uma elaboração “desviada” dessa “norma”).

Labov (2008[1972]) propôs também o termo variável linguística, que é o elemento que possui diferentes significações.

Podemos definir uma variável sociolinguística como correlacionada com alguma variável não linguística do contexto social: o falante, o interlocutor, o público, o ambiente, etc. Alguns traços linguísticos (que chamaremos de indicadores) mostram uma distribuição regular pelos grupos socioeconômicos, étnicos ou etários, mas são usados por cada indivíduo mais ou menos do mesmo modo em qualquer contexto. (LABOV, 2008[1972], p. 275).

Assim, a variável é onde ocorre a variação, e a variante acontece quando um termo, palavra de um certo idioma proveniente de um “padrão ideal” dessa linguagem é pronunciada no “ideal culto” ou no “desvio gramatical” desse ideal. Tomemos como exemplo o objeto deste estudo: A preposição *PARA* e sua variante *PRA*.

As variantes são percebidas como diferentes modos de se dizer a mesma coisa, e são entendidas como competindo na língua. Segundo Tarallo (1990), “as variantes linguísticas são, portanto, diversas maneiras de se dizer a mesma coisa em um mesmo contexto, e com o mesmo valor de verdade. A um conjunto de variantes dá-se o nome de variável”.

As variantes podem permanecer estáveis nos sistemas por um período curto ou até por séculos, ou podem sofrer mudanças quando uma das formas desaparece. As alternâncias de usos analisadas quando estudamos a variação, são influenciadas por fatores sociais que se apresentam sistemáticos e estatisticamente previsíveis; dessa maneira, é possível fazer análises e descrições científicas.

Logo, um dos objetivos da Sociolinguística é compreender os fatores determinantes que promovem a variação linguística e também a importância de cada um deles na efetivação do surgimento da variável. Segundo Tarallo (1990):

O modelo de análise proposto por Labov apresenta-se como uma reação à ausência do componente social no modelo gerativo. Foi, portanto, William Labov quem, mais veemente, voltou a insistir na relação entre língua e sociedade e na possibilidade, virtual e real, de se sistematizar a variação existente e própria da língua falada. (TARALLO. 1990, p. 7).

Coelho *et al* (2015) apontam que a Sociolinguística é uma das áreas da Linguística que possui métodos que se propõem a compreender as variações e as mudanças da língua. Assim, as autoras nos dizem que a Sociolinguística entende a língua como um sistema que possui organização, tendo suas regras de categorias e suas

regras de variações.

Dessa maneira, ao tratar das regras variáveis da língua, as autoras nos dizem que elas são sistemáticas, ou seja, apontam padrões linguísticos, padrões sociais e padrões estilísticos de comportamento. Assim, “o termo padrão, nesse caso, é entendido como um uso regular e frequente de uma dada variante, isto é, como uma tendência de comportamento linguístico. (Coelho *et al*, 2015, p. 62).

Além disso, se de um lado está a língua em sua heterogeneidade, do outro lado está a chamada competência linguística dos falantes. Tal competência é o que abrange o domínio do falante para o tratamento da heterogeneidade do sistema. Portanto:

Ora, se tanto o sistema linguístico quanto a competência dos falantes abrem espaço para estruturas heterogêneas e se o comportamento linguístico variável dos falantes pode ser entendido como o domínio de diversos estilos, chegamos naturalmente a mais um princípio da Teoria da Variação e Mudança: Não existe falante de estilo único. (COELHO *et al*, 2015, p. 63).

A autora aponta também o fato que envolve a negação de muitos em relação à heterogeneidade das variedades linguísticas que uma língua suporta. Essa negação reflete o julgamento diante da língua e dos falantes, o que permite a existência do chamado preconceito linguístico, que nada mais é do que tratar por “erro” as diversidades de fala que não são provenientes da construção padrão da linguagem, determinada pelo *status* econômico e social mais elevado do país.

O preconceito linguístico abre espaço para a marginalização dos usuários que fazem uso das mais diversas variedades construídas no falar de uma língua, e isso ocasiona o preconceito social, pois apontar que a fala de outrem é “errada”, faz julgamento, também, do *status* social desse usuário da língua.

Assim, a Sociolinguística com seus diversos estudos e pesquisas relacionadas ao tema da variação, ajuda seriamente no combate ao preconceito linguístico, pois apresenta a língua como um sistema heterogêneo, não estático, passível de mudança, além de não apontar as variações como “erro”.

3.2 A FONOLOGIA

A Fonologia é o estudo que trata de assuntos referentes ao aspecto sonoro de uma língua. A Fonologia se preocupa em “tratar de sons que distinguem o significado das palavras, além de organizar, postular regras e entender como se dá a variação na realização efetiva dos sons.” (SEARA, NUNES e LAZZAROTTO-VOLCÃO, 2015, p. 21).

Diversas alternâncias podem ser verificadas nas produções de diferentes falantes do português brasileiro. Esse fato pode ser averiguado quando vemos a comunicação entre falantes de diferentes regiões do país. Além disso, essas alternâncias podem ser verificadas também, entre falantes do nosso idioma e falantes de outros países em que se fala a língua portuguesa.

As alternâncias que os falantes de uma língua realizam podem ocorrer devido à preferência do usuário ou a pressão do sistema. Dessa maneira, a escolha dos usuários ou o próprio sistema linguístico podem fazer com que uma das escolhas possa sobressair-se sobre a outra, fazendo com que a outra variedade desapareça ocasionando assim, uma mudança linguística.

Bisol & Brescancini (2002) esclarecem que:

A ideia de que a variabilidade é uma característica inerente a qualquer sistema linguístico conduz naturalmente à busca por uma explicação para o fato de o falante, ou grupo de falantes, efetuar uma determinada *escolha* e não outra. (BISOL & BRESCANCINI, 2002, p.14).

Neste estudo, vamos tratar de um dos processos de estruturação silábica que, Seara, Lazzarotto-Volcão e Nunes (2011 apud ROBERTO, 2016, p. 118) definem:

Os processos de assimilação são entendidos como aqueles em que um segmento se torna semelhante a outro, assumindo traços de um segmento vizinho... Os processos de estruturação silábica são aqueles em que ocorre mudança na distribuição dos elementos silábicos, seja por acréscimo, inversão ou supressão... Os processos de enfraquecimento e reforço envolvem modificações de elementos conforme sua posição na palavra... Por fim, processos de neutralização ocorrem quando dois segmentos distintos perdem suas diferenças em determinados contextos. (ROBERTO, 2016, p. 118).

3.2.1 A redução silábica de *PARA*

Ao realizar a produção da variante *PRA* ao invés da forma *PARA*, o falante está produzindo uma redução silábica, pois apaga um segmento dentro do vocábulo padrão

PARA, mas não fere a escala de soância (Clements, 1990 apud AMARAL, 2002) que propõe a ordem preferida de segmentos dentro da sílaba:

A partir de uma escala de soância, os segmentos com posição mais alta ficam no núcleo da sílaba, e os segmentos com posição mais baixa ficam nas margens. A escala tem a seguinte ordem: $O > N > L > G > V$ que indica serem os sons menos soantes os obstruintes e os mais soantes, as vogais. (AMARAL, 2002 apud BISOL & BRESCANCINI, p. 102).

Além disso, percebe-se que a sílaba *PRA* obedece às condições de boa-formação, já que, segundo Amaral (2000, p. 80-81) “O ataque do português permite, no máximo, dois elementos, sendo o primeiro uma oclusiva ou fricativa labial, e o segundo, uma soante não-nasal. Esta Condição de Ataque forma a sequência OL - obstruinte + líquida (vibrante simples ou lateral)”.

Assim fica a variante *PRA* na escala de soância: P (obstruinte), R (líquida), A (vogal) – $O > L > V$.

Escala de Soância

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|-----------------|
| O | < | N | < | L | < | G | < | V | |
| - | | - | | - | | - | | + | Silábico |
| - | | - | | - | | + | | + | Vocóide |
| - | | - | | + | | + | | + | Aproximante |
| - | | + | | + | | + | | + | Soante |
| 0 | | 1 | | 2 | | 3 | | 4 | Grau de soância |
| P | | | | R | | | | A | |

A redução da preposição *PARA* em sua variante *PRA* não fere a escala de Clements citada por Amaral (2002 apud BISOL & BRESCANCINI, 2002), já que a sequência cresce em direção ao núcleo: 0+2+4 formando um ataque complexo próprio da língua.

Amaral (2000, p. 148) afirma, em sua análise sobre a síncope nas proparoxítonas, que das duas líquidas (lateral e vibrante simples) que formam a segunda posição do ataque complexo, a vibrante é o fator que mais efeito provoca no

apagamento da postônica não-final: *víspora* > *vispra*, já que a lateral em *pétala* > *petla* está próxima do ponto neutro, o que sugere ser a vibrante uma melhor formadora de ataque complexo do que a lateral.

A autora buscou apoio à sua hipótese em outras escalas, uma vez que a de Clements (acima) coloca ambas as líquidas em igualdade. Tanto em Kiparsky (1979), Goldsmith (1990), Bonet e Mascaró (1996) a vibrante apresenta um valor de soância maior do que a lateral, pois aquela está mais próxima das vogais.

Em Bonet e Mascaró (1996, p.4 apud AMARAL, 2000, p. 149), por exemplo, na escala para o catalão, a vibrante simples (r-fraco) equivale ao flap, anexo aos glides, com valor maior de soância do que a lateral:

Escala de Soância para o Catalão

0 1 2 3 4 5

oclusivas – vibrante, fricativas – nasais – laterais – flap, glides – vogais

Com isso, em nossa pesquisa, queremos mostrar que o apagamento da vogal *a* na primeira sílaba de *PARA* provoca uma redução silábica *PRA* que favorece o surgimento de um grupo consonantal bem-formado, por atender às restrições fonotáticas da língua.

No entanto, a gramática padrão, tão usada como referência nas aulas de Língua Portuguesa no ensino escolar, desconsidera a produção da variante *PRA* no lugar da forma *PARA* e, devido a esse fator, achamos de suma importância um estudo que verifique a produção da variante *PRA* na escrita, já que seria considerado um “erro” gramatical.

Assim, temos a representação da construção silábica da variação da forma *PARA* na variante *PRA* da seguinte maneira: *PARA* (2 sílabas), com o padrão silábico CV-CV – *P’RA* (1 sílaba), com o padrão silábico CCV. Podemos perceber que a preposição *PARA* passa por um processo de supressão que origina a sua variante *PRA*. Temos então, uma palavra dissílaba transformada para uma monossílaba, como se vê abaixo:

| | |
|---------|-------|
| PA RA | P R A |
| CV + CV | CCV |

Quanto à acentuação, a preposição *PARA* tem então duas sílabas, e o acento recai na primeira sílaba. Já sua variante *PRA* tem apenas uma sílaba e perde o acento, tornando-se um monossílabo átono. Conforme Bisol (2010), as partículas átonas são chamadas clíticos, pois são elementos com acentuação fraca, ou seja, dependem do acento primário da palavra adjacente com que fazem associação. Para a autora, os clíticos seriam palavras monossilábicas sem acentuação que podem sofrer processos fonológicos, como elevação vocálica ou processo de sândi.

A forma reduzida da preposição *PARA*, encaixa-se na situação citada acima, pois ao reduzirmos a norma *PARA*, tal forma pode apresentar as variantes *PRA* e *PA*. Dessa maneira, *PARA* não seria um clítico, mas *PRA* e *PA*, sim.

Além disso, os clíticos não pronominais podem passar por um *processo de fusão*, o que pode ocorrer nos casos entre clíticos e artigos definidos e indefinidos como: (pra+o = pro, pra + a = pra), como temos nos exemplos das produções dos escolares presentes no *corpus* desta Dissertação.

Ex: *motivos pro nosso país; apresentem pra sociedade; veio pra nossa sociedade; pro pão de cada dia; não olha pro povo; e mostre pros outros (...).*

3.3 DA FALA PARA A ESCRITA

Marcuschi (2007) aponta para duas distinções que tratam das relações da língua falada e da língua escrita. Segundo o autor, existe: “(a) de um lado, oralidade e letramento e (b) de outro lado, fala e escrita” (2007, p.25).

Para Marcuschi (2007), oralidade e letramento tratam das práticas sociais, e fala e escrita tratam de uma distinção entre modalidades de uso da língua. O autor relata que a oralidade é uma interação social praticada pelos indivíduos com o intuito de comunicar, ela apresenta vários gêneros textuais que se fundam na sonoridade. A oralidade acontece desde uma realização mais informal a mais formal diante dos variados contextos de usos possíveis de serem produzidos por meio da linguagem humana.

Em relação ao letramento, o autor aponta que o mesmo envolveria diferentes práticas de escrita dentro da sociedade. O letramento aconteceria desde a apropriação mínima da escrita (ex: uma pessoa analfabeta que sabe qual ônibus pegar), até uma apropriação bem mais profunda (ex: um escritor de romances). Para o autor, “Letrado é o indivíduo que participa de forma significativa de eventos de letramento e não apenas aquele que faz um uso formal da escrita”. Marcuschi (2007, p. 25).

A fala é apontada como forma de produção textual-discursiva com finalidade de comunicação diante da oralidade. A fala é caracterizada pelo uso da língua na forma de sons em um sistema articulado e com significância, além da gestualidade, movimentos corporais e também a mímica.

Marcuschi (2007) descreve que a escrita é uma modalidade de produção textual-discursiva com finalidade de comunicação que abrange algumas especificidades materiais que se caracteriza devido à constituição gráfica, ou seja, para o autor, a escrita situa-se dentro do plano de letramentos, e é um modo de uso da língua que complementa a fala.

Em certo sentido, a distinção entre *fala* e *escrita* aqui sugerida contempla, de modo particular aspectos formais, estruturais e semiológicos, ou seja, os modos de representarmos a língua em sua condição de código. São os aspectos *sonoro* e *gráfico* que contam de modo essencial neste caso. Note-se, no entanto, que o aspecto gráfico não está aqui sendo equiparado a uma de suas formas de realização, isto é, a forma alfabética, pois a escrita abrange todos os tipos de escrita, sejam eles alfabéticos ou ideográficos, entre outros. (MARCUSCHI, 2007, p. 26).

Marcuschi (2007) aponta para uma ampliação de sua visão ao que foi dito acima: engloba na fala todas as manifestações textuais da modalidade oral e engloba na escrita todas as manifestações textuais da modalidade escrita.

Assim, segundo o autor, “**fala e escrita** são usadas para designar formas e atividades comunicativas, não se restringindo ao plano do código. Trata-se muito mais de processos e eventos do que produtos”. Marcuschi (2007, p. 26).

3.4 ESTUDOS SOCIOLINGUÍSTICOS

Dentre os estudos sociolinguísticos pesquisados com a variação entre *PARA* e *PRA*, foram referenciados para este trabalho três da oralidade e dois da escrita: SILVA (2010), FERREIRA (2014 - 2018), FRANCO (2017) e NASCIMENTO & VIEIRA (2019).

A maioria dos trabalhos realizados com o tema desta pesquisa tem a sua análise oriunda da oralidade, assim, já que o *corpus* deste trabalho é proveniente da escrita em sala de aula, pensamos ser necessária uma abordagem na escrita também, para que haja um retorno ao meio escolar e social.

Os únicos trabalhos realizados na escrita, de que se têm conhecimento, são os de Franco (2017) “A Preposição *PARA* e sua variante *PRA* na escrita de alunos do Primeiro Ano do Ensino Médio”, que terá seus resultados comparados com os desta pesquisa, e o de Nascimento & Vieira (2019), “Um estudo da variação entre *para* e *pra* em verbetes publicados no *Instagram*”, publicado na revista Coralina, em fevereiro de 2019, no qual as autoras realizam uma pesquisa da variação entre *PARA* e *PRA* nos verbetes de Edgard Abbelhusen, publicados no *Instagram* (Fotocitando).

3.4.1 O estudo de Franco (2017)

O estudo de Franco (2017) analisou a preposição *PARA* e da sua variante *PRA* na escrita de seis alunos, três meninas e três meninos, com idade entre quatorze e dezesseis anos, cursando o primeiro ano do ensino médio. A pesquisa realizou-se dentro da Teoria Sociolinguística Laboviana e teve como *corpus* as produções textuais coletadas em sala de aula por meio de cinco gêneros textuais: e-mail, autobiografia, relato de passeio escolar, crônica da *internet* e teatro.

Em Franco (2017) foram analisados os seguintes grupos de fatores: posição da variante no texto (inicial, intermediária e final); gênero textual; formalidade do texto

(mais formal, menos formal); sexo (feminino e masculino).

De um total de 196 produções e reescritas textuais, 135 favoreceram a forma *PARA* e 61 favoreceram a variante *PRA*. Os grupos de fatores que favoreceram mais a variante *PRA* foram: Gênero Textual (teatro), Formalidade do texto (reescrita menos formal), posição da variável no texto (final) e Sexo (masculino).

Os resultados, segundo a autora, mostram que há variação na escrita formal dos alunos, em relação às variantes *PARA* e *PRA*. A forma padrão *PARA* apresentou um bom índice de ocorrência, mas é perceptível a presença da variante *PRA* nos textos escolares, especialmente nos que remetem à oralidade e/ou linguagem da *internet*.

3.4.2 O Estudo de Nascimento & Vieira (2019)

No estudo de Nascimento & Vieira (2019), as autoras dividiram os verbetes em *nomes* e *signos* (onde *nomes* são uma tentativa de interação por meio da escrita do autor com o público e *signos* são a descrição do horóscopo do público feita pelo autor), apontando para questão de que, apesar de o verbebo ser um tipo textual normalmente apresentado na norma padrão de linguagem, nos verbetes *nomes*, o modo de aproximação e intimidade com que o autor (Edgard Abbelhusen) propõe abordá-los com o seu público, em sua página no *Instagram*, faz com que apareça muito mais a presença da variante *PRA* ao invés da norma *PARA*.

Já nos verbetes *signos* Edgard faz mais uso da linguagem padrão, pois assim, não há a tentativa de aproximação com determinado público, e sim, uma escrita formal, sem interesse de intimidade com o leitor, por parte do autor. Vejamos o que dizem as autoras:

Percebe-se de modo geral que o *para* é mais formal e bastante utilizado na escrita, já o *pra* é característico da oralidade e está presente na fala dos indivíduos e há também situações em que ele aparece na escrita. Partindo da hipótese de que o uso do *pra* pode não ser a falta de conhecimento da língua padrão, mas uma forma de tornar o texto mais próximo da fala das pessoas, da oralidade, tornando o texto mais descontraído, pensou-se em discutir o uso da variação *para/prá* em verbetes. (NASCIMENTO & VIEIRA, 2019, p. 2).

Dessa maneira, as autoras chegaram à confirmação de suas hipóteses de que, ao usar a variedade não padrão *PRA*, há por trás uma intenção de aproximação da linguagem oral das pessoas, em uma tentativa de aproximação e intimidade.

O estudo de Nascimento & Vieira (2019) não será comparado com os resultados deste trabalho, devido ao fato de não apresentar uma análise estatística.

3.4.3 O estudo de Silva (2010)

Embasada na Sociolinguística de Labov (1966), Silva (2010) aponta para a disparidade entre a língua falada e a escrita em nosso país. A autora afirma que no Brasil, o nivelamento da língua não é possível devido a diferentes fatores, tais como: sociais, históricos, geográficos e culturais, e que é necessário um entendimento do processo evolutivo da língua para uma compreensão melhor da realidade linguística.

Em seu trabalho, Silva (2010) comenta sobre a variação entre *PARA* e *PRA*:

As gramáticas normativas, base do estudo nas escolas, conceituam a preposição como uma classe gramatical invariável que estabelece relações de sentido e de dependência, porém em sua maioria não incluem as formas não padrão do “para”, ignoram a variante “pra”. (SILVA, 2010, p. 17).

Em seu estudo “A PREPOSIÇÃO *PARA* E SUAS VARIANTES NO FALAR ARAGUATINENSE”, Silva (2010) aborda a preposição *PARA* com o intuito de analisar o que ocorre no uso da língua falada em Araguatins –TO, em relação à variação. Para a realização de seu trabalho, a autora utiliza o *corpus* coletado pelo Projeto Variação Linguística no Estado do Tocantins – VALTINS.

Silva (2010) buscou mostrar as variações existentes, a partir da análise das influências dos fatores linguísticos e extralinguísticos no processo de escolhas das variantes, aqui duas: a variante *PRA* e a variante *PA*. Em relação às variáveis linguísticas que foram analisadas no trabalho de Silva (2010), os resultados apontaram que, de um total de 3.210 ocorrências, 35 foram da variante *PARA*, 1.852 da variante *PRA* e 1.323 da variante *PA*.

As variantes que exerceram papel principal na pesquisa de Silva (2010) pelo programa GOLDVARB foram: *Contexto Fonológico Seguinte, Escolarização, Faixa Etária e Sexo*. Em *Contexto Fonológico Seguinte*, os fatores que favoreceram a variante *PRA* (valores acima de 0.50) foram: dorsais, vogais anteriores e as vogais posteriores.

Em relação às mesmas variáveis os fatores *vogal central, labial e coronal* favoreceram menos a variante *PRA* (resultados abaixo de 0.50) e é a vogal central a que mais favorece a aparição de *PA*. Segundo Silva (2010), os resultados indicam que a escolha de *PRA* ou *PA* (redução de uma sílaba do vocábulo *PARA*) é influenciada pela proximidade do ponto labial.

Os fatores *Escolarização e Faixa Etária* foram decisivos em relação às variáveis sociais, pois mostraram valores consideráveis. Em relação à escolaridade mais elevada (+ 49 anos) há predominância da variante *PRA*, e nos discursos dos falantes de menor escolaridade (15-25) esse número cai.

O fator *Sexo* não mostrou muita força, mas o uso da variante *PRA* mostrou uma fraca vantagem com relação às mulheres (996/1582 – 63% - 0.53), homens (856/1593 – 53% - 0.46). A autora acredita que esse resultado acontece devido às mulheres estarem igualando-se aos homens quanto aos papéis sociais.

Segundo Silva (2010), os fatores descartados foram: *Presença de Vibrante no Item Seguinte - Paralelismo Formal*. A respeito dos resultados obtidos em sua pesquisa, a autora comenta:

O fenômeno da variação da preposição *PARA* tem um caráter abrangente, atingindo, talvez, a totalidade do território brasileiro. Os surpreendentes resultados obtidos, nesta pesquisa, atestam que no uso da fala do araguatinsense, hoje, sinaliza um desaparecimento desta variante na oralidade. Com relação às outras variantes, *PRA* e *PA*, destacou-se um fenômeno de variação estável na fala do nativo de Araguatins – TO. (SILVA, 2010, p. 64).

3.4.4 O estudo de Ferreira (2014)

Em outro estudo sobre a preposição *PARA*, Ferreira (2014) teve como objetivo investigar a variação da preposição *PARA/PRA/PA* na fala da cidade de Londrina/PR. A autora buscou verificar se há motivações linguísticas e/ou sociais para a ocorrência a

respeito do fenômeno da variação.

Segundo a autora: “O percurso histórico da preposição *para* mostra que sua variante *pra* tem sido realizada há bastante tempo, principalmente na fala, mas também na escrita informal”. (FERREIRA, 2014, p. 14). E ainda, que segundo Williams (1961, p. 71), já no século XVI a modificação *para* < *pra* ocorria devido ao ritmo do verso de Sá de Miranda e Gil Vicente. (FERREIRA, 2014, p. 14).

Sob a luz da Variação Fonológica, Ferreira (2014) nos diz que:

A preposição *para* parece sofrer uma perda fonética quando realizada como *pra* ou *pa*, comumente utilizadas na forma falada. Nesta perda sonora, há também a supressão de uma sílaba e consequências para o acento. (FERREIRA, 2014, p. 14).

Em sua análise, de um total de 349 dados, foram obtidas 257 (73.6%) ocorrências de *PRA*, 90 (25.8%) de *PA* e somente duas (0.6%) de *PARA*. A hipótese da autora de que a ocorrência de *PARA* seria insignificante confirmou-se. Dessa maneira, Ferreira (2014) decidiu realizar uma análise binária, definindo como variável dependente a variante *PRA*.

Em relação à descrição das variáveis linguísticas que foram selecionadas pelo programa VARBRUL como possíveis condicionantes para a ocorrência da preposição *PRA*, foram selecionadas as seguintes variáveis: *Idade*, *Posição em relação a pausas e Sexo*. Ferreira (2014) considerou como aplicações de *PA* as ocorrências que não estão incluídas nos resultados da variável dependente *PRA*, pois a variante *PARA* obteve escassa quantidade de ocorrências na pesquisa da autora.

A variável linguística *Posição em relação a pausas* foi inicialmente dividida em contexto *Sem Pausas* e *Com Pausas*. O resultado obtido foi o seguinte: de um total de 246 aplicações *Sem Pausa* 170 (69%) dos casos favoreceram o aparecimento de *PRA* e de 52 aplicações do contexto *Com Pausas*, 46 (88.5%) dos dados favoreceram o aparecimento de *PRA*.

Os grupos *Contexto Morfológico Seguinte*, *Tonicidade da Sílaba Seguinte*, *Paralelismo*, *Número de Sílabas do Item Seguinte*, *Pausa*, *Processo de Sândi com a Sílaba Seguinte* e *Escolaridade* não favoreceram a variável *PRA* ou foram eliminados

pelo programa.

Quanto à variável Idade, a autora explica que a dividiu em – *de 50 anos e + de 50 anos*, faixa etária igual a do Projeto VARSUL, donde a amostra de seu trabalho foi retirada do *corpus* do respectivo Projeto.

Segundo Ferreira (2014), apesar de os informantes menores de 50 anos terem apresentado quantidades maiores de dados (de um total de 171, 111 favoreceram o uso de *PRA* – 65% - 0.39), foram os maiores de 50 anos que utilizaram mais a forma *PRA* (de um total de 127 aplicações, 105 favoreceram o uso do *PRA* – 82.7% - 0.64).

Em relação à variável *Sexo*, têm-se no trabalho de Ferreira (2014), uma maior influência ao percentual de *PRA* relativo às mulheres (de um total de 139, 111 favorecem o *PRA* – 80% - 0.62) contra o percentual dos homens em que de um total de 159, 105 favoreceram o uso do *PRA* – 66% - 0.39).

A autora acredita que este fato ocorra pela opção que as mulheres fazem em escolher variáveis linguísticas inovadoras, o que se confirma em seu estudo no caso das mulheres mais jovens. Porém, segundo Ferreira (2014), se essa variável for desprestigiada socialmente, as mulheres tendem a rejeitá-las.

3.4.5 O estudo de Ferreira (2018)

Ferreira (2018) faz um retrospecto da preposição *PARA* do Latim ao Português, além de uma abordagem do estatuto prosódico da respectiva preposição. Alicerçando sua pesquisa em relação à preposição *PARA* do Latim ao Português, relata que a preposição, a princípio, era denominada como uma partícula adverbial, precedente ao complemento, sendo vista também para a expressão de noções temporais, como: tempo, lugar, modo, e outras noções de mesma categoria.

Ao longo do tempo, as preposições tiveram seu valor acrescido por meio da gradativa queda das flexões gramaticais, passando a determinar também, relações de caso. Mesmo assim, segundo a autora, grande parte das preposições se perdeu devido ao

fato dos processos de simplificação e economia.

Em relação ao objeto de estudo, Ferreira (2018) investigou a variação da preposição *PARA* na fala das cidades de Curitiba (PR) e de Florianópolis (SC) com base na Teoria da Variação e Mudança Linguística buscando verificar as motivações linguísticas e/ou sociais existentes para a ocorrência do fenômeno variável.

Os resultados da pesquisa em relação à análise eneária foram os seguintes: de um total de 2591 dados, 2086 (80,5%) das aplicações apontaram para o uso da variante *PRA*, 456 (17,6%) apontaram para o uso de *PA* e apenas 49 (1,9%) dos dados apontaram para a variante padrão *PARA*.

Ferreira (2018) confirma sua hipótese de que a variante *PRA* seria predominante em sua pesquisa, dessa maneira, a autora optou por realizar uma análise binária definindo a variante *PRA* como variável dependente.

Para dar sequência em seu trabalho, Ferreira (2018) juntou as variantes *PARA* e *PA*, compreendendo que entre as duas há bastante diferença em termos segmentais e prosódicos, alegando que tal união não acarretaria, estatisticamente, interferência nos resultados da pesquisa.

A autora alega ter examinado os poucos casos de *PARA* manualmente antes de unir a variável padrão à variante *PA*. As variáveis selecionadas pelo programa foram: *Contexto Morfológico Seguinte*, *Contexto Fonológico Seguinte*, *Paralelismo*, *Idade*, *Cidade*, *Escolaridade* e *Número de Sílabas do Item Seguinte*.

Em relação ao *Contexto Morfológico Seguinte*, inicialmente com as categorias *Palavra Lexical Nominal* e *Palavra Lexical Verbal* (que após foi amalgamada, tornando-se *Palavra Lexical*) e a categoria *Palavra Gramatical* (88,4% - 0.65). A categoria *Palavra Gramatical* mostrou-se mais significativa para o uso de *PRA* (935/1058) do que a *Palavra Lexical* (75,8% - 0.37) e os resultados, segundo Ferreira (2018), podem indicar que o *PRA* não é condicionado em termos de peso relativo se for seguido por palavra lexical.

Em *Contexto Fonológico Seguinte*, a categoria *Vogal Coronal* (309/341 – 90,6%

- 0.65) apresentou mais favorecimento da variante *PRA*, após veio à categoria *Consoante Dorsal* (305/361 – 84,5% - 0.56). Além disso, para Ferreira (2018) apesar de uma menor quantidade de contextos, chamou a atenção o percentual de 97% de aplicação do uso de *PRA* antes de palavras monossílabas tônicas iniciadas por consoantes dorsais.

Em relação à variável *Paralelismo*, os resultados evidenciaram a preferência pela variante *PRA* quando é antecedida por ela mesma (374/422), com um percentual de 88,6% e peso relativo de 0.64. Apesar de haver uma quantidade maior de contextos em *Ocorrência Isolada/Primeiro da série* (1539/1927 – 80% - 0.46), é o fator *Antecedido de PRA* que tem percentual maior.

No grupo *Número de Sílabas do Item Seguinte*, o fator *Duas ou mais sílabas* mostrou-se mais favorável à realização da variante *PRA* (1334/1643 – 81,2% - 0.53) em relação ao fator *Uma sílaba* (579/706 – 82% - 0.43), o que pode segundo Ferreira (2018), representar uma preferência de *PRA* ocorrer antes de uma palavra com duas ou mais sílabas. Mesmo assim, segundo a autora, os pesos relativos de ambos encontram-se próximos (diferença de 0.10 entre si e pouca diferença do ponto neutro 0.50).

No grupo *Tonicidade da Sílaba Seguinte*, a autora cruzou esse grupo de fatores com o *Contexto Fonológico Seguinte* e os resultados mostraram que o uso de *PRA* é favorecido por palavras iniciadas por *Consoante Dorsal e Tônicas monossilábicas* (97%) - apesar de uma menor quantidade de contextos, e *Vogal Coronal e Tônica* (95%).

Em relação às variáveis extralinguísticas que foram controladas em sua pesquisa, à autora descreve os resultados por ordem de significância. Primeiramente é descrito o grupo *Idade*, que foi dividido em (- de 50 anos e + de 50 anos), onde, segundo a autora, apesar de os informantes com menos de 50 anos apresentarem maior quantidade de dados (993/1267 – 78,4% - 0.44), são os maiores de 50 anos (920/1082 – 85% - 0.56) que usam mais o *PRA*.

Em relação à escolaridade, relata que o fator *Até 11 anos de estudo* favorece o

aparecimento da variável *PRA* (1132/1350 – 84% - 0.53), já o fator *Até 4 anos de estudo* (781/999 – 78,2% - 0.45), além de apresentarem menos contextos de ocorrências, favorece menos a aplicação de *PRA*.

Por fim, Ferreira (2018) analisa o grupo *Cidade*, que indica preferência pelo uso de *PRA* na capital de Curitiba (PR), com resultados de um total de 1143, 965 para *PRA* (84,4% - 0.55), enquanto que para Florianópolis (SC), esses dados são de um total 1206, 948 favorecem o *PRA* (78,6% - 0.45).

Após a realização do detalhamento dos estudos referenciados neste trabalho, procederemos, a seguir, com a metodologia, as hipóteses e os objetivos da pesquisa.

4 METODOLOGIA

A metodologia segue os pressupostos da Sociolinguística Quantitativa de William Labov (2008[1972]) e utiliza o programa estatístico GOLDVARB X, para ver a possível influência dos fatores linguísticos e extralinguísticos que podem intervir ou não, a variação no fenômeno estudado.

Segundo Onofre (2001), o programa estatístico GOLDVARB foi desenvolvido por Rand & Sankoff (1990), sendo uma opção de escolha do pacote do VARBRUL (oriundo do inglês Variable Rules Analysis), que funciona como um conjunto de programas para computador de múltipla análise. Assim, o GOLDVARB X é uma alternativa que atua no Windows para aclimatar dados referentes à variação Sociolinguística.

É a partir do programa GOLDVARB X que as hipóteses da pesquisa são ou não confirmadas.

Assim, se um fenômeno linguístico tem seus grupos de fatores apontados como não significativos pelo programa, a hipótese é rejeitada / se os grupos de fatores são significativos, mas a influência não é como se previu no valor da aplicação, a hipótese também é rejeitada / se os grupos de fatores são significativos e a influência dos fatores é como a prevista no valor de aplicação, a hipótese é confirmada. (ONOFRE, 2001, p. 1).

Dessa maneira, usamos o GOLDVARB X na presente pesquisa, a fim de aprofundar nossa análise quantitativa através de resultados oriundos de rodadas de dados que foram elaboradas no referente programa para a confirmação das hipóteses.

4.1 COMPOSIÇÃO DA AMOSTRA

Os sujeitos que compõem a amostra são estudantes (83 meninas e 69 meninos, num total de 152) cursando os primeiros anos (60 meninas e 61 meninos, num total de 121) e o terceiro ano (23 meninas e 8 meninos, num total de 31) do Ensino Médio (com idade entre catorze e dezoito anos), de uma escola da rede pública estadual, localizada em um bairro próximo ao centro da cidade de Rio Grande/(RS), local onde a autora desta pesquisa atua como professora, desde o ano de 2012.

As 1116 ocorrências são oriundas de produções textuais das aulas de Literatura da autora deste trabalho e de outras duas profissionais da escola, ministrantes da

disciplina de Língua Portuguesa, que orientaram as turmas na elaboração textual e nos critérios de avaliação dos textos.

Os textos em *prosa*, conforme orientação da professora dos primeiros anos, tiveram um caráter argumentativo, pois a professora escolheu temas atuais em destaque para que os escolares se posicionassem criticamente a respeito dos assuntos tratados.

Com relação ao terceiro ano, a professora da turma pediu produções dissertativas também com temas atuais, em formato de produções para o ENEM e vestibulares.

Assim, somente os textos em *verso* tiveram realmente um caráter mais lúdico, pois foram aplicados nas aulas de literatura, pela autora desta dissertação, e foram realizados a partir de uma palestra que tratava do dia mundial da Consciência Negra, onde os escolares produziram versos livres a respeito da temática abordada.

Todas as produções foram realizadas nas salas de aula com as professoras das respectivas turmas e disciplinas presentes.

4.2 VARIÁVEIS CONTROLADAS

As variáveis controladas nesta pesquisa serão apresentadas a seguir:

4.2.1 Variável Dependente

Neste estudo, a variável dependente é a preposição *PARA* e sua variante *PRA* que foi escolhida como valor de aplicação para ser controlada na análise estatística.

4.2.2 Variáveis Independentes

Os fatores linguísticos internos que foram considerados são: *Contexto morfológico seguinte*, *Contexto fonológico seguinte*, *Tonicidade da sílaba seguinte*, *Número de sílabas do item seguinte* e *Tipo textual*. Foram considerados também, os seguintes fatores extralinguísticos: *Sexo*, *Faixa etária* e *Nível de escolaridade*, já que teremos turmas de níveis diferentes de aprendizagem (1º anos e 3º ano).

Cada grupo foi analisado de maneira minuciosa para a averiguação de fatores que favoreceram ou não, o surgimento da variante *PRA*, a forma inovadora, mais presente na fala do que na escrita.

4.2.2.1 Grupo de fatores linguísticos

a) *Contexto morfológico seguinte*: trata da aparição da variante *PRA* antes de palavras específicas, ou seja, de classes de palavras. Exemplo:

PRA antes de *Pronome*: *pra você, pra esse, pra tu, pra mim...*

PRA antes de *Outros*: *prá lá, pro (pra+o); pra que, pra não...*

PRA antes de *Substantivo*: *pra sociedade, pra adulto, pra guerra, pra mãe...*

PRA antes de *Verbo*: *pra chamar, pra estudar, pra dormir, pra nascer...*

b) *Contexto fonológico seguinte*: trata da ocorrência da variante *PRA* antes de diferentes segmentos. Exemplos:

PRA antes de *Fricativas e lateral*: *pra sustentar, pra você, pra lá, pra sociedade, pra chamar...*

Pra antes de *Vogais altas e não altas*: *pra esse, pro (pra+o), pra adulto, pra*

estudar...

Pra antes de Oclusivas: pra casa, pra tu, pra que, pra guerra, pra dormir...

Pra antes de Nasal: pra mim, pra não, pra mãe, pra nascer...

c) *Tonicidade da sílaba seguinte*: trata de analisar a força na produção da próxima sílaba da palavra seguinte ao surgimento da variante *PRA*. Neste item analisamos se a força produzida pela primeira sílaba da palavra seguinte à ocorrência da variante *PRA*, influenciou ou não o surgimento de tal variante nas produções dos escolares.

d) *Número de sílabas do item seguinte*: averiguamos se a extensão da palavra seguinte, com maior ou menor número de sílabas, pôde ser um fator de influência na escolha dos estudantes em relação ao não uso da forma *PARA*.

e) *Tipo textual*: trata da verificação se o surgimento da variante *PRA* aconteceu mais nas produções em *prosa* ou nas produções em *verso*.

4.2.2.2 Grupo de fatores extralinguísticos

a) *Sexo*: trata de um fator que permitiu averiguar se o sexo dos sujeitos que compõem a amostra causou influência na escolha dos objetos de análise da pesquisa.

b) *Faixa etária*: trata de analisar se a idade dos estudantes influenciou na ocorrência da proposição *PARA* ou de sua variante *PRA*. Como a pesquisa analisou turmas de primeiros e terceiro ano, a faixa etária dos sujeitos que compõem essa pesquisa são diferentes, o que permitiu uma análise das ocorrências diante da idade dos estudantes.

c) *Nível de escolaridade*: trata de verificar se o ano em que se encontram os sujeitos da amostra influenciou na escolha do padrão gramatical *PARA* ou de sua variante *PRA*. Na análise desta pesquisa, os estudantes cursam anos diferentes, os primeiros e o terceiro ano do Ensino Médio.

4.3 OBJETIVOS DO ESTUDO

É de interesse da pesquisa, observar o tema deste trabalho, uso das formas *PARA* e *PRA*, na escrita de alunos de cinco turmas de Ensino Médio, a fim de trazer respostas e possibilidades de trabalho diante do fenômeno analisado.

O estudo em questão tem como propósito a análise da preposição *PARA* e sua variante *PRA* na escrita de escolares, através de uma pesquisa quantitativa com coleta de dados realizada em uma escola pública, tendo como *corpus* os textos produzidos em *verso* e *prosa*.

Para Faraco & Moura, na obra *Língua e Literatura*, a realização de atividades com diferentes textos, “permite trabalho constante de intertextualidade, além de variadas formas de leitura, visando permitir que o aluno possa articular diferentes linguagens e seus contextos” (Faraco & Moura, p. 3, 2002).

4.3.1 Objetivo geral

O trabalho tem como objetivo geral investigar o uso da preposição *PARA* e de sua variante *PRA* no contexto escolar, na modalidade escrita, em cinco turmas de uma escola pública, do Ensino Médio, na cidade de Rio Grande – RS.

4.3.2 Objetivos específicos

Os objetivos específicos são:

- a) analisar que as variáveis linguísticas e extralinguísticas da preposição *PARA* e sua variante *PRA* que poderiam afetar o fenômeno estudado;
- b) determinar uma frequência de ocorrência *PARA* e *PRA* nos textos escritos por estudantes dos 1ºs anos e do 3º ano do Ensino Médio;
- c) verificar se há diferença na ocorrência da variável *PARA* e de sua variante *PRA* nos textos escritos em *verso* e *prosa*.

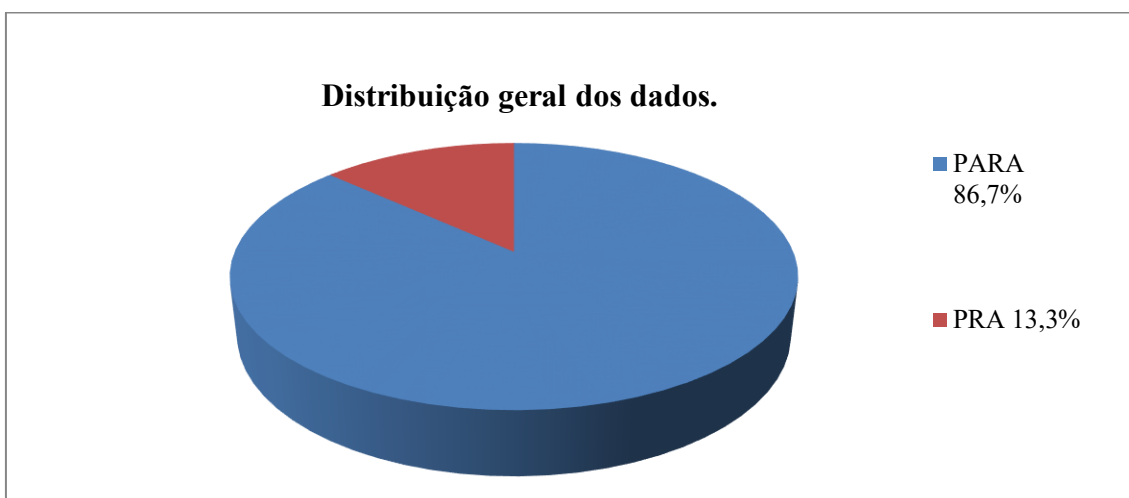
5 RESULTADOS DA ANÁLISE ESTATÍSTICA¹

A primeira rodada de dados realizada, através do programa GOLDVARB X, apresentou dois *knockouts* em seu resultado. As classes morfológicas: *Numeral* e *Preposição*, presentes no grupo de fatores *Contexto morfológico seguinte*, apresentaram 100% de aplicação para o *PARA* e 0% para a sua variante *PRA*, respectivamente. Dessa maneira, decidimos eliminar ambas as classes (*knockouts*) e fazer uma nova rodada de dados para averiguar a influência ou não, dos fatores internos e externos no fenômeno pesquisado.

5.1 DISTRIBUIÇÃO DAS VARIANTES *PRA* E *PARA*

Na análise deste estudo, de um total de 1116 dados, 968 (86,7%) ocorrem com a forma *PARA* e apenas 148 (13,3%), com a variante *PRA*. Vejamos a situação ilustrada no gráfico a seguir:

Gráfico 1– Distribuição geral dos dados.



¹ Os dados de Nascimento & Vieira (2019) não serão comparados com os desta pesquisa, pois as autoras não fizeram o detalhamento dos mesmos em seu artigo.

Comparando com as pesquisas abordadas nesta dissertação, os resultados podem ser vistos na Tabela 1:

Tabela 1- Distribuição dos dados por variante em cada pesquisa.

| | Variável Dependente | <i>Para</i> Apl. % | <i>Pra</i> Apl. % | <i>Pa</i> Apl. % | Nº de Dados |
|-----------------|----------------------------|-----------------------|----------------------|---------------------|--------------------|
| Silva (2010) | <i>Pra</i> | 35 1 | 1852 54 | 1323 45 | 3210 |
| Ferreira (2014) | <i>Pra</i> | 2 0,6 | 257 73,6 | 90 25,8 | 349 |
| Franco (2017) | <i>Pra</i> | 135 61 | 61 39 | / / | 196 |
| Ferreira (2018) | <i>Pra</i> | 49 1,9 | 2086 80,5 | 456 17,6 | 2591 |
| Franco (2020) | <i>Pra</i> | 968 86,7 | 148 13,3 | / / | 1116 |

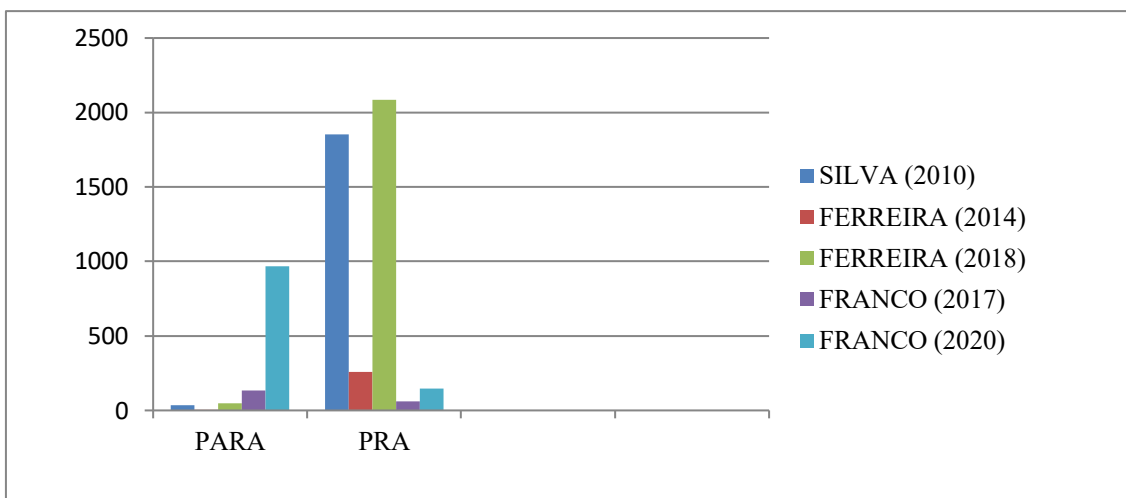
Podemos ver, na Tabela 1, que Silva (2010) e Ferreira (2014-2018) analisaram também a variante *PA* na fala de seus informantes. Os resultados apontam que todas as pesquisas obtiveram percentual muito baixo da norma *PARA*, com exceção das pesquisas oriundas da escrita realizadas por Franco (2017 – 2020), com 61% e 86,7% dos dados referentes ao padrão *PARA* respectivamente.

A variante *PRA* teve preferência na maioria dos trabalhos. Os resultados mais próximos em favor de tal variante foram os de Ferreira (2014) – aproximadamente 73% e Ferreira (2018) – aproximadamente 80,5% dos dados.

Acredita-se que o baixo percentual obtido da variante *PRA* nos trabalhos de Franco (2017 - 2020) é proveniente de os estudos serem oriundos da escrita, onde há mais formalidade linguística. Além disso, todas as produções em *prosa* e *verso*, que serviram de material de apoio para esta pesquisa, assim como todos os textos coletados por Franco (2017) foram escritos em salas de aula, ou seja, foram criados dentro do contexto escolar, o que pode fazer com que os alunos sintam-se mais pressionados a usar a linguagem padrão, na dúvida em relação ao uso da norma ou da variante *PRA*.

Podemos visualizar melhor, no gráfico a seguir, a comparação de dados entre todas as pesquisas.

Gráfico 2 – Comparação dos dados entre pesquisas da variável *PARA*.



O Gráfico 2 ilustra a distribuição dos resultados de *PARA/PRA* em cada pesquisa, expostos na Tabela 1.

Com exceção de Franco (2017 - 2020), cujos dados são originários da escrita, pode-se observar, pelo Gráfico acima, que a variante *PRA* é a preferida na modalidade de fala.

5.2 VARIÁVEIS LINGUÍSTICAS E EXTRALINGUÍSTICAS SIGNIFICATIVAS PARA O USO DA VARIANTE *PRA* NA MODALIDADE ESCRITA

Após terem sido retirados os *knockouts* e feitas às amalgamações que achamos pertinentes nos grupos de fatores *Contexto morfológico seguinte* **Outros: conjunções, advérbios e artigos (definidos e indefinidos)*, **Pronomes: pronomes pessoais (oblíquo e reto), pronomes indefinidos, pronomes demonstrativos, pronomes possessivos e pronomes de tratamento*); *Contexto fonológico seguinte* **Fricativas e lateral: fricativas, lateral e vibrante*, **Vogais altas e não altas: vogais altas, vogal baixa e vogais médias*, foi realizada uma nova rodada de dados.

Na segunda rodada, os resultados apontaram duas variáveis em sobreposição² (*Nível de escolaridade* e *Faixa etária*), ao realizarmos uma rodada sem o grupo de fatores (*Faixa etária*) e outra rodada sem o outro grupo de fatores (*Nível de*

² Falaremos melhor adiante.

escolaridade), confirmamos a sobreposição. Assim, achamos que a sobreposição foi o que fez o Programa selecionar apenas um grupo de fatores (*Nível de escolaridade*) e excluir o outro (*Faixa etária*).

Dessa maneira, foram selecionados como favoráveis ao aparecimento da variante *PRA* os grupos que vêm a seguir:

- (i) *Tipo textual*
- (ii) *Nível de escolaridade*
- (v) *Sexo*
- (iv) *Contexto fonológico seguinte*

5.2.1 *Tipo textual*

A partir da rodada no programa computacional apontamos o grupo *Tipo textual* como o mais relevante para a aplicação da variante *PRA*. Tal grupo foi dividido em dois fatores: *prosa* e *verso*, e o material coletado para esta pesquisa foi todo produzido dentro desses dois tipos de modalidades textuais.

Na tabela 2, vemos que o contexto que mais favorece a variante *PRA* é a produção em *verso*.

Tabela 2 – Variante *PRA* considerando o *Tipo textual*.

| GRUPO | APLICAÇÃO/TOTAL | % | PESO RELATIVO |
|-------|-----------------|------|---------------|
| Prosa | 103/1001 | 10,3 | 0.45 |
| Verso | 45/115 | 39,1 | 0.83 |

Input : 0.098 Significância: 0.001

Podemos observar pela tabela 2 que o fator *verso* mesmo tendo uma menor quantidade de dados (45/115) é o que favorece mais a variante *PRA* (0.83), em oposição ao fator *prosa* (0.45), que desfavorece o surgimento de tal variante.

Acreditamos que este resultado ocorra devido ao fato de as produções em *verso* permitirem uma maior influência da oralidade, a fim de uma busca pelo melódico e também pela possibilidade de recitação. Tais produções foram solicitadas por meio de

versos livres, ou seja, não havia rigidez na construção dos versos, permitindo aos alunos uma maior flexibilidade na linguagem escrita.

Através dessa modalidade, podemos confirmar as hipóteses de que as variantes *PARA* e *PRA* iriam se comportar de forma distinta nos diferentes tipos de textos e também que os alunos considerariam a escolha das variantes a partir do grau de formalidade dos textos propostos.

Em Franco (2017) foram analisados *cinco gêneros textuais* diferentes (*Autobiografia, Relato, E-mail, Crônica da Internet e Teatro*). Tal pesquisa mostrou que as variantes *PARA* e *PRA* são sensíveis ao *gênero textual* em que os textos dos estudantes foram produzidos.

Os resultados da pesquisa mostraram que o uso da norma *PARA* está mais envolvido com os gêneros em que os alunos produziram o texto: *Autobiografia* (0.89) e *Relato* (0.84). O uso de *PRA*, segundo Franco (2017) mostrou-se mais frequente nos gêneros textuais em que havia uma *linguagem da internet* como foco: *E-mail* (0.60) e *Crônica da Internet* (0.47).

Por fim, o gênero *Teatro* (0.33) que é proveniente de uma atuação da oralidade, pouco influenciou na escrita dos escolares para o uso da variante menos formal *PRA*.³

5.2.2 Nível de escolaridade⁴

Averiguamos que o grupo de fatores *Nível de escolaridade* foi o primeiro grupo de fatores social selecionado como relevante para o processo de realização da variante *PRA*.

³ Não foi possível comparar os resultados da variável *gênero textual* com as demais pesquisas: Silva (2010), Ferreira (2014 – 2018) - pelo fato de as autoras não a terem incluído em suas análises.

⁴ Ao analisarmos os resultados da pesquisa, percebemos que os grupos de fatores *Faixa etária* e *Nível de escolaridade* estavam sobrepondo-se um ao outro. Desta maneira, realizamos uma primeira rodada sem o grupo de fatores *Faixa etária* e logo após, uma nova segunda rodada sem o grupo de fatores *Nível de escolaridade*. Ambos os grupos mostraram-se relevantes para o resultado desta pesquisa e apresentaram resultados iguais na rodada de dados. Talvez seja esse o motivo de o Programa ter escolhido apenas um e descartado o outro.

Tabela 3 – Variante *PRA* considerando o *Nível de escolaridade*.

| GRUPO | APLICAÇÃO/TOTAL | % | PESO RELATIVO |
|-------------|-----------------|----------------------|---------------|
| 1ºs anos | 140/904 | 15,5 | 0.57 |
| 3º ano | 8/212 | 3,8 | 0.21 |
| Input 0.098 | | Significância: 0.001 | |

Como observamos na Tabela acima, os *1ºs anos* (15,5%) produziram mais a variante *PRA* do que o *3º ano* (3,8%), apesar de o fator *3º ano* ter menor quantidade de dados (8/212) em relação ao fator *1ºs anos* (140/904).

O que podemos pensar em relação ao grupo de fatores escolaridade, é que os escolares dos primeiros anos tendem a produzir mais a variante *PRA* por estarem chegando ao ensino médio e reproduzirem mais os reflexos da oralidade em seus textos, além do contato que têm com a linguagem da *internet*.

Dessa maneira, conforme os estudantes avançam de ano, vão tendo mais contato com a linguagem formal, através das exigências de construções textuais e adequação verbal/oral elaboradas nas atividades realizadas pelos professores.

Outro ponto a refletir é que estudantes de terceiros anos preocupam-se mais com a escrita formal devido à avaliação dissertativa exigida pelo ENEM⁵ para a entrada em uma Universidade.

Em seu trabalho, Silva (2010), analisou o grupo de fatores *Escolarização*, dividindo-o em três grupos: *Baixa* (0.37), *Média* (0.50) e *Alta* (0.61). A autora aponta que no discurso dos falantes de escolaridade mais elevada há uma predominância pela variante *PRA* (considerada pela autora como referência padrão em sua pesquisa), enquanto no discurso dos menos escolarizados o número baixa bastante.

No estudo de Ferreira (2014), a autora aponta que houve um favorecimento maior entre os informantes com *até 11 anos de escolaridade* pela variante *PRA* (0.57) em relação aos informantes com *até 4 anos de estudo* que apresentaram peso relativo de (0.45) para o aparecimento da variante *PRA* na oralidade. Os informantes com grau intermediário de escolaridade não apresentam preferência por nenhuma das duas variantes apontadas pela autora (*PRA* e *PA*), com peso relativo de (0.50) – ponto neutro.

Em Ferreira (2018) a autora alega que o grupo de fatores *Escolaridade* desempenhou um papel moderado, pois os valores dos pesos relativos de ambas as

⁵ Exame Nacional do Ensino Médio.

faixas etárias ficaram muito próximas do ponto neutro: *até 11 anos de estudo* (0.53), *até 4 anos de estudo* (0.45).

Mesmo assim, os informantes com *até 11 anos de estudo* apresentaram favorecimento maior para a aplicação da variante *PRA*.

Não foi possível comparar os resultados do grupo de fatores social *Escolaridade* desta pesquisa com o trabalho de Franco (2017), pois tal grupo de fatores não foi utilizado no trabalho da autora por seus dados serem todos referentes ao primeiro ano do ensino médio.

5.2.3 Sexo

No trabalho em questão, o *sexo* foi o segundo grupo de fatores social analisado que apontou o aparecimento da variante *PRA*. Vejamos os resultados na Tabela 4:

Tabela 4 – Variante *PRA* considerando o *Sexo* do informante.

| GRUPO | APLICAÇÃO/TOTAL | % | PESO RELATIVO |
|-----------|-----------------|------|---------------|
| Feminino | 69/715 | 9,7 | 0.42 |
| Masculino | 79/401 | 19,7 | 0.62 |

Input: 0.098

Significância: 0.001

Ao analisarmos os resultados do grupo de fatores *Sexo*, vimos que o *sexo masculino* (0.62) tende a favorecer mais o surgimento da variante *PRA* do que o *sexo feminino* (0.42), sinalizando uma preferência dos meninos pela forma inovadora *PRA* na escrita.

Podemos pensar diante do resultado obtido, que o *sexo feminino* tende a seguir mais as relações formais da escrita, e que o *sexo masculino* demora mais a ter percepção dessas relações, talvez por serem menos contidos do que as meninas ao desempenharem seus papéis sociais no período da adolescência. Assim, os meninos são mais “desligados” das regras e padrões sociais, como a formalidade na escrita.

No trabalho de Silva (2010), oriundo da oralidade, o resultado do grupo de fatores *Sexo* revelou a escolha da forma *PRA* entre as mulheres, com peso relativo de (0.53), contra (0.46) para os homens.

A autora aponta que, atualmente, os papéis sociais estão cada vez mais parecidos e que isso poderia ser a resposta para o resultado de sua pesquisa em relação ao *sexo* e a escolha de *PRA* como inovação na linguagem.

No trabalho de Ferreira (2014), o resultado em relação ao *PRA* também foi favorável às mulheres (0.62), contra (0.39) dos homens.

Quanto à variante *PRA* na modalidade oral, acredita-se que as mulheres optem mais que os homens por variantes inovadoras, o que se confirmou nos estudo de Ferreira (2014) em relação às mulheres mais jovens.

No trabalho de Franco (2017), que é proveniente da escrita, os resultados mostraram um peso relativo de (0.66) do uso de *PARA* na escrita das meninas contra um peso relativo (0.34) na escrita dos meninos do padrão formal de linguagem. Tal pesquisa obteve resultados favoráveis ao surgimento de *PRA* também por parte do *sexo masculino*.

Assim, em relação ao objeto de estudo desse trabalho (a preposição *PARA* e sua variante *PRA*), podemos pensar que: na escrita, o *sexo feminino* tende a optar pela forma padrão (*PARA*); já na oralidade, os resultados das pesquisas realizadas mostraram um favorecimento da forma inovadora *PRA* por parte das mulheres.

Contudo, na pesquisa de Ferreira (2014), ao cruzar os grupos de fatores *Idade* e *Sexo*, os resultados apontam que essa tendência é por parte das *mulheres mais jovens* (-de 50 anos: 79% dos dados). Assim, refletimos que o *sexo feminino* (mais jovem) segue os padrões normativos na linguagem escrita, mas é mais “despojado” na linguagem oral, apreciando e aderindo formas inovadoras, caso sejam socialmente prestigiadas (caso do *PRA*).

Não foi possível comparar os resultados do grupo de fatores social *Sexo* desta pesquisa com o trabalho de Ferreira (2018), pois tal grupo de fatores não foi utilizado no trabalho da autora.

5.2.4 Contexto fonológico seguinte

Este grupo de fatores foi dividido, inicialmente em: *vogal baixa*, *vogal média*, *vogal alta*, *vibrante*, *fricativa*, *oclusiva*, *nasal* e *lateral*. Após a rodada de dados, por alguns fatores apresentarem baixas ocorrências, decidimos amalgamar os seguintes fatores: *fricativa*, *vibrante* e *lateral*, formando o grupo *Fricativa e lateral* e os fatores

vogal baixa, vogal média e vogal alta, criando o grupo de fatores *Vogais altas e não altas*.

O resultado final da amalgamação está descrito na Tabela 5, a seguir:

Tabela 5 – Variante *PRA* considerando o *Contexto Fonológico Seguinte*.

| GRUPO | APLICAÇÃO TOTAL | % | PESO RELATIVO |
|---|-----------------|------|---------------|
| Fricativa e lateral <i>pra sustentar</i> | 33/170 | 19,4 | 0.63 |
| Vogais altas e não altas <i>pra esse</i> | 43/470 | 9,1 | 0.39 |
| Oclusiva <i>pra casa</i> | 54/370 | 14,6 | 0.55 |
| Nasais <i>pra mim</i> | 18/106 | 17,0 | 0.56 |

Input: 0.098 Significância: 0.001

Com peso relativo de (0.63), o grupo de fatores *Fricativa e lateral* foi o que mais apontou a variante *PRA*, apresentando 19,4% dos dados. Os fatores *Oclusiva* e *Nasais* foram os que vieram após, em relação ao uso de *PRA* na escrita dos escolares, apresentando pesos relativos de (0.55) e (0,56), respectivamente, ambos um pouco acima do ponto neutro.

Por fim, o fator *Vogais altas e não altas* foi o que apresentou o surgimento do *PRA* na escrita, apresentando 9,1% dos dados e peso relativo abaixo do ponto neutro (0.39). A partir dos resultados, pode-se pensar que consoantes favorecem mais o uso de *PRA* do que vogais.

Em Silva (2010), foram obtidos resultados que apontaram a *consoante dorsal* (61%, peso relativo 0.53) e a *vogal anterior* (69%, peso relativo 0.60) para a escolha de *PRA*, além da *vogal posterior* que apresentou 73% dos dados e peso relativo de (0.66).

Apesar de apresentar peso relativo maior, o fator *vogal posterior* apresentou baixa quantidade de dados em comparação com os citados anteriormente, sendo assim, a autora acredita que seria necessário uma quantidade maior de dados deste fator para se ter conclusões mais precisas.

No trabalho de Ferreira (2014), o resultado mostra pouca diferença entre os fatores *Consoante* (74,6%) e *Vogal* (71,9%) e peso relativo igual (0.50). A autora

concluiu que o grupo *Contexto Fonológico Seguinte* não apresentou significância, pois o peso relativo de ambos encontra-se no ponto neutro.

Em Ferreira (2018), o fator *Vogal Coronal* apresentou 90,6% das aplicações e peso relativo de (0.65), mostrando favorecimento para a realização do *PRA*. Após, veio a *Consoante Dorsal*, com percentual de 84,5% e peso relativo de (0.56). Os demais fatores, segundo a autora, obtiveram peso relativo abaixo do ponto neutro: *Consoante [-posterior]* (0.48) e *Vogal Dorsal* (0.29).

Quanto ao estudo de Franco (2017), a autora não utilizou o grupo de fatores *Contexto fonológico seguinte* em seu trabalho, não sendo possível a comparação com os dados apresentados nesta pesquisa.

5.3 VARIÁVEIS LINGUÍSTICAS DESCARTADAS

Em nossa análise, os grupos de fatores que não apontaram condicionamento para o fenômeno estudado na escrita foram: *Número de sílabas do item seguinte* e *Faixa etária*. Os grupos de fatores *Tonicidade da sílaba seguinte* e *Contexto morfológico seguinte* mostraram-se irrelevantes, pois não foram selecionados nem descartados pelo programa.

O programa GOLDVARB X apontou que essas variáveis linguísticas não foram condicionantes para o surgimento do *PRA*. Dessa maneira, não foi possível comprovar se tais fatores condicionam o uso de *PRA* na produção dos escolares. Mesmo assim, optou-se por fazer cruzamentos entre alguns grupos de fatores.

Além disso, o grupo de fatores *Contexto fonológico seguinte* primeiramente foi selecionado pelo *step-up*, como foi apresentado neste trabalho, porém, no final da análise, o *step-down* o descartou.

Tal situação poderá melhor ser compreendida na citação a seguir:

Outra possibilidade de resultado seria aquele em que não há coincidência entre as duas análises, isto é, grupos de fatores podem ser selecionados e depois eliminados ou ainda não serem selecionados nem eliminados, situação que os caracteriza como variáveis com status indefinido. (SANKOFF 1998, p. 991-992 in: BRESCANCINI 2002, p.47).

Por esse motivo, resolvemos fazer uma análise alternativa, retirando o grupo de fatores primeiro selecionado, o *Tipo textual*, que já mostrou sua importância ao apontar o fator *verso* como o mais significativo para o uso da variante *PRA*. Assim, esperamos dar mais chance de interações entre os fatores linguísticos, em sua maioria, descartados.

Conforme Brescancini (2002, p. 49), “uma pesquisa em regra variável construída apenas estatisticamente, sem aprofundamento linguístico, torna-se desinteressante e inadequada”, o que corrobora Guy (2007, p. 65) ao dizer que “o trabalho quantitativo não é um substituto, mas apenas um acessório para a análise linguística”. Mas sabemos que aplicar apenas a teoria nos resultados da pesquisa, pode-se lidar com equívocos nesses resultados. Em concordância, com os autores, o ideal é unir o linguístico e o estatístico na análise.

5.4 FAIXA ETÁRIA (uma variável sobreposta)

Embora descartado, pelo Programa, resolvemos apresentar o resultado desse grupo de fatores porque achamos conveniente para a nossa análise linguística.

Para tanto, mostraremos primeiro o cruzamento de dados referentes ao grupo de fatores *Faixa etária* e o grupo de fatores *Nível de escolaridade*, que foi selecionado pelo Programa, a fim de apontar o que se passou quando realizamos o cruzamento dos respectivos grupos de fatores.

Faixa etária x Nível de escolaridade

Tabela 6– Cruzamento entre os grupos de fatores *Faixa etária* e *Nível de escolaridade*.

| | 14 a 16 anos Aplic. / % | 17 a 19 anos Aplic. / % | Total Aplic. / % |
|-----------------|------------------------------------|------------------------------------|-----------------------------|
| 1ºs anos | 140 15 T= 904 | 0 - T= 0 | 140 15 T= 904 |
| 3º ano | 0 - T= 0 | 8 4 T= 212 | 8 4 T= 212 |
| Total | 140 15 T= 904 | 8 4 T= 212 | 148 13 T= 1116 |

Podemos ver na Tabela 6 que aparecem nos resultados a presença de células vazias, o que ocasionou a percepção de que talvez ambos os grupos tivessem dados iguais, apontando para a sobreposição dos grupos *Faixa etária* e *Nível de escolaridade*.

Amaral (2000) nos diz que: “Nos cruzamentos quando todas as combinações apresentam dados (ocorrem na amostra) pode-se realizar comparação entre os efeitos das células”, o que não foi possível no cruzamento acima.

Vejamos o que aponta outro autor a respeito da presença de células vazias.

Tal situação configura-se como a ideal para a operação do VARB2000: todas as células formadas pelo cruzamento das variáveis contêm dados, ou em outras palavras, todos os grupos de fatores são ortogonais. (GUY, 1998, p.31).

Na ocorrência de células vazias, o indicado é retirar o grupo de fatores que ocasionou a célula vazia e realizar uma rodada de dados sem ele para a averiguação de melhores resultados.⁶

Feita a explicação, apresentaremos a seguir, os resultados do grupo de fatores *Faixa etária*.

Tabela 7 – Variante *PRA* considerando a *Faixa etária*.

| GRUPO | APLICAÇÃO/TOTAL | % | PESO RELATIVO |
|--------------|-----------------|------|---------------|
| 14 a 16 anos | 140/904 | 15,5 | 0.57 |
| 17 a 19 anos | 8/212 | 3,8 | 0.21 |

Input: 0.098 Significância: 0.001

Ao analisar esse resultado, é possível apontar que os estudantes com *faixa etária entre 14 e 16 anos*, fazem mais uso da variante *PRA* em suas produções escritas, do que os estudantes com *faixa etária de 17 a 19 anos*.

Com relação à variante *PRA*, os escolares com a faixa etária menor, ou seja, com idade entre *14 e 16 anos* tendem a produzir mais a variante *PRA*, o que nos faz pensar que por estarem chegando ao ensino médio, eles tendem a reproduzir mais os reflexos da oralidade em seus textos, além do contato que têm com a linguagem da *internet*.

Podemos ponderar que, conforme os estudantes avançam na idade, vão tendo mais contato com a linguagem formal, também por interação com o meio escolar e as exigências de construções textuais e adequação verbal/oral, que são requeridas pelos professores, além da preocupação com as entrevistas de estágios e empregos, que nessa faixa etária é bem comum.

Não podemos descartar que, o fator *faixa etária de 17 a 19 anos* teve uma quantidade significativa menor de dados (8/212) em relação ao fator de *14 a 16 anos* (140/904), pois tivemos na pesquisa uma turma de 3º ano e quatro turmas de 1ºs anos.

⁶ Falaremos novamente sobre o fenômeno mais adiante, no subcapítulo (5.5 CRUZAMENTO ENTRE FATORES LINGUÍSTICOS E EXTRALINGUÍSTICOS).

No trabalho de Silva (2010), a autora dividiu a idade dos informantes em três faixas etárias, os mais jovens (15 a 25 anos), os mais velhos (+ de 49 anos) e a faixa etária intermediária (26 a 49 anos). O trabalho aponta que os mais jovens e os mais velhos apresentaram percentuais quase idênticos de realização da variante *PRA*, com peso relativo de 0.55 e 0.56, respectivamente; e em relação ao grupo de fator intermediário, a ocorrência diminui (0.46).

Segundo a autora, o fato de os mais jovens escolherem mais a variante *PRA* ao invés da variante *PA*, seria por estarem sofrendo interferência da escola, devido ao fato de alguns livros considerarem a variante *PRA* em sua escrita. Já com relação aos mais velhos usarem mais *PRA* do que *PA* em sua oralidade, a autora argumenta que seria uma situação de variação estável.

Em Ferreira (2014), têm-se uma situação semelhante ao trabalho de Ferreira (2018). Segundo os resultados das pesquisas da autora, os informantes mais velhos também favoreceram o uso do *PRA*.

A autora dividiu sua faixa etária em dois fatores (- de 50 anos) e (+ de 50 anos); em Ferreira (2014) o fator (+ de 50 anos) apresentou (0.64) em relação aos mais jovens (0.39) e em Ferreira (2018), o fator “mais velhos” (0.56) também apresentou peso relativo maior (0.44) em relação a fator dos “mais jovens” no uso de *PRA* na oralidade.

Ferreira (2018) acredita que o fato de os mais velhos estarem favorecendo mais o uso de *PRA* pareça indicar que tal variante não seja a forma inovadora.

Os resultados das pesquisas (acima) selecionadas são referentes à fala. Já os desta pesquisa são referentes à escrita.

5.5 CRUZAMENTO ENTRE FATORES LINGUÍSTICOS E EXTRALINGUÍSTICOS

O cruzamento de fatores tem por finalidade averiguar se há uma relação entre duas variáveis, assim, para ver se tal relação acontece cruzam-se essas variáveis.

Além disso, o cruzamento de fatores também tem por finalidade a investigação e relação entre os fatores a fim de descobrir a presença de células vazias, que, segundo (Ferreira, 2018, p. 56) “são as combinações de fatores que não geraram nenhuma ocorrência da variável que esteja sendo estudada”, aqui, a variante *PRA*.

Caso ocorra a presença de célula vazia, o indicado é retirar o grupo de fatores que ocasionou a célula vazia, após ser feita a análise desejada.

Segundo Ferreira (2018):

Esse problema pode ocorrer de uma distribuição desequilibrada dos dados, o que pode acarretar que a variante controlada não ocorra nenhuma vez antes de fatores que sejam mais escassos em determinadas análises, como uma conjunção ou uma palavra com quatro sílabas, por exemplo. (FERREIRA, 2018, p. 56- 57).

Na presente pesquisa, foram realizados quatro cruzamentos⁷ com os seguintes grupos: *Nível de escolaridade*, *Contexto fonológico seguinte*, *Contexto morfológico seguinte*, *Tipo textual*, *Sexo e Número de sílabas do item seguinte*.

5.5.1 Sexo x Nível de escolaridade

A Tabela 8, a seguir, apresenta os dados referentes ao cruzamento entre *Sexo* e *Nível de escolaridade*:

Tabela 8– Cruzamento entre os grupos de fatores *Sexo* e *Nível de escolaridade*.

| | Feminino Aplic. / % | Masculino Aplic. / % | Total Aplic. / % |
|-----------------|--------------------------------|---------------------------------|-----------------------------|
| 1ºs anos | 64 11 T= 570 | 76 23 T= 334 | 140 15 T= 904 |
| 3º ano | 5 3 T= 145 | 3 4 T= 67 | 8 4 T= 212 |
| Total | 69 10 T= 715 | 79 20 T= 401 | 148 13 T= 1116 |

A partir do cruzamento entre os grupos de fatores *Sexo* e *Nível de escolaridade*, observamos os seguintes resultados:

- 1) Meninas dos 1ºs anos escolheram a variante *PRA* em 11% das vezes (64/570);
- 2) Meninas do 3º ano escolheram a variante *PRA* em 3% das vezes (5/145);
- 3) Meninos dos 1ºs anos escolheram a variante *PRA* em 23% das vezes (76/334);
- 4) Meninos do 3º ano escolheram a variante *PRA* em 3% das vezes (4/67).

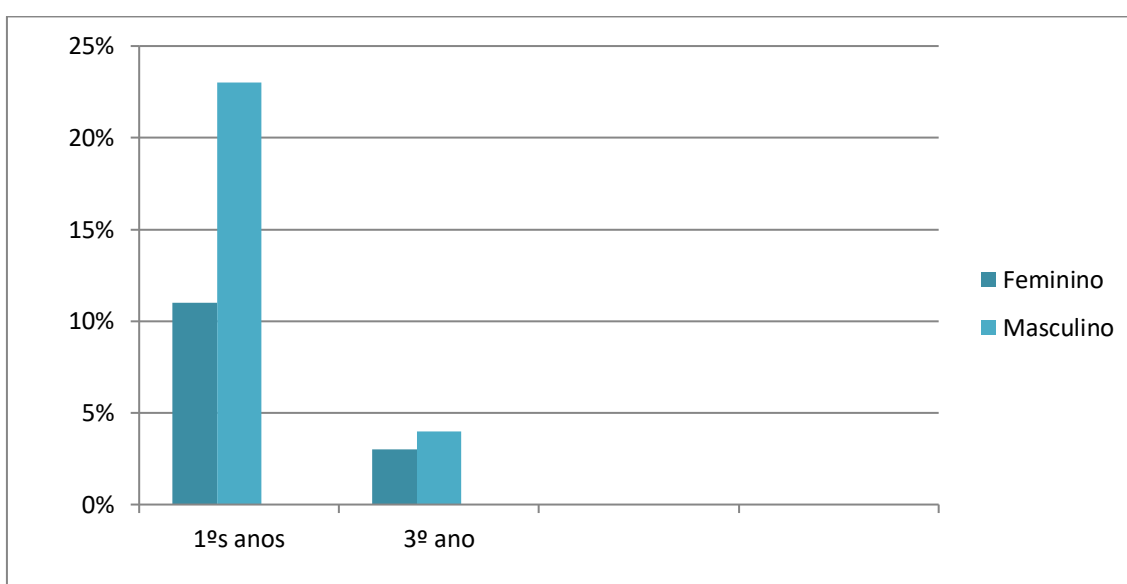
⁷ Foi realizado também o cruzamento dos grupos de fatores *Faixa etária* e *Nível de escolaridade* que foi citado acima e que ocasionou a presença de células vazias, assim, tal cruzamento não será analisado neste capítulo).

Percebemos a partir desse cruzamento que são os 1^{os} anos que utilizam mais a variante *PRA* na modalidade escrita (15%) e nesses, os meninos (23%).

Os meninos, em geral, apresentaram maior percentual de uso da variante *PRA* (20%) contra apenas (10%) das meninas, sinalizando uma preferência dos meninos pela forma inovadora *PRA* na escrita, como já havíamos mencionado anteriormente, nos resultados da rodada do grupo de fatores *Sexo*.

Podemos visualizar melhor, no gráfico a seguir, o cruzamento dos grupos de fatores *Sexo* e *Nível de escolaridade*.

Gráfico 3 – Dados referentes ao cruzamento entre os grupos de fatores *Sexo* e *Nível de escolaridade*.



O que percebemos nesse cruzamento de dados é que tanto os informantes do *sexo feminino* dos 1^{os} anos quanto os informantes do *sexo feminino* do 3^o ano utilizam um total de 10% da variante *PRA* e 90% do padrão *PARA* em suas produções na modalidade escrita. Dessa maneira, o estudo aponta que quem mais utiliza a variante *PRA* na modalidade escrita é o *sexo masculino*, porque utilizaram 20% da variante *PRA* e 80% do padrão *PARA*.

Assim, confirmamos nosso pensamento de que o *sexo masculino* talvez não se preocupe tanto ou demore mais a ter percepção das relações formais da modalidade escrita, talvez por serem menos contidos ao desempenharem seus papéis sociais no período da adolescência ou por serem mais “desligados” das regras e padrões sociais, como formalidade na escrita.

O interessante é que em relação ao 3º ano há quase uma paridade do uso tanto do *PARA* quanto do *PRA*, pois tanto os informantes do *sexo feminino* quanto os informantes do *sexo masculino* utilizaram 3% e 4% de *PRA* e 97% - 96% do *PARA* respectivamente, o que mostra um aumento do uso da forma *PARA*, pois os estudantes encontram-se em um nível mais elevado de estudo, ou seja, já estão no 3º ano e não mais nos 1ºs anos do Ensino Médio.

Isso tudo confirma a superioridade do uso da forma padrão *PARA* na modalidade escrita, apesar de nossa pesquisa ter o interesse de ver o uso do *PRA* a fim de verificar se há algum avanço no uso de tal variante linguística nos textos escritos dos alunos.

5.5.2 Contexto fonológico seguinte x Contexto morfológico seguinte

A próxima Tabela mostrará os resultados referentes ao cruzamento dos seguintes grupos de fatores: *Contexto fonológico seguinte* e *Contexto morfológico seguinte*:

Tabela 9 – Cruzamento entre os grupos de fatores *Contexto fonológico seguinte* e *Contexto morfológico seguinte*

| | Pronomes Aplic. /% | Outros Aplic. / % | Substantivo Aplic. / % | Verbo Aplic./% | Total Aplic./% |
|---------------------------------|----------------------------------|---------------------------------------|---------------------------------------|--------------------------------------|---------------------------|
| Fricativa e lateral | 7 21 T= 34 <i>pra você</i> | 2 20 T= 10 <i>pra lá</i> | 8 33 T= 24 <i>pra sociedade</i> | 16 16 T= 99 <i>pra chamar</i> | 33 20 T= 167 |
| Vogais altas e não altas | 7 7 T= 104 <i>pra esse</i> | 23 10 T= 238 <i>pro (pra+o)</i> | 4 15 T= 26 <i>pra adulto</i> | 9 9 T= 102 <i>pra estudar</i> | 43 9 T= 470 |
| Oclusivas | 2 5 T= 40 <i>pra tu</i> | 7 9 T= 79 <i>pra que</i> | 13 24 T= 55 <i>pra guerra</i> | 32 17 T= 189 <i>pra dormir</i> | 54 15 T= 363 |
| Nasais | 13 22 T= 59 <i>pra mim</i> | 3 19 T= 16 <i>pra não</i> | 1 6 T= 16 <i>pra mãe</i> | 1 7 T=14 <i>pra nascer</i> | 18 17 T= 105 |
| Total | 29 12 T= 237 | 35 10 T= 343 | 26 21 T= 121 | 58 14 T= 404 | 148 13 T= 1105 |

Ao fazer o cruzamento entre os grupos de fatores *Contexto fonológico seguinte* e *Contexto morfológico seguinte*, as classes morfológicas tinham os seguintes percentuais: *Substantivo* (21%), *Verbo* (14%), *Pronomes* (12%) e *Outros* (10%).

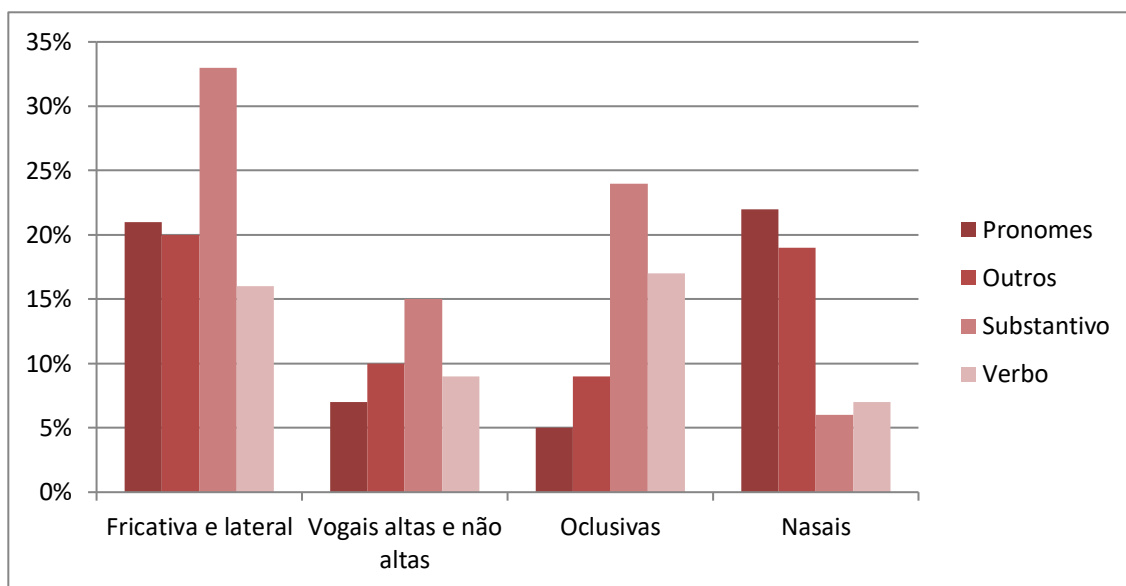
O cruzamento dos fatores morfológicos com os fatores fonológicos *Fricativa e lateral* (33%), *Oclusivas* (24%) e *Vogais altas e não altas* (15%), apresentou maior

percentual para a classe morfológica Substantivo (21%) num total de 121 ocorrências assim, os resultados nos mostraram que as palavras que vinham a seguir do *PRA* eram mais os *substantivos* que iniciavam com *Fricativa e lateral*, seguido dos *verbos* com (14%) num total de 404 ocorrências.

O fator *Nasais* apresentou percentual de (22%) para a classe morfológica *Pronomes*, isso nos mostra que a variante *PRA*, aconteceu mais antes de *Pronomes* com inicial nasal, como, por exemplo: “*pra mim*”, “*pra ninguém*”, “*pra nós*”...

O gráfico 4 mostrará melhor os resultados citados anteriormente:

Gráfico 4 – Dados referentes ao cruzamento entre os grupos de fatores *Contexto fonológico seguinte* e *Contexto morfológico seguinte*.



5.5.3 Tipo textual x Contexto morfológico seguinte

A tabela seguinte mostrará os resultados referentes ao cruzamento dos grupos de fatores *Tipo textual* e *Contexto morfológico seguinte*:

Tabela 10 – Cruzamento entre os grupos de fatores *Tipo textual* e *Contexto morfológico seguinte*

| | Pronomes Aplic. /% | Outros Aplic. /% | Substantivo Aplic. /% | Verbo Aplic. /% | Total Aplic. /% |
|--------------|-------------------------------|-----------------------------|----------------------------------|----------------------------|----------------------------|
| Prosa | 24 11 T= 215 | 18 6 T= 305 | 24 22 T= 111 | 37 10 T= 360 | 103 10 T= 991 |
| Verso | 5 23 T= 22 | 17 45 T= 38 | 2 20 T= 10 | 21 48 T= 44 | 45 39 T= 114 |
| Total | 29 12 T= 237 | 35 10 T= 343 | 26 21 T= 121 | 58 14 T= 404 | 148 13 T= 1105 |

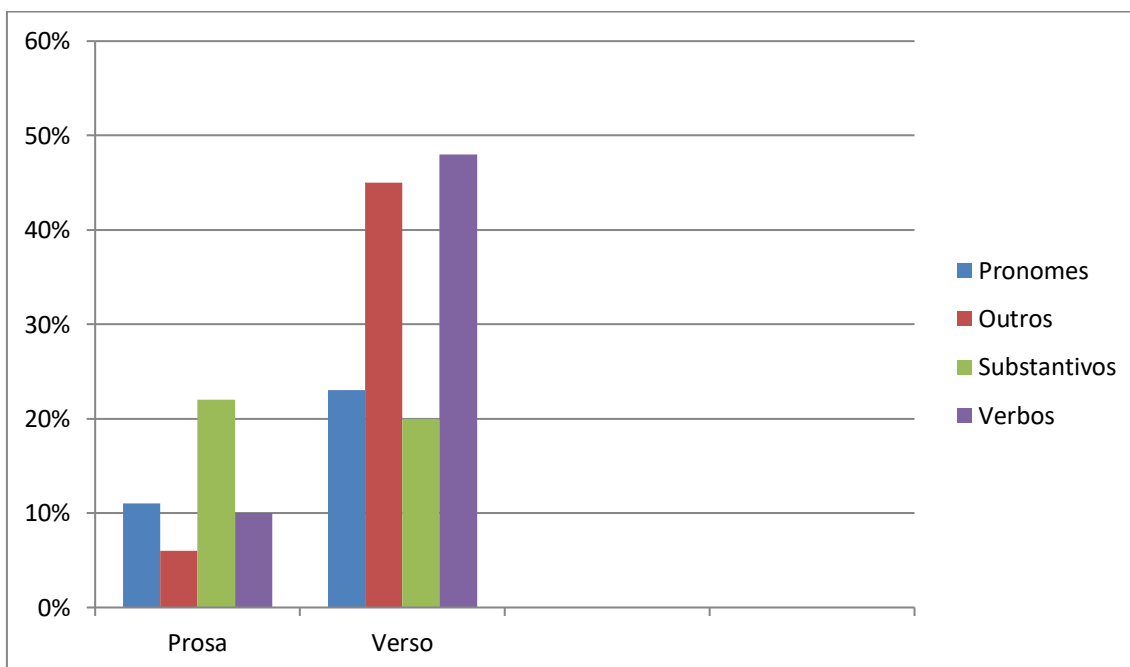
Os resultados do cruzamento entre os grupos de fatores *Tipo textual* e *Contexto morfológico seguinte* apontaram os seguintes resultados:

- a) Quando cruzamos o fator *Prosa*, o fator morfológico que mais influencia o surgimento da variante *PRA* é o *Substantivo*, com percentual de (22%);
- b) Quando cruzamos a modalidade *Verso*, o fator que mais influencia o surgimento da variante *PRA* é o *Verbo*, com percentual de (48%).

A partir dos resultados, vimos que nos textos em *Prosa* o fator escolhido como favorável ao surgimento da variante *PRA* é o *Substantivo* e nos textos em *Verso*, o fator escolhido é o *Verbo*. Isso nos faz supor que a variante *PRA*, em produções escritas, é favorecida antes de palavras lexicais, ao invés de palavras gramaticais, já que em ambos os grupos de fatores, as classes morfológicas escolhidas são as palavras lexicais (*substantivo* e *verbo*), em detrimento das palavras gramaticais (*Pronomes* e *Outros - conjunções, advérbios e artigos*).⁸

O gráfico 5 ilustrará melhor os resultados obtidos, vejamos:

Gráfico 5 – Dados referentes ao cruzamento entre os grupos de fatores *Contexto morfológico seguinte* e *Tipo textual*.



⁸ As classes gramaticais *Numeral* e *Pronomes* apresentaram *knockouts* (com 100% de aplicações para o padrão *PARA*) e foram retiradas após a primeira rodada de dados.

6. RODADA DE DADOS SEM OS GRUPOS DE FATORES TIPO TEXTUAL E FAIXA ETÁRIA

Ao realizarmos a segunda rodada de dados (após a primeira rodada decidimos amalgamar alguns fatores referentes aos grupos *Contexto fonológico seguinte* e *Contexto morfológico seguinte*), o programa apontou sobreposição dos grupos de fatores *Nível de escolaridade* e *Faixa etária*, assim, Brescancini (2002) nos diz que “para que pudéssemos obter valores confiáveis para as variáveis em destaque, seria aconselhável desconsiderar uma delas na rodada em que deseja analisar o comportamento da outra”, o que foi realizado, apontando resultados iguais em ambos os grupos

O grupo de fatores *Tipo textual* apresentou resultados plausíveis, na rodada anterior, porém havia disparidade em relação à quantidade de dados referentes aos subfatores *verso* (13%) e *prosa* (87%). Como a maioria dos dados estão em *prosa*, esperamos que haja maior interação e significância entre os demais grupos de fatores.

Então, nem sempre vamos querer simplesmente concluir a análise mediante a apresentação dos resultados escolhidos como a melhor rodada escolhida pelo *step-up/ step-down*. Podemos dividir as possibilidades analíticas nas quatro categorias seguintes, sendo as mais comuns a primeira e a terceira, enquanto a segunda e a quarta são alternativas usadas somente em função de certas justificativas:

- 1) retirar grupos idênticos pelo *step* como sem significância;
- 2) manter, pelo menos para certos fins grupos sem significância;
- 3) manter grupos idênticos pelo *step* como tendo significância;
- 4) retirar, ou não manter para todas as análises, grupos com significância. (GUY, 1998, p. 182-183).

Dessa maneira, decidimos realizar mais uma rodada de dados, desta vez, sem os grupos de fatores *Faixa etária* (devido ao fato de o programa ter selecionado o grupo *Nível de escolaridade* e ambos os grupos apresentarem dados e resultados iguais) e o grupo *Tipo textual* (devido ao fato da disparidade em relação aos dados), para verificarmos uma última vez os resultados obtidos para esta análise.

6.1 GRUPOS DE FATORES SELECIONADOS PELO PROGRAMA GOLDVARB X

Após realizarmos nossa última rodada de dados (sem os grupos de fatores *Faixa etária* e *Tipo textual*), o Programa selecionou os seguintes grupos de fatores como favoráveis ao surgimento da variante *PRA*:

- (i) *Contexto morfológico seguinte*
- (ii) *Contexto fonológico seguinte*
- (iii) *Tonicidade da sílaba seguinte*
- (iv) *Sexo*
- (v) *Nível de escolaridade*

Dos cinco grupos de fatores selecionados pelo Programa nesta última rodada realizada nesta pesquisa, podemos perceber que foram selecionados dois grupos de fatores que não haviam sido selecionados anteriormente: *Contexto morfológico seguinte* e *Tonicidade da sílaba seguinte*.

Um fato importante é que o Programa descartou os mesmos grupos de fatores da rodada anterior: *Número de sílabas do item seguinte* (este grupo não foi selecionado em nenhuma das rodadas realizadas nesta pesquisa), e *Contexto fonológico seguinte* (este grupo foi novamente selecionado e também descartado pelo Programa, assim como aconteceu em todas as rodadas realizadas).

Assim, decidimos analisar apenas os dois grupos de fatores que foram selecionados nesta última rodada de dados e que haviam sido “ignorados” na rodada anterior. São eles: *Contexto morfológico seguinte* e *Tonicidade da sílaba seguinte*.

6.1.1 *Contexto morfológico seguinte*

O grupo de fatores *Contexto morfológico seguinte* foi o primeiro a ser selecionado pelo Programa. É de interesse nosso saber os resultados desse grupo de fatores, já que gostaríamos de saber quais as classes gramaticais (as palavras) que vêm depois da preposição *PARA* e de sua variante *PRA*.

Assim, apresentamos na Tabela a seguir, os resultados do grupo de fatores *Contexto morfológico seguinte*.

Tabela 11 – Variante *PRA* considerando o *Contexto morfológico seguinte*

| GRUPO | APLICAÇÃO/TOTAL | % | PESO RELATIVO |
|---------------------------------|-----------------|------|---------------|
| Pronomes <i>pra mim</i> | 29/237 | 12,2 | 0.39 |
| Outros <i>pra lá</i> | 35/343 | 10,2 | 0.48 |
| Substantivos <i>pra casa</i> | 26/121 | 21,5 | 0.63 |
| Verbos <i>pra viver</i> | 58/404 | 14,4 | 0.52 |

Input: 0.107 Significância: 0.035

O fator que mais favoreceu o surgimento da variante *PRA* nas produções dos escolares foi o *Substantivo*, com peso relativo de (0.63), em seguida veio o *Verbo* com (0.52). Os demais fatores (*Outros* e *Pronomes*) encontram-se abaixo do peso neutro.

O que pensamos a respeito dos resultados é que a variante *PRA* é favorecida antes de palavras lexicais, como já foi apontado anteriormente (nos resultados do cruzamento de dados realizado na rodada anterior). Assim, achamos que o *PRA* necessita “ligar-se” a uma palavra lexical na modalidade escrita.

No trabalho de Ferreira (2014), a autora relata que o fator (*Outros* – numerais, advérbios, pronomes, etc...) mostrou-se mais favorável ao uso de *PRA* (0.57), assim, a autora concluiu que a variante *PRA* não necessita realizar processo de juntura com palavras lexicais – verbais (0.51) e, principalmente, nominais (0.39).

No trabalho de Ferreira (2018), o fator *Palavra gramatical* (engloba todos os casos de não verbos e não nomes, como pronomes, numerais, advérbios, etc.) teve peso relativo (0.65) ao favorecimento do *PRA*, em detrimento da categoria *Palavra Nominal* (0.37).

Se pararmos para analisar, veremos que na oralidade, são as *Palavras gramaticais* que favorecem o *PRA*, ao contrário de nossa pesquisa, que é oriunda da escrita, onde são as *Palavras lexicais* (*Substantivo* e *Verbo*) que favorecem o seu uso.

Não foram comparados os resultados de Silva (2010) e de Franco (2017), pois as autoras não abordaram tais grupos de fatores em suas pesquisas.

6.1.2 Tonicidade da sílaba seguinte

Este grupo de fatores mostrou-se relevante em nossa última rodada de dados, vejamos na Tabela a seguir os resultados apresentados pelo programa GOLDVARB X a respeito do respectivo.

Tabela 12 – Variante *PRA* considerando a *Tonicidade da sílaba seguinte*

| GRUPO | APLICAÇÃO/TOTAL | % | PESO RELATIVO |
|--------|-----------------|------|---------------|
| Átona | 90/779 | 11,6 | 0.45 |
| Tônica | 58/337 | 17,2 | 0.60 |

Input: 0.107 Significância: 0.035

Ao analisarmos os resultados acima, pudemos perceber que a sílaba *Tônica* é a que mais favorece o aparecimento da variante *PRA* na escrita dos estudantes, tal subfator apresentou 17,2 % dos dados e um peso relativo de (0.60).

O que pensamos a respeito do resultado é que talvez, as sílabas tônicas dêem força para o surgimento da variante, devido ao fato de *PRA* ser uma monossílaba átona, favorecendo a colocação de uma tônica após seu uso.

Em Ferreira (2014) é também a sílaba *Tônica* que favorece o surgimento do *PRA* (0.54), porém a diferença entre os percentuais é de apenas 7% e isso faz com que a autora conclua que a tonicidade da sílaba posterior exerce influência moderada no surgimento da variante *PRA* na oralidade.

Não foi possível comparar nossos resultados com o de Silva (2010) e Franco (2017), pois as autoras não analisaram este grupo de fatores. Em Ferreira (2018), a autora analisa o grupo de fatores *Tonicidade da sílaba seguinte* apenas em cruzamentos.

6.2 CRUZAMENTO ENTRE OS GRUPOS DE FATORES CONTEXTO MORFOLÓGICO SEGUINTE E NÚMERO DE SÍLABAS DO ITEM SEGUINTE

Por fim, decidimos realizar um último cruzamento entre os grupos de fatores *Contexto morfológico seguinte* (que foi selecionado pelo Programa nesta rodada de dados e “ignorado” nas demais) e *Número de sílabas do item seguinte* (que foi descartado em todas as rodadas realizadas pelo programa GOLDVARB X).

6.2.1 Número de sílabas do item seguinte x Contexto morfológico seguinte

A Tabela a seguir mostrará os resultados referentes ao cruzamento dos dados:

Tabela 13 – Cruzamento entre os grupos de fatores *Número de sílabas do item seguinte* e *Contexto morfológico seguinte*

| | Pronomes Aplic. /% | Outros Aplic. /% | Substantivos Aplic. /% | Verbos Aplic. /% | Total Aplic. /% |
|------------------------------|-------------------------------|-----------------------------|-----------------------------------|-----------------------------|----------------------------|
| 1 sílaba | 16 19 T= 86 | 30 10 T= 296 | 2 25 T= 8 | 10 20 T= 51 | 58 13 T= 441 |
| 2 sílabas | 13 9 T= 142 | 3 8 T= 36 | 11 26 T= 43 | 31 19 T= 167 | 58 15 T= 388 |
| 3 sílabas | 0 0 T= 8 | 2 20 T= 10 | 10 20 T= 49 | 13 10 T= 125 | 25 13 T= 192 |
| 4 ou mais sílabas | 0 0 T= 1 | 0 0 T= 1 | 3 14 T= 21 | 4 7 T= 61 | 7 8 T= 84 |
| Total | 29 12 T= 237 | 35 10 T= 343 | 26 21 T= 121 | 58 14 T= 404 | 148 13 T= 1105 |

Pela distribuição das ocorrências nas células que se formaram neste cruzamento, vimos que nos textos dos escolares, não ocorrem *Pronomes* e *Outros* com *mais de três sílabas* e nem *Pronomes* com três sílabas. Os 29 *Pronomes* que seguem o *PRA* tem 1 ou 2 sílabas no máximo, e no fator *Outros*, também aparecem palavras de até 3 sílabas.

Os fatores que estão distribuídos em todas as células (vertical e horizontal) são os *Substantivos* com (21%) dissílabos e os *Verbos* que tem (14%) monossílabos.

7. CONCLUSÃO

O presente trabalho buscou investigar a variação da preposição *PARA* e de sua variante *PRA* na escrita de escolares cursando os primeiros e terceiro ano de uma escola pública estadual localizada próxima ao centro da cidade de Rio Grande – RS.

O *corpus* da pesquisa é constituído por textos em *verso* e *prosa*. As rodadas de dados foram realizadas através do programa estatístico GOLDVARB X e a metodologia adotada foi a da Teoria da Variação de William Labov (2008[1972]).

Através dos resultados obtidos no grupo de fatores *Tipo textual (verso e prosa)* confirmamos nossas hipóteses de que as variantes *PARA* e *PRA* iriam se comportar de forma distinta nos diferentes tipos de textos e também que os alunos considerariam a escolha das variantes a partir do grau de formalidade dos textos propostos.

Assim, acreditamos que os escolares ao realizarem suas produções textuais no tipo textual *verso*, sentem-se mais “livres”, por poderem produzir algo próximo do melódico, o que os fazem aludir sua escrita mais próxima da oralidade e o que facilita o surgimento da variante *PRA* ao invés do padrão *PARA*.

Alguns fatores linguísticos mostraram-se relevantes para o aparecimento da variante *PRA* na escrita dos escolares como, por exemplo: o fator *Verso*, a *Faixa etária entre 14 e 16 anos*, o *Nível de escolaridade primeiros anos*, o *Sexo masculino* e o *Contexto fonológico seguinte fricativa e lateral*.

O cruzamento de dados que realizamos entre os grupos de fatores sociais *Sexo* e *Nível de escolaridade* mostrou um favorecimento dos meninos dos primeiros anos ao escolherem a variedade linguística *PRA* em suas produções escritas.

O cruzamento entre os grupos de fatores *Contexto fonológico seguinte* e *Contexto morfológico seguinte* mostrou o *Substantivo* antecedido de *consoantes Fricativas e lateral* na realização do aparecimento da variante *PRA* na escrita dos escolares.

Por fim, o cruzamento dos grupos de fatores *Tipo textual* com o *Contexto morfológico seguinte* também apontou percentual maior para o *Substantivo* nas

produções em *Prosa* e dos *Verbos*, nas produções em *Verso*.

Em relação à questão ampla que norteou este trabalho, de como se comportariam as variantes *PARA* e *PRA* na escrita dos alunos dos primeiros e terceiro ano do Ensino Médio, pôde ser respondida através de nossos resultados, pois os escolares dos 1ºs anos usam mais o *PRA* do que os escolares do 3º ano, ou por estarem entrando no ensino médio e não terem tido tanto contato com as regras normativas na escola ainda, ou por carregarem mais reflexos da fala em produções escritas.

Quanto a nossa última rodada de dados sem os grupos de fatores *Tipo textual* e *Faixa etária*, as variáveis linguísticas que se mostraram relevantes para o aparecimento da variante *PRA* na escrita dos escolares foram: *Contexto morfológico seguinte*, *Contexto fonológico seguinte*, *Tonicidade da sílaba seguinte*, *Sexo* e *Nível de escolaridade*.

Com referência ao cruzamento que foi realizado na última análise entre os grupos de fatores *Contexto morfológico seguinte* e *Número de sílabas do item seguinte* são os *Substantivos dissílabos* (26%) e os *Verbos monossílabos* (20%) que favorecem mais o aparecimento da variante *PRA* na escrita dos escolares, o que aponta que são as palavras lexicais que favorecem mais o *PRA*.

O trabalho justificou-se pela escassez do estudo da preposição *PARA* e de sua variante *PRA* na escrita, e também por possibilitar futuras pesquisas em torno do tema abordado. Além disso, o trabalho fundamentou-se também, pela necessidade de um estudo entre as formas *PARA* e *PRA* com vistas a uma discussão mais ampla entre os registros formais e informais, ampliando o interesse pela pesquisa referente a esse assunto.

A pesquisa também se justificou pelo retorno ao meio escolar, já que a sala de aula foi o local onde a coleta de dados se realizou. Dessa maneira, por partir de uma prática escolar, a pesquisa permite, também, questionamentos pertinentes do professor frente às questões de mudança na língua, o que é tradicional e o que é inovador, a fim de realizar um trabalho mais direcionado na adequação textual e não somente visando o “certo” e o “errado” na modalidade escrita.

Apesar de apontarmos suposições para os resultados desta dissertação, acreditamos ser necessária a realização de mais pesquisas na área da Sociolinguística e da Fonologia com a variante *PRA* na modalidade escrita, a fim de compará-las com nossos resultados.

Esperamos que os resultados obtidos nesta pesquisa possam contribuir para outros estudos da área e para o conhecimento do uso da preposição *PARA* e de sua variante *PRA* na escrita de escolares. Acreditamos que a pesquisa possa contribuir também para o desenvolvimento de atividades na sala de aula que abram espaço para a adequação linguística, com o intuito de ampliar os conhecimentos sobre a variação na escrita de textos oriundos do meio escolar e também como uma tentativa de minimizar o problema em relação ao uso das variáveis em contextos pré-estabelecidos.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- AMARAL, Marisa. P. do. **A síncope em proparoxítonas: uma regra variável.** In: BISOL, Leda; BRESCANCINI, Cláudia (org). **Fonologia e variação: recortes do português brasileiro.** – Porto Alegre: EDIPUCRS, 2002, p. 99-126.
- AMARAL, Marisa. P. do. **As Proparoxítonas: teoria e variação.** Tese de Doutorado. Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2000.
- BAGNO, Marcos. **Gramática pedagógica do português brasileiro.** São Paulo: Parábola Editorial, 2012.
- BECHARA, Evanildo. **Moderna gramática portuguesa.** Rio de Janeiro: Lucerna, 2004.
- BISOL, Leda; BRESCANCINI, Cláudia (org). **Fonologia e variação: recortes do português brasileiro.** – Porto Alegre: EDIPUCRS, 2002.
- CASTILHO, Ataliba T. de. **Nova gramática do português brasileiro / Ataliba T. de Castilho.** – 1. ed., 2ª reimpressão – São Paulo: Contexto, 2012.
- CASTILHO, Ataliba T. de. **Gramática do português culto falado no Brasil: volume VII: a construção fonológica da palavra / coordenador geral Ataliba T. de Castilho: organizadora Maria Bernadete M. Abaurre.** – São Paulo: Contexto, 2013.
- CEGALLA, Domingos Paschoal, 1920-. **Novíssima gramática da língua portuguesa / Domingos Paschoal Cegalla.** – 46. ed. – São Paulo: Companhia Editora Nacional, 2005.
- COELHO... [et al.]. **Para conhecer sociolinguística / Izete Lehmkuhl Coelho... [et al.].** – São Paulo: Contexto, 2015. – Coleção para conhecer linguística.
- CUNHA, Celso; CINTRA, Luís. **Nova gramática do português contemporâneo.** Rio de Janeiro: Nova Fronteira, 2001.
- CÂMARA JR, Joaquim Mattoso. **História e estrutura da língua portuguesa.** 2ª Ed., Rio de Janeiro, Padrão, 1979.
- FARACO, Carlos. Emílio; MOURA, Francisco. Marto. **Língua e Literatura.** Vol. 1. São Paulo : Ática, 2002.
- FERREIRA, Mauro. **Aprender e praticar gramática.** São Paulo: FDT, 2003.
- FERREIRA, Melissa Osterlund. **A variação da preposição PARA na fala de Curitiba e de Florianópolis pelos Dados do VARSUL.** Dissertação de Mestrado – Universidade Federal do Rio Grande do Sul. Porto Alegre, 2018.

- FERREIRA, Melissa Osterlund. **A variação da preposição PARA na fala de Londrina pelos dados do VARSUL**. Trabalho de conclusão de curso (Graduação) - Universidade Federal do Rio Grande do Sul. Porto Alegre, 2014.
- FRANCO, Andressa Coelho. **A preposição PARA e sua variante PRA na escrita de alunos do primeiro ano do Ensino Médio**. Pós-graduação *Lato Sensu* em Linguística e Ensino de Língua Portuguesa - Univeridade Fedral do Rio Grande – FURG. Rio Grande, 2017.
- GUY, G. **Varbrul: análise avançada**. Tradução de Ana Maria Stahl Zilles. Cadernos de Tradução, Porto Alegre, UFRGS, Instituto de Letras, 1998.
- LABOV, Willian. **Padrões sociolinguísticos** / William Labov; tradução Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. – São Paulo, Parábola Editorial, 2008.
- LUFT, Celso Pedro, 1921 – **Gramática resumida: explicação da Nomenclatura Gramatical Brasileira**. / Celso Pedro Luft. – 9ªed.– Porto Alegre: Globo, 1987.
- MARCUSCHI, Luiz Antônio. **Da fala para a escrita : atividades de retextualização** / Luiz Antônio Marcuschi – 8. ed. – São Paulo : Cortez, 2007.
- MOLLICA, Maria Cecília, Maria Luiza Braga (org). **Introdução à Sociolinguística: o tratamento da variação**. 2ª Ed. – São Paulo: Contexto, 2004.
- NASCIMENTO, A. M. F.; VIEIRA, M. S.; **Um estudo da variação entre para e pra em verbetes publicados no Instagram**/ Aline Moreira da Fonseca Nascimento & Marília Silva Vieira. - Revista Coralina– Cidade de Goiás, vol. 1, n. 1, fev./2019.
- ONOFRE, Diana Pilatti. **Trabalhando com o GOLDVARB 2001**. SlideShare, 2001. <https://pt.slideshare.net/dianapilatti/trabalhando-com-goldvarb-2001-diana-pilatti-onofre>: pesquisa realizada em 02/01/202.
- SEARA, Izabel Christine. **Para conhecer fonética e fonologia do português brasileiro** / Izabel Christine Seara, Vanessa Gonzaga Nunes, Cristiane Lazzarotto – Volcão. – São Paulo: Contexto, 2015.
- SILVA, Nahete de Alcantara. **A Preposição para e suas variantes no falar araguanense** / Nahete de Alcantara Silva. - - João Pessoa, 2010.
- TARALLO, Fernando. **A Pesquisa Sociolinguística**. 3ª Edição. São Paulo: Ática, 1990.

ANEXO A – Texto de um escolar com a presença da variante PRA (1º ano - menino).

Lehrstaller nastelns turma: 104

A pena de morte

248

A pena de morte é uma punição que no meu
país é uma violência.

Aqui no Brasil não existe a pena de morte,
mas seria uma boa, mesmo eu achando isso um
pouco ruim e mesmo pois tem muita liberdade,
e muito estudo, porém estando em número, muita
gente roubando, e matando e a lei não é seguida.
Logo fígura, ou no outro dia tá libertando, isso
na nos brasileiros é uma vergonha, temos
políticos que não ajudam só sabem sacar do
seu dinheiro libertando.

Em outros países ou em alguns tem a
pena de morte, e não como o país vai pra
frente, pois sabem se roubar, matar, estudar, vai
ter a pena de morte, e uma coisa triste é se fa-
mília pois perde um parente, pra mãe e pra pais
mais também se eles perdem a outra família
também e deixar na rua a pessoa vai conti-
nuar matando, todos não aceitam ninguém se
libera, e se tivesse pena no nosso país, agente
estava com poucas problemas, pois nem todos
direitos tem pra todos ladrões, estupradores.

Então eu acho que teria que ter punição,
pois ninguém respeita mais ninguém e no
Brasil está uma vergonha, ninguém consegue sair.

ANEXO B – Texto de um escolar com a presença da variante PRA (3º ano – menina).

Nome: Rafaela Martins Brandelt

Turma: 301

(15)

taxonomizado, for ser feito
condenado, amarrado, nem julgado
se' humilhado, me monte calado
frente os homens de poder
só me resta escrever

Não me deem escolhas
Deus me fez assim
como conta a Iza
dona de mim
mãe foi um amor
minha cor é linda
sou negro e não cinza

libertado, soni, se é **faa** viver, vivo
ainda luto, eu sim de gente
sou adriado mesma sendo preto
quem disse que eu mãe podia
me encontrar aqui no comitê
eu nasci negro e nasci **faa** viver

ANEXO C – Texto de um escolar com a presença da variante PRA (1º ano – menina).

331

data

S T O O S S D

Negro, claro, bege a sua cor tanto faz
 Você tem que saber que somos todos iguais,
 o povo negro veio pra revolucionar
 veio pra nossa sociedade melhorar

Por muito tempo foram escravizados
 e mesmo com a liberdade ainda são
 deixados de lado
 alguns brancos acham que são importante
 mas pra mim não passam de seres repugnantes

Mas eu não to aqui pra desmerecer ninguém
 o que mas tem no mundo é gente
 preconceituosa e arrogante mas isso
 vai melhorar é o meu sonho mas importante.

Natamiele Souza 102

NSA

ANEXO D – Texto de um escolar com a presença da variante PRA (3º ano – menina).

E.E.E.M. Marechal Mascarenhas de Moraes

Redação – Prof.ª Ana Paula de Andrade

NOME: Lauro J. Garcia

TURMA: 303

DATA: 21/ 8 / 19

402

P

TEMA: A vacinação deve ser obrigatória?

REDAÇÃO

Atualmente, podemos observar o baixo índice de pessoas ^{em} favor da vacina. A obrigatoriedade da vacinação é algo que causa medo e dúvidas aos cidadãos devido os pensamentos enraizados ao longo dos anos.

Durante muito tempo, se conhece uma cultura imposta pela sociedade de que "os vacinos têm muitos efeitos colaterais que não são divulgados", mas, na verdade, isto é a falta de conhecimento e informações.

Conforme a vacinação vai sendo "excluída", há chance de grandes vírus e doenças graves voltarem com todo furo, ou como ou por exemplo, o sarampo, que poderia ter sido eliminado do país, mas, por falta de cuidado da população, voltou com um grande surto. Principalmente nesse país, que recebe muitos estrangeiros e fugitivos de outros países, com isso, a entrada e saída das doenças é mútua e precisa ser controlada.

solto

??

Para a vacinação ser novamente aceita, deve-se tomar atitudes, dando início com o mínimo de campanhas feitas pelo Ministério da Saúde e com ajuda em que onde representem pra sociedade que não há maléficos, mas que nos protegem de muitas doenças e o que pode acontecer são inações de nosso organismo.

pontuação

para

coisas

Questão sociocultural

| Competências | Avaliação | | | | | |
|--------------------------------------|-----------|----|----|-----|-----|-----|
| | 0 | 40 | 80 | 120 | 160 | 200 |
| 1-Norma culta | 0 | 40 | 80 | 120 | 160 | 200 |
| 2-Compreensão da proposta de Redação | 0 | 40 | 80 | 120 | 160 | 200 |

ANEXO E – Rodada de dados sem *Knockouts* e com *Amalgamações*.

Name of token file: 2205.tkn

Name of condition file: 2205 ESSE.cnd

(

(1)

(2 (o (COL 2 a))

(o (COL 2 q))

(o (COL 2 o))

(p (COL 2 i))

(p (COL 2 d))

(p (COL 2 b))

(p (COL 2 t))

(p (COL 2 p))

(s (COL 2 s))

(/ (COL 2 l))

(v (COL 2 v))

(/ (COL 2 R))

(3 (r (COL 3 j))

(r (COL 3 u))

(r (COL 3 r))

(2 (COL 3 3))

(2 (COL 3 4))

(2 (COL 3 2))

(k (COL 3 k))

(n (COL 3 n))

(4)

(5)

(6)

(7)

(8)

(9)

)

Number of cells: 243

Application value(s): 1

Total no. of factors: 22

| | | Non- | | | |
|-------|---|------|------|-------|------|
| Group | | Apps | apps | Total | % |
| ----- | | | | | |
| 1 (2) | | | | | |
| p | N | 29 | 208 | 237 | 21.4 |
| | % | 12.2 | 87.8 | | |
| o | N | 35 | 308 | 343 | 31.0 |
| | % | 10.2 | 89.8 | | |
| s | N | 26 | 95 | 121 | 11.0 |
| | % | 21.5 | 78.5 | | |
| v | N | 58 | 346 | 404 | 36.6 |
| | % | 14.4 | 85.6 | | |
| | | | | | |
| Total | N | 148 | 957 | 1105 | |
| | % | 13.4 | 86.6 | | |
| ----- | | | | | |
| 2 (3) | | | | | |
| r | N | 33 | 137 | 170 | 15.2 |
| | % | 19.4 | 80.6 | | |
| 2 | N | 43 | 427 | 470 | 42.1 |
| | % | 9.1 | 90.9 | | |
| k | N | 54 | 316 | 370 | 33.2 |
| | % | 14.6 | 85.4 | | |
| n | N | 18 | 88 | 106 | 9.5 |

% 17.0 83.0

Total N 148 968 1116

% 13.3 86.7

3 (4)

c N 90 689 779 69.8

% 11.6 88.4

e N 58 279 337 30.2

% 17.2 82.8

Total N 148 968 1116

% 13.3 86.7

4 (5)

8 N 7 79 86 7.7

% 8.1 91.9

5 N 58 389 447 40.1

% 13.0 87.0

7 N 25 167 192 17.2

% 13.0 87.0

6 N 58 333 391 35.0

% 14.8 85.2

Total N 148 968 1116

% 13.3 86.7

5 (6)

g N 103 898 1001 89.7

% 10.3 89.7

h N 45 70 115 10.3
 % 39.1 60.9

Total N 148 968 1116
 % 13.3 86.7

6 (7)

f N 69 646 715 64.1
 % 9.7 90.3

m N 79 322 401 35.9
 % 19.7 80.3

Total N 148 968 1116
 % 13.3 86.7

7 (8)

P N 140 764 904 81.0
 % 15.5 84.5

T N 8 204 212 19.0
 % 3.8 96.2

Total N 148 968 1116
 % 13.3 86.7

8 (9)

X N 140 764 904 81.0
 % 15.5 84.5

Z N 8 204 212 19.0
 % 3.8 96.2

Total N 148 968 1116
 % 13.3 86.7

 TOTAL N 148 968 1116
 % 13.3 86.7

Name of new cell file: 2205 ESSE.cel

• BINOMIAL VARBRUL • 22/06/2020 13:39:15

Name of cell file: 2205 ESSE.cel

Averaging by weighting factors.

Threshold, step-up/down: 0.050001

Stepping up...

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

Run # 1, 1 cells:

Convergence at Iteration 2

Input 0.133

Log likelihood = -436.725

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

Run # 2, 5 cells:

Convergence at Iteration 4

Input 0.131

Group # 1 -- p: 0.481, o: 0.430, s: 0.645, v: 0.526

Log likelihood = -431.829 Significance = 0.021

Run # 3, 4 cells:

Convergence at Iteration 4

Input 0.128

Group # 2 -- r: 0.621, 2: 0.408, k: 0.538, n: 0.582

Log likelihood = -429.538 Significance = 0.004

Run # 4, 2 cells:

Convergence at Iteration 4

Input 0.131

Group # 3 -- c: 0.465, e: 0.580

Log likelihood = -433.581 Significance = 0.013

Run # 5, 4 cells:

Convergence at Iteration 5

Input 0.131

Group # 4 -- 8: 0.370, 5: 0.496, 7: 0.497, 6: 0.535

Log likelihood = -435.182 Significance = 0.389

Run # 6, 2 cells:

Convergence at Iteration 5

Input 0.121

Group # 5 -- g: 0.456, h: 0.824

Log likelihood = -408.706 Significance = 0.000

Run # 7, 2 cells:

Convergence at Iteration 4

Input 0.126

Group # 6 -- f: 0.426, m: 0.630

Log likelihood = -425.879 Significance = 0.000

Run # 8, 2 cells:

Convergence at Iteration 5

Input 0.120

Group # 7 -- P: 0.573, T: 0.223

Log likelihood = -423.742 Significance = 0.000

Run # 9, 2 cells:

Convergence at Iteration 5

Input 0.120

Group # 8 -- X: 0.573, Z: 0.223

Log likelihood = -423.742 Significance = 0.000

Add Group # 5 with factors gh

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

Run # 10, 10 cells:

Convergence at Iteration 6

Input 0.117

Group # 1 -- p: 0.486, o: 0.420, s: 0.667, v: 0.525

Group # 5 -- g: 0.454, h: 0.831

Log likelihood = -402.747 Significance = 0.009

Run # 11, 8 cells:

Convergence at Iteration 5

Input 0.115

Group # 2 -- r: 0.632, 2: 0.398, k: 0.552, n: 0.558

Group # 5 -- g: 0.455, h: 0.830

Log likelihood = -400.858 Significance = 0.002

Run # 12, 4 cells:

Convergence at Iteration 5

Input 0.119

Group # 3 -- c: 0.468, e: 0.574

Group # 5 -- g: 0.456, h: 0.822

Log likelihood = -406.201 Significance = 0.028

Run # 13, 8 cells:

Convergence at Iteration 5

Input 0.120

Group # 4 -- 8: 0.392, 5: 0.493, 7: 0.506, 6: 0.529

Group # 5 -- g: 0.456, h: 0.822

Log likelihood = -407.711 Significance = 0.580

Run # 14, 4 cells:

Convergence at Iteration 5

Input 0.115

Group # 5 -- g: 0.456, h: 0.821

Group # 6 -- f: 0.428, m: 0.626

Log likelihood = -399.248 Significance = 0.000

Run # 15, 4 cells:

Convergence at Iteration 5

Input 0.108

Group # 5 -- g: 0.454, h: 0.831

Group # 7 -- P: 0.576, T: 0.213

Log likelihood = -395.185 Significance = 0.000

Run # 16, 4 cells:

Convergence at Iteration 5

Input 0.108

Group # 5 -- g: 0.454, h: 0.831

Group # 8 -- X: 0.576, Z: 0.213

Log likelihood = -395.185 Significance = 0.000

Add Group # 8 with factors XZ

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

Run # 17, 19 cells:

Convergence at Iteration 5

Input 0.106

Group # 1 -- p: 0.472, o: 0.437, s: 0.662, v: 0.520

Group # 5 -- g: 0.453, h: 0.836

Group # 8 -- X: 0.574, Z: 0.218

Log likelihood = -390.227 Significance = 0.019

Run # 18, 16 cells:

Convergence at Iteration 6

Input 0.103

Group # 2 -- r: 0.636, 2: 0.396, k: 0.552, n: 0.562

Group # 5 -- g: 0.453, h: 0.839

Group # 8 -- X: 0.577, Z: 0.209

Log likelihood = -387.162 Significance = 0.001

Run # 19, 8 cells:

Convergence at Iteration 5

Input 0.106

Group # 3 -- c: 0.462, e: 0.586

Group # 5 -- g: 0.455, h: 0.829

Group # 8 -- X: 0.578, Z: 0.206

Log likelihood = -391.869 Significance = 0.010

Run # 20, 15 cells:

Convergence at Iteration 5

Input 0.108

Group # 4 -- 8: 0.405, 5: 0.505, 7: 0.495, 6: 0.518

Group # 5 -- g: 0.455, h: 0.830

Group # 8 -- X: 0.576, Z: 0.214

Log likelihood = -394.591 Significance = 0.756

Run # 21, 8 cells:

Convergence at Iteration 5

Input 0.103

Group # 5 -- g: 0.455, h: 0.828

Group # 6 -- f: 0.430, m: 0.623

Group # 8 -- X: 0.575, Z: 0.216

Log likelihood = -386.402 Significance = 0.000

Run # 22, 4 cells:

Convergence at Iteration 5

Input 0.108

Group # 5 -- g: 0.454, h: 0.831

Group # 7 -- P: 0.558, T: 0.269

Group # 8 -- X: 0.518, Z: 0.423

Log likelihood = -395.185 Significance = 0.994

Add Group # 6 with factors fm

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

Run # 23, 36 cells:

Convergence at Iteration 6

Input 0.101

Group # 1 -- p: 0.473, o: 0.436, s: 0.669, v: 0.518

Group # 5 -- g: 0.454, h: 0.834

Group # 6 -- f: 0.429, m: 0.625

Group # 8 -- X: 0.572, Z: 0.224

Log likelihood = -381.268 Significance = 0.017

Run # 24, 32 cells:

Convergence at Iteration 6

Input 0.098

Group # 2 -- r: 0.637, 2: 0.391, k: 0.557, n: 0.563

Group # 5 -- g: 0.453, h: 0.837

Group # 6 -- f: 0.428, m: 0.626

Group # 8 -- X: 0.576, Z: 0.213

Log likelihood = -378.011 Significance = 0.001

Run # 25, 16 cells:

Convergence at Iteration 6

Input 0.101

Group # 3 -- c: 0.462, e: 0.588

Group # 5 -- g: 0.455, h: 0.826

Group # 6 -- f: 0.430, m: 0.623

Group # 8 -- X: 0.578, Z: 0.207

Log likelihood = -383.053 Significance = 0.010

Run # 26, 29 cells:

Convergence at Iteration 5

Input 0.103

Group # 4 -- 8: 0.410, 5: 0.505, 7: 0.499, 6: 0.515

Group # 5 -- g: 0.455, h: 0.827

Group # 6 -- f: 0.430, m: 0.622

Group # 8 -- X: 0.575, Z: 0.217

Log likelihood = -385.898 Significance = 0.799

Run # 27, 8 cells:

Convergence at Iteration 5

Input 0.103

Group # 5 -- g: 0.455, h: 0.828

Group # 6 -- f: 0.430, m: 0.623

Group # 7 -- P: 0.558, T: 0.271

Group # 8 -- X: 0.518, Z: 0.425

Log likelihood = -386.402 Significance = 0.994

Add Group # 2 with factors r2kn

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

Run # 28, 103 cells:

Convergence at Iteration 7

Input 0.097

Group # 1 -- p: 0.463, o: 0.491, s: 0.640, v: 0.486

Group # 2 -- r: 0.632, 2: 0.398, k: 0.549, n: 0.572

Group # 5 -- g: 0.453, h: 0.839

Group # 6 -- f: 0.427, m: 0.628

Group # 8 -- X: 0.575, Z: 0.215

Log likelihood = -374.996 Significance = 0.114

Run # 29, 57 cells:

Convergence at Iteration 6

Input 0.096

Group # 2 -- r: 0.636, 2: 0.403, k: 0.553, n: 0.528

Group # 3 -- c: 0.470, e: 0.568

Group # 5 -- g: 0.453, h: 0.836

Group # 6 -- f: 0.428, m: 0.627

Group # 8 -- X: 0.578, Z: 0.207

Log likelihood = -376.194 Significance = 0.059

Run # 30, 97 cells:

Convergence at Iteration 6

Input 0.097

Group # 2 -- r: 0.642, 2: 0.389, k: 0.560, n: 0.555

Group # 4 -- 8: 0.396, 5: 0.521, 7: 0.505, 6: 0.497

Group # 5 -- g: 0.453, h: 0.836

Group # 6 -- f: 0.428, m: 0.626

Group # 8 -- X: 0.576, Z: 0.213

Log likelihood = -377.280 Significance = 0.692

Run # 31, 32 cells:

Convergence at Iteration 5

Input 0.098

Group # 2 -- r: 0.637, 2: 0.391, k: 0.557, n: 0.563

Group # 5 -- g: 0.453, h: 0.837

Group # 6 -- f: 0.428, m: 0.626

Group # 7 -- P: 0.559, T: 0.268

Group # 8 -- X: 0.518, Z: 0.425

*** Warning, negative change in likelihood (-0.00013847) replaced by 0.0.

Log likelihood = -378.011 Significance = 1.000

No remaining groups significant

Groups selected while stepping up: 5 8 6 2

Best stepping up run: #24

Stepping down...

----- Level # 8 -----

Run # 32, 243 cells:

Convergence at Iteration 6

Input 0.095

Group # 1 -- p: 0.416, o: 0.460, s: 0.672, v: 0.530

Group # 2 -- r: 0.619, 2: 0.424, k: 0.529, n: 0.546

Group # 3 -- c: 0.465, e: 0.581

Group # 4 -- 8: 0.379, 5: 0.530, 7: 0.484, 6: 0.501

Group # 5 -- g: 0.453, h: 0.835

Group # 6 -- f: 0.428, m: 0.626

Group # 7 -- P: 0.560, T: 0.264

Group # 8 -- X: 0.518, Z: 0.425

Log likelihood = -371.350

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

Run # 33, 134 cells:

Convergence at Iteration 6

Input 0.096

Group # 2 -- r: 0.639, 2: 0.402, k: 0.552, n: 0.528

Group # 3 -- c: 0.469, e: 0.570

Group # 4 -- 8: 0.428, 5: 0.498, 7: 0.537, 6: 0.500

Group # 5 -- g: 0.453, h: 0.835

Group # 6 -- f: 0.428, m: 0.627

Group # 7 -- P: 0.560, T: 0.263

Group # 8 -- X: 0.518, Z: 0.424

Log likelihood = -375.726 Significance = 0.036

Run # 34, 119 cells:

Convergence at Iteration 10

Input 0.097

Group # 1 -- p: 0.408, o: 0.413, s: 0.701, v: 0.565

Group # 3 -- c: 0.456, e: 0.600

Group # 4 -- 8: 0.366, 5: 0.541, 7: 0.462, 6: 0.502

Group # 5 -- g: 0.454, h: 0.831

Group # 6 -- f: 0.430, m: 0.623

Group # 7 -- P: 0.558, T: 0.269

Group # 8 -- X: 0.517, Z: 0.426

Log likelihood = -374.883 Significance = 0.074

Run # 35, 206 cells:

Convergence at Iteration 6

Input 0.095

Group # 1 -- p: 0.449, o: 0.444, s: 0.684, v: 0.520

Group # 2 -- r: 0.623, 2: 0.409, k: 0.542, n: 0.558

Group # 4 -- 8: 0.344, 5: 0.561, 7: 0.448, 6: 0.492

Group # 5 -- g: 0.453, h: 0.838

Group # 6 -- f: 0.428, m: 0.626

Group # 7 -- P: 0.558, T: 0.270

Group # 8 -- X: 0.517, Z: 0.427

Log likelihood = -372.984 Significance = 0.075

Run # 36, 145 cells:

Convergence at Iteration 6

Input 0.095

Group # 1 -- p: 0.418, o: 0.481, s: 0.651, v: 0.518

Group # 2 -- r: 0.622, 2: 0.421, k: 0.530, n: 0.549

Group # 3 -- c: 0.459, e: 0.595
 Group # 5 -- g: 0.453, h: 0.836
 Group # 6 -- f: 0.428, m: 0.627
 Group # 7 -- P: 0.560, T: 0.262
 Group # 8 -- X: 0.518, Z: 0.425
 Log likelihood = -372.162 Significance = 0.658

Run # 37, 190 cells:

Convergence at Iteration 7
 Input 0.107
 Group # 1 -- p: 0.395, o: 0.472, s: 0.656, v: 0.538
 Group # 2 -- r: 0.607, 2: 0.440, k: 0.509, n: 0.564
 Group # 3 -- c: 0.460, e: 0.590
 Group # 4 -- 8: 0.371, 5: 0.523, 7: 0.484, 6: 0.510
 Group # 6 -- f: 0.427, m: 0.629
 Group # 7 -- P: 0.556, T: 0.277
 Group # 8 -- X: 0.517, Z: 0.429
 Log likelihood = -398.910 Significance = 0.000

Run # 38, 162 cells:

Convergence at Iteration 7
 Input 0.099
 Group # 1 -- p: 0.414, o: 0.456, s: 0.673, v: 0.535
 Group # 2 -- r: 0.617, 2: 0.430, k: 0.521, n: 0.546
 Group # 3 -- c: 0.464, e: 0.582
 Group # 4 -- 8: 0.367, 5: 0.532, 7: 0.480, 6: 0.503
 Group # 5 -- g: 0.453, h: 0.838
 Group # 7 -- P: 0.561, T: 0.262
 Group # 8 -- X: 0.518, Z: 0.423
 Log likelihood = -380.209 Significance = 0.000

Run # 39, 243 cells:

Convergence at Iteration 6
 Input 0.095

Group # 1 -- p: 0.416, o: 0.460, s: 0.672, v: 0.530
 Group # 2 -- r: 0.619, 2: 0.424, k: 0.529, n: 0.546
 Group # 3 -- c: 0.465, e: 0.580
 Group # 4 -- 8: 0.378, 5: 0.530, 7: 0.484, 6: 0.501
 Group # 5 -- g: 0.453, h: 0.835
 Group # 6 -- f: 0.429, m: 0.626
 Group # 8 -- X: 0.577, Z: 0.210
 Log likelihood = -371.354 Significance = 0.933

Run # 40, 243 cells:

Convergence at Iteration 6

Input 0.095

Group # 1 -- p: 0.416, o: 0.460, s: 0.672, v: 0.530
 Group # 2 -- r: 0.619, 2: 0.424, k: 0.529, n: 0.546
 Group # 3 -- c: 0.465, e: 0.580
 Group # 4 -- 8: 0.378, 5: 0.530, 7: 0.484, 6: 0.501
 Group # 5 -- g: 0.453, h: 0.835
 Group # 6 -- f: 0.429, m: 0.626
 Group # 7 -- P: 0.577, T: 0.210
 Log likelihood = -371.354 Significance = 0.933

Cut Group # 7 with factors PT

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

Run # 41, 134 cells:

Convergence at Iteration 7

Input 0.096

Group # 2 -- r: 0.639, 2: 0.402, k: 0.552, n: 0.527
 Group # 3 -- c: 0.469, e: 0.571
 Group # 4 -- 8: 0.428, 5: 0.498, 7: 0.538, 6: 0.500
 Group # 5 -- g: 0.453, h: 0.835
 Group # 6 -- f: 0.428, m: 0.627
 Group # 8 -- X: 0.578, Z: 0.208

Log likelihood = -375.726 Significance = 0.036

Run # 42, 119 cells:

Convergence at Iteration 10

Input 0.097

Group # 1 -- p: 0.408, o: 0.413, s: 0.701, v: 0.565

Group # 3 -- c: 0.456, e: 0.600

Group # 4 -- 8: 0.366, 5: 0.541, 7: 0.462, 6: 0.502

Group # 5 -- g: 0.454, h: 0.831

Group # 6 -- f: 0.430, m: 0.623

Group # 8 -- X: 0.576, Z: 0.215

Log likelihood = -374.888 Significance = 0.074

Run # 43, 206 cells:

Convergence at Iteration 6

Input 0.095

Group # 1 -- p: 0.449, o: 0.444, s: 0.684, v: 0.520

Group # 2 -- r: 0.623, 2: 0.409, k: 0.542, n: 0.558

Group # 4 -- 8: 0.344, 5: 0.561, 7: 0.448, 6: 0.492

Group # 5 -- g: 0.453, h: 0.838

Group # 6 -- f: 0.428, m: 0.626

Group # 8 -- X: 0.575, Z: 0.216

Log likelihood = -372.988 Significance = 0.075

Run # 44, 145 cells:

Convergence at Iteration 6

Input 0.095

Group # 1 -- p: 0.418, o: 0.481, s: 0.651, v: 0.518

Group # 2 -- r: 0.622, 2: 0.421, k: 0.530, n: 0.549

Group # 3 -- c: 0.459, e: 0.595

Group # 5 -- g: 0.453, h: 0.836

Group # 6 -- f: 0.428, m: 0.627

Group # 8 -- X: 0.578, Z: 0.208

Log likelihood = -372.165 Significance = 0.658

Run # 45, 190 cells:

Convergence at Iteration 7

Input 0.107

Group # 1 -- p: 0.395, o: 0.472, s: 0.656, v: 0.538

Group # 2 -- r: 0.607, 2: 0.440, k: 0.509, n: 0.564

Group # 3 -- c: 0.461, e: 0.590

Group # 4 -- 8: 0.370, 5: 0.524, 7: 0.484, 6: 0.510

Group # 6 -- f: 0.427, m: 0.629

Group # 8 -- X: 0.573, Z: 0.223

Log likelihood = -398.912 Significance = 0.000

Run # 46, 162 cells:

Convergence at Iteration 7

Input 0.099

Group # 1 -- p: 0.414, o: 0.456, s: 0.673, v: 0.535

Group # 2 -- r: 0.617, 2: 0.430, k: 0.522, n: 0.546

Group # 3 -- c: 0.465, e: 0.581

Group # 4 -- 8: 0.366, 5: 0.533, 7: 0.480, 6: 0.503

Group # 5 -- g: 0.453, h: 0.838

Group # 8 -- X: 0.578, Z: 0.206

Log likelihood = -380.212 Significance = 0.000

Run # 47, 181 cells:

Convergence at Iteration 8

Input 0.106

Group # 1 -- p: 0.440, o: 0.429, s: 0.683, v: 0.539

Group # 2 -- r: 0.611, 2: 0.429, k: 0.529, n: 0.533

Group # 3 -- c: 0.472, e: 0.564

Group # 4 -- 8: 0.354, 5: 0.534, 7: 0.481, 6: 0.503

Group # 5 -- g: 0.455, h: 0.828

Group # 6 -- f: 0.426, m: 0.630

Log likelihood = -383.974 Significance = 0.000

Cut Group # 4 with factors 8576

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

Run # 48, 57 cells:

Convergence at Iteration 6

Input 0.096

Group # 2 -- r: 0.636, 2: 0.403, k: 0.553, n: 0.528

Group # 3 -- c: 0.470, e: 0.568

Group # 5 -- g: 0.453, h: 0.836

Group # 6 -- f: 0.428, m: 0.627

Group # 8 -- X: 0.578, Z: 0.207

Log likelihood = -376.194 Significance = 0.046

Run # 49, 61 cells:

Convergence at Iteration 7

Input 0.098

Group # 1 -- p: 0.409, o: 0.441, s: 0.674, v: 0.550

Group # 3 -- c: 0.447, e: 0.621

Group # 5 -- g: 0.454, h: 0.831

Group # 6 -- f: 0.429, m: 0.624

Group # 8 -- X: 0.576, Z: 0.212

Log likelihood = -376.071 Significance = 0.050

Run # 50, 103 cells:

Convergence at Iteration 7

Input 0.097

Group # 1 -- p: 0.463, o: 0.491, s: 0.640, v: 0.486

Group # 2 -- r: 0.632, 2: 0.398, k: 0.549, n: 0.572

Group # 5 -- g: 0.453, h: 0.839

Group # 6 -- f: 0.427, m: 0.628

Group # 8 -- X: 0.575, Z: 0.215

Log likelihood = -374.996 Significance = 0.018

Run # 51, 103 cells:

Convergence at Iteration 7

Input 0.107

Group # 1 -- p: 0.399, o: 0.489, s: 0.636, v: 0.528

Group # 2 -- r: 0.609, 2: 0.437, k: 0.510, n: 0.565

Group # 3 -- c: 0.454, e: 0.604

Group # 6 -- f: 0.426, m: 0.630

Group # 8 -- X: 0.574, Z: 0.219

Log likelihood = -399.900 Significance = 0.000

Run # 52, 94 cells:

Convergence at Iteration 7

Input 0.100

Group # 1 -- p: 0.415, o: 0.478, s: 0.650, v: 0.522

Group # 2 -- r: 0.619, 2: 0.428, k: 0.523, n: 0.549

Group # 3 -- c: 0.457, e: 0.597

Group # 5 -- g: 0.453, h: 0.839

Group # 8 -- X: 0.579, Z: 0.204

Log likelihood = -381.241 Significance = 0.000

Run # 53, 100 cells:

Convergence at Iteration 6

Input 0.106

Group # 1 -- p: 0.443, o: 0.453, s: 0.661, v: 0.524

Group # 2 -- r: 0.613, 2: 0.426, k: 0.531, n: 0.537

Group # 3 -- c: 0.465, e: 0.580

Group # 5 -- g: 0.455, h: 0.829

Group # 6 -- f: 0.425, m: 0.631

Log likelihood = -385.333 Significance = 0.000

Cut Group # 2 with factors r2kn

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

Run # 54, 16 cells:

Convergence at Iteration 6

Input 0.101

Group # 3 -- c: 0.462, e: 0.588

Group # 5 -- g: 0.455, h: 0.826

Group # 6 -- f: 0.430, m: 0.623

Group # 8 -- X: 0.578, Z: 0.207

Log likelihood = -383.053 Significance = 0.005

Run # 55, 36 cells:

Convergence at Iteration 6

Input 0.101

Group # 1 -- p: 0.473, o: 0.436, s: 0.669, v: 0.518

Group # 5 -- g: 0.454, h: 0.834

Group # 6 -- f: 0.429, m: 0.625

Group # 8 -- X: 0.572, Z: 0.224

Log likelihood = -381.268 Significance = 0.002

Run # 56, 36 cells:

Convergence at Iteration 7

Input 0.109

Group # 1 -- p: 0.399, o: 0.455, s: 0.654, v: 0.551

Group # 3 -- c: 0.445, e: 0.626

Group # 6 -- f: 0.427, m: 0.629

Group # 8 -- X: 0.573, Z: 0.222

Log likelihood = -403.136 Significance = 0.000

Run # 57, 33 cells:

Convergence at Iteration 7

Input 0.103

Group # 1 -- p: 0.407, o: 0.441, s: 0.671, v: 0.552

Group # 3 -- c: 0.446, e: 0.622

Group # 5 -- g: 0.454, h: 0.834

Group # 8 -- X: 0.577, Z: 0.209

Log likelihood = -384.804 Significance = 0.000

Run # 58, 36 cells:

Convergence at Iteration 8

Input 0.109

Group # 1 -- p: 0.429, o: 0.417, s: 0.682, v: 0.555

Group # 3 -- c: 0.454, e: 0.605

Group # 5 -- g: 0.456, h: 0.825

Group # 6 -- f: 0.427, m: 0.628

Log likelihood = -388.784 Significance = 0.000

All remaining groups significant

Groups eliminated while stepping down: 7 4 2

Best stepping up run: #24

Best stepping down run: #49

