

**UNIVERSIDADE ESTADUAL DO CENTRO-OESTE – UNICENTRO
PROGRAMA DE PÓS-GRADUAÇÃO - MESTRADO EM LETRAS**

TATIANE DA LUZ WALCZAK DA FONSECA

CONCORDÂNCIA NOMINAL EM GUARAPUAVA, PARANÁ

**GUARAPUAVA
2016**

**UNIVERSIDADE ESTADUAL DO CENTRO-OESTE – UNICENTRO
PROGRAMA DE PÓS-GRADUAÇÃO - MESTRADO EM LETRAS**

TATIANE DA LUZ WALCZAK DA FONSECA

CONCORDÂNCIA NOMINAL EM GUARAPUAVA, PARANÁ

Dissertação apresentada ao Programa de Pós-Graduação-
Mestrado em Letras, da Universidade Estadual do
Centro-Oeste – Unicentro, para a obtenção do título de
mestre em Letras.

Orientadora: Prof^a. Dr^a. Loremi Loregian-Penkal
Coorientadora: Prof^a. Dr^a. Lucelene Teresinha
Franceschini.

**GUARAPUAVA
2016**

Catálogo na Publicação
Biblioteca Central da Unicentro, Campus Cedeteg

Fonseca, Tatiane da Luz Walczak da

F676c Concordância nominal em Guarapuava, Paraná / Tatiane da Luz
Walczak da Fonseca. -- Guarapuava, 2016

xi, 100 f. : il. ; 28 cm

Dissertação (mestrado) - Universidade Estadual do Centro-Oeste,
Programa de Pós-Graduação em Letras, 2016

Orientadora: Loremi Loregian-Penkal

Coorientadora: Lucelene Teresinha Franceschini

Banca examinadora: Loremi Loregian-Penkal, Lucelene Teresinha
Franceschini, Edson Fagundes, Maria Cleci Venturini

Bibliografia

1. Concordância nominal. 2. Sintagma nominal. 3. Sociolinguística
variacionista. 4. VARLINGUA. I. Título. II. Programa de Pós-Graduação em
Letras.

CDD 469.5



TERMO DE APROVAÇÃO

Tatiane da Luz Walczak da Fonseca

Concordância Nominal em Guarapuava, Paraná

Dissertação aprovada em 14/12/2016 como requisito parcial para obtenção do grau de Mestre no Curso de pós-Graduação em Letras, da Unicentro Estadual do Centro-Oeste, UNICENTRO, área de concentração Interfaces entre Língua e Literatura, pela seguinte Banca Examinadora:

Prof.(a) Dr.(a) Loremi Loregian Penkal - UNICENTRO – Presidente/Orientador(a)

Prof.(a) Dr.(a) Lucelene Teresinha Francheschini – UNICENTRO- Co-orientador(a)

Prof.(a) Dr.(a) Edson Fagundes - UTFPR – Membro Titular

Prof.(a) Dr.(a) Maria Cleci Venturini- UNICENTRO – Membro Titular

“Quase me apetece dizer que não há uma língua portuguesa, há línguas em português.”

José Saramago

AGRADECIMENTOS

A Deus, pela saúde e por mais um objetivo realizado.

À minha orientadora, Loremi Loregian Penkal e à minha coorientadora, Lucelene Franceschini, pelas valiosas orientações, dedicação e comprometimento.

Aos professores, Maria Cleci Venturini e Edson Fagundes, pela leitura do trabalho, e pelas relevantes contribuições.

A todos os professores do mestrado em Letras da Unicentro.

À minha família, especialmente ao meu esposo Gelson e à minha filha Valéria, pelo apoio, amor e paciência.

Aos colegas do mestrado, pelas palavras amigas.

Aos informantes, fundamentais para a realização desta pesquisa, pela colaboração e paciência.

A todos aqueles que contribuíram, direta ou indiretamente, para a execução deste trabalho acadêmico.

FONSECA, Tatiane da Luz Walczak da. **Concordância nominal em Guarapuava, Paraná.** 101f. Dissertação (Mestrado em Letras) – Universidade Estadual do Centro-Oeste. Orientadora: Prof^a. Dr^a. Loremi Loregian Penkal. Coorientadora: Prof^a. Dr^a. Lucelene Teresinha Franceschini. Guarapuava, 2016.

RESUMO

O presente estudo objetiva analisar o uso da regra de concordância nominal de número plural em Guarapuava, Paraná. Primeiramente, discorremos brevemente sobre a teoria Sociolinguística Variacionista, na qual nos apoiamos nesta pesquisa e, também, apresentamos alguns trabalhos de seu precursor, William Labov ([1972] 2008). Na sequência, abordamos o nosso objeto de estudo, a concordância nominal e apresentamos brevemente alguns estudos anteriores sobre o tema, são eles: Braga & Scherre (1976); Braga (1977); Poplack (1980); Scherre (1988; 1998); e Fernandes (1996). Para analisar a concordância nominal em Guarapuava, as variáveis linguísticas inicialmente consideradas neste trabalho foram: *saliência fônica, tonicidade, posição do elemento no SN, classe gramatical do elemento e marcas precedentes*. Em seguida, realizamos outra rodada aliando as variáveis *saliência fônica e tonicidade*, assim como as variáveis *posição do elemento no SN e classe gramatical*. As variáveis sociais consideradas foram: *sexo, faixa etária e escolaridade*. Os dados analisados são pertencentes ao banco de dados de Guarapuava (VARLINGUA) e compostos por 24 entrevistas, distribuídas por sexo (*feminino e masculino*), faixa etária (*25 a 45 anos e 50 anos ou mais*) e escolaridade (*fundamental I, fundamental II e ensino médio*), transcritas e codificadas. Para a análise dos dados, utilizamos o programa estatístico GoldVarbX. Os fatores linguísticos que favoreceram a *presença de CN* foram: os *determinantes* (artigos, possessivos, indefinidos e quantificadores) *antepostos ao núcleo na 2ª e 1ª posição*, a *ausência de elemento precedente*, a *ausência de marca morfológica no elemento precedente*, o *plural duplo*, os *itens terminados em -l* e os *itens terminados em -s*. Os fatores sociais que favoreceram a *presença de CN* foram: os falantes com *ensino médio* e os falantes do sexo *feminino*. Com este estudo, esperamos contribuir para a descrição da concordância nominal e para os estudos da variação linguística na região Sul do Brasil.

Palavras-chave: concordância nominal; sintagma nominal; sociolinguística variacionista; VARLINGUA.

FONSECA, Tatiane da Luz Walczak da. **Nominal agreement in Guarapuava, Paraná.** 101 f. Dissertation (Master of Arts) – Universidade Estadual do Centro-Oeste. Supervisor: Prof. Dr. Loremi Loregian Penkal. Joint supervisor: Prof. Dr. Lucelene Teresinha Franceschini. Guarapuava, 2016.

ABSTRACT

The current study has as objective to analyze the use of the nominal agreement rule for plural numbers in the municipality of Guarapuava, Paraná. At first, we briefly discuss the Variationist Sociolinguistic theory, in which this research is based, and then it is presented some of the works of its precursor, William Labov ([1972] 2008). In the following, we approach our study subject, the nominal agreement and briefly present some previous studies on the subject, being them: Braga & Scherre (1976); Braga (1977); Poplack (1980); Scherre (1988; 1998); and Fernandes (1996). For the analysis of the nominal agreement in Guarapuava, the linguistic variables initially considered were: *phonic salience*, *tonicity*, *position of the element in SN*, *grammatical class of the element and previous marks*. Afterwards, it was performed another round combining the variables allying the variables *phonic salience* and *tonicity*, as wells as the variables *position of the element in SN* and *grammatical class*. The social variables considered were: gender, age range and education. The analyzed data belong to the Guarapuava database (VARLINGUA), being composed by 24 interviews, distributed by gender (*female* and *male*), age range (*25 to 45 years old and 50 years or more*) and education (*primary, elementary and high school*), transcribed and encoded. For the data analysis, we used the statistical software GoldVarbX. Linguistic factors that favored the *presence of CN* were: determinants (articles, possessives, indefinite and quantifiers), *prefixed to the word in the 2nd and 1st position*, *the absence of a preceding element*, *the absence of morphological mark in the previous element*, *the double plural*, *the items ending in -l* and *the items ending in -s*. The social factors that favored the *presence of CN* were: the speakers at studied only to the *elementary school* and the *female* speakers. With this study, we hope to contribute for the description of the nominal agreement and for the linguistic variation studies in the South of Brazil.

Keywords: nominal agreement; noun phrase; variationist sociolinguistic; VARLINGUA.

LISTA DE TABELAS

TABELA 1: Análise final de Processos e Tonicidade dos substantivos, adjetivos e categorias substantivadas dos adultos por escolarização.....	24
TABELA 2: Classe gramatical e posição no SN: Análise C.....	25
TABELA 3: Distribuição dos dados dos falantes adultos em função da classe gramatical e da posição dos elementos no SN.....	27
TABELA 4: Distribuição das classes gramaticais não nucleares em função da posição e da relação com o núcleo.....	28
TABELA 5: Processos e tonicidade dos substantivos, adjetivos e categorias substantivadas.	31
TABELA 6: Distribuição das classes gramaticais do SN em função da posição e da relação com o núcleo.....	32
TABELA 7: Graus da variável saliência fônica.....	44
TABELA 8: Constituição da variável tonicidade.....	45
TABELA 9: Resultados da CN em Guarapuava/PR- rodada final.....	57
TABELA 10: Resultados da posição na presença de CN.....	59
TABELA 11: Resultados da classe gramatical do elemento na presença de CN.....	60
TABELA 12: Resultados da posição e classe gramatical aliadas na presença de CN.....	63
TABELA 13: Resultados de marcas precedentes na presença de CN.....	66
TABELA 14: Resultados da escala da saliência fônica na presença de CN.....	69
TABELA 15: Resultados da tonicidade na presença de CN.....	70
TABELA 16: Resultados da saliência e tonicidade aliadas na presença de CN.....	72
TABELA 17: Resultados da saliência e tonicidade aliadas - Substantivos, Adjetivos e Categorias Substantivadas.....	74
TABELA 18: Resultados da escolaridade na presença de CN.....	77
TABELA 19: Resultados do sexo na presença de CN.....	79

LISTA DE QUADRO, FIGURAS E GRÁFICOS

QUADRO 1: Estratos sociais dos falantes de Guarapuava.....	39
FIGURA 1: Catedral Nossa Senhora de Belém – Guarapuava/Paraná.....	36
FIGURA 2: Mapa de Guarapuava/Paraná e região.....	37
GRÁFICO 1: Cruzamento da saliência e tonicidade aliadas com a escolaridade - Substantivos, adjetivos e categorias substantivadas.....	76
GRÁFICO 2: Cruzamento da escolaridade e sexo.....	80
GRÁFICO 3: Cruzamento da escolaridade e faixa etária.....	81
GRÁFICO 4: Cruzamento da faixa etária e sexo.....	82

SUMÁRIO

INTRODUÇÃO.....	11
CAPÍTULO 1 – A SOCIOLINGUÍSTICA VARIACIONISTA.....	14
CAPÍTULO 2 – CONCORDÂNCIA NOMINAL E ESTUDOS ANTERIORES.....	19
2.1. Concordância nominal.....	19
2.2. Revisão de estudos anteriores sobre CN.....	20
CAPÍTULO 3 – PROCEDIMENTOS METODOLÓGICOS.....	34
3.1. A comunidade de Guarapuava/Paraná.....	34
3.2. A constituição da amostra.....	37
3.3. A coleta de dados.....	39
3.4. O programa estatístico GoldVarb X.....	41
3.5. As variáveis e os grupos de fatores.....	42
3.5.1. A variável dependente CN.....	42
3.5.2. As variáveis independentes.....	43
3.5.2.1. Saliência fônica.....	43
3.5.2.2. Tonicidade.....	45
3.5.2.3. Posição do elemento no SN.....	46
3.5.2.4. Classe gramatical do elemento.....	48
3.5.2.5. Marcas precedentes.....	50
3.6. As variáveis sociais.....	52
3.6.1. A variável sexo.....	52
3.6.2. A variável faixa etária.....	53
3.6.3. A variável escolaridade.....	54
CAPÍTULO 4 – RESULTADOS DE CONCORDÂNCIA NOMINAL EM GUARAPUAVA.....	55
4.1. Resultados gerais da CN em Guarapuava.....	56
4.2. Resultados das variáveis linguísticas em Guarapuava.....	58
4.2.1. Posição e classe gramatical aliadas.....	58
4.2.2. Marcas precedentes.....	65
4.2.3. Saliência e tonicidade aliadas.....	68
4.3. Resultados das variáveis sociais em Guarapuava.....	77

4.3.1. Escolaridade.....	77
4.3.2. Sexo.....	78
4.4. Resultados dos cruzamentos das variáveis sociais em Guarapuava.....	79
4.4.1. Cruzamento: escolaridade e sexo.....	79
4.4.2. Cruzamento: Escolaridade e faixa etária.....	81
4.4.3. Cruzamento: Faixa etária e sexo.....	82
CONSIDERAÇÕES FINAIS.....	84
REFERÊNCIAS.....	88
ANEXOS.....	90

INTRODUÇÃO

A relação entre os seres humanos ocorre, principalmente, por meio da linguagem e do convívio em uma sociedade organizada. Nas línguas naturais, existem muitas variações que não impedem a comunicação. A variação linguística é um dos modos como uma língua pode fazer referência, por exemplo, a um elemento do mundo por mais de um termo.

Nesse sentido, Calvet (2002, p. 90) destaca: “Entenderemos aqui por *variável* o conjunto constituído pelos diferentes modos de realizar a mesma coisa (um fonema, um signo...) e por *variante* cada uma das formas de realizar a mesma coisa.”. Podemos dizer, então, que variantes são diferentes formas linguísticas que veiculam um mesmo sentido referencial acerca de um determinado fenômeno que possa variar.

Em Guarapuava, a título de ilustração, é muito comum na língua falada o emprego da palavra “piá”, que em outra comunidade pode ser utilizada de outra forma para corresponder à mesma variação lexical, “menino”, “garoto”, “guri”, entre outros. A área específica da Linguística que trata da relação intrínseca entre linguagem e sociedade e a questão da variação e da mudança linguística é a Sociolinguística Variacionista.

Esta área da ciência da linguagem, entre outros interesses, procura verificar de que modo fatores de natureza linguística e extralinguística estão relacionados ao uso das variantes em diferentes níveis. A Sociolinguística pretende desvendar como a heterogeneidade da língua – a variação – se organiza, e de que modo essa variação é regulada e se transforma, ou não, em mudança na língua.

Um dos temas de estudo da Sociolinguística, a concordância nominal (doravante CN), já foi analisado no Brasil com amostras de diversas regiões, tais como: o NURC (Norma Urbana Culta); o Projeto Censo da Variação Linguística do Estado do Rio de Janeiro; o VARSUL (Variação Linguística Urbana na Região Sul), que abrange Paraná, Santa Catarina e Rio Grande do Sul, entre outras. Por exemplo, o trabalho de Scherre (1988), e o de Fernandes (1996), referências para este estudo, analisaram a CN em amostras do Projeto Censo do Rio de Janeiro e do Projeto VARSUL.

Nossa análise dos dados foi feita por meio de uma amostra oriunda de Guarapuava do banco VARLINGUA¹ (Variação Linguística de Guarapuava), composta de 24 entrevistas, distribuídas por sexo (feminino e masculino), faixa etária (1 - 25 a 45 anos e 2 - 50 anos ou

¹ O banco VARLINGUA foi criado por pesquisadores da Universidade Estadual do Centro-Oeste – Unicentro, do Programa de Pós-Graduação em Letras. O banco de dados apoia-se teoricamente na Sociolinguística Variacionista e assemelha-se ao banco do Projeto VARSUL.

mais) e escolaridade² (Fundamental I, Fundamental II e Ensino Médio). A escolha do tema se fortaleceu pela carência de pesquisas sociolinguísticas na região de Guarapuava³.

Neste estudo, pretendemos analisar a CN em Guarapuava, Paraná, que apresenta duas variantes: *ausência de CN* e *presença de CN*, como procuram demonstrar os exemplos⁴ abaixo:

(1) Eu vejo *as cidade* crescendo, i se fosse pra mim saí de Guarapuava seria pro interior, lugar menor. (VARLINGUA⁵ Inf4 M1g);

(2) A grande coisa que *as pessoas* sempre falam é no friu daqui, porque é notícia né, o friu de Guarapuava. (VARLINGUA Inf12 M2c).

Como podemos observar, no exemplo 1, no sintagma nominal (doravante SN) “*as cidade*” o artigo *as* apresenta a *presença de CN* e o substantivo *cidade*, segundo elemento e núcleo do SN, apresenta a *ausência de CN*, ou seja, não recebe a marcação de plural. Já no exemplo 2, no SN “*as pessoas*”, tanto o determinante (artigo) *as*, quanto o núcleo (substantivo) *pessoas* apresentam a *presença de CN*.

O objetivo geral deste trabalho é analisar o uso da regra variável de CN de número plural em Guarapuava. Os objetivos específicos são:

1. Descrever e analisar as variantes usadas por falantes de diferentes estratos sociais pertencentes à zona urbana de Guarapuava;
2. Verificar de que modo os fatores linguísticos e sociais atuam no condicionamento da variação de número no SN na língua falada em Guarapuava;
3. Contribuir à descrição da CN e para os estudos da variação linguística do Português falado no Brasil (doravante PB), sobretudo da região do Paraná.

As principais hipóteses levantadas foram as seguintes:

1. A comunidade de Guarapuava utiliza um maior número de ocorrências constituídas pela *ausência de CN* na língua falada que pelas formas marcadas pela *presença de CN* no SN;

² A proposta inicial deste estudo era de analisar o Ensino Superior na variável escolaridade. No entanto, não foi possível devido ao tempo estipulado para a realização do mesmo. Futuramente, em novas pesquisas na comunidade de Guarapuava, esse nível de escolaridade será adotado.

³ A primeira pesquisa sociolinguística realizada na comunidade de Guarapuava foi apresentada pela mestranda Vanessa Aparecida Deon, em 2015, cujo tema é “Variação pronominal nós/a gente em Guarapuava – PR”, ao Programa de Pós-Graduação em Letras da Universidade Estadual do Centro-Oeste – Unicentro.

⁴ Para destacar os exemplos apresentados na dissertação, colocamos o elemento em análise em negrito e todo o SN em itálico.

⁵ Os dados entre parênteses identificam o *sexo*: (F) *Feminino*, (M) *Masculino*; a *faixa etária*: (1) *25 a 45 anos*, (2) *50 anos ou mais*; e a *escolaridade* dos informantes: (p) *Fundamental I*, (g) *Fundamental II*, (c) *Ensino médio*.

2. Os moradores dessa comunidade com maior grau de escolaridade apresentam um maior número de marcas de plural na língua falada que os moradores com menor grau de escolaridade;
3. Os itens mais salientes favorecem a *presença de CN* na língua falada em Guarapuava e os resultados são semelhantes aos de outras regiões como Rio de Janeiro e Florianópolis, por exemplo;
4. As classes gramaticais dos substantivos, dos adjetivos e das categorias substantivadas desfavorecem a aplicação da regra de CN em Guarapuava.

Este estudo justifica-se pelo fato de o PB apresentar casos de variação linguística e a CN ser uma das variáveis passíveis de análise, reforçada pela evidência de que ocorrem frequentemente casos de *ausência de CN* entre os elementos do SN em Guarapuava. Procuraremos também verificar se ocorre uma mudança em curso da forma *presença de CN* para a forma *ausência de CN* na comunidade analisada.

Em linhas gerais, nosso estudo assim se organiza: no primeiro capítulo, apresentamos brevemente a teoria norteadora deste estudo, a Sociolinguística Variacionista e, na sequência, alguns trabalhos de seu precursor, William Labov. No segundo capítulo, vamos discorrer sobre o nosso objeto de estudo, a CN de número plural. Apresentamos, na perspectiva da Sociolinguística, a regra variável de CN e uma síntese de alguns trabalhos sobre a CN, incluindo principalmente os estudos de Maria Marta Pereira Scherre, pioneira no estudo do tema aqui no Brasil.

No terceiro capítulo, abordamos os procedimentos metodológicos desta pesquisa, que inclui a comunidade de Guarapuava, a constituição da amostra, a coleta de dados, o programa estatístico e as variáveis analisadas. O quarto capítulo traz os resultados gerais da pesquisa e os resultados de cada variável analisada no programa estatístico GoldVarbX para o fenômeno em questão, a CN em Guarapuava. Por fim, apresentamos as considerações finais deste estudo, com as conclusões obtidas e algumas reflexões para dar luz às futuras pesquisas.

CAPÍTULO 1 - A SOCIOLINGUÍSTICA VARIACIONISTA

O século XX foi marcado por diferentes correntes teóricas, no que se refere à Linguística, e as principais foram o estruturalismo (Saussure) e o gerativismo (Chomsky). Ferdinand de Saussure (2008) definiu a língua como o principal objeto de estudos da Linguística em sua obra Curso de Linguística Geral (CLG), publicada em 1916, a partir de anotações de aula de seus alunos, Charles Bally e Albert Sechehaye, durante cursos que ele ministrou na Universidade de Genebra.

Saussure estabeleceu a dicotomia língua (*langue*) e fala (*parole*) e considerou apenas a língua, tida como um sistema fechado de signos, homogênea e social, como objeto de estudo da Linguística. O linguista não considerou a fala (*parole*), que resulta das combinações feitas pelo falante por meio do código da língua, por ser individual e heterogênea. A partir do recorte teórico adotado, o objetivo do linguista seria descrever uma língua em um único estado, em um único momento, sem eventuais mudanças e variações.

Segundo Petter (2011), em oposição à teoria de Saussure, a partir dos anos de 1950, o linguista americano Noam Chomsky trouxe uma nova teoria para os estudos linguísticos, o gerativismo. Esta corrente teórica afirmava que o conhecimento linguístico era adquirido pela faculdade da linguagem, uma propriedade inata transmitida geneticamente a todo ser humano. Este estudo consistia em explicar as frases existentes na língua e as frases potencialmente realizáveis pelos falantes, posto que, a partir de um número limitado de regras, era possível gerar um número infinito de sentenças, independente da relação entre a linguagem e o mundo.

Muitos pesquisadores estudaram a relação entre linguagem e sociedade, entre eles, conforme Pietroforte (2011), o linguista Coseriu, que percebeu a realização da fala com suas diversas formas (sotaques, vocabulários próprios de grupos sociais) não de maneira individual, tampouco comum a todos os falantes de uma determinada língua, ou seja, Coseriu percebeu que cada comunidade linguística adota na língua falada algumas variantes linguísticas.

Dadas essas constatações, esse linguista redefiniu a dicotomia língua e fala de Saussure, na qual incluiu a noção de norma, passando a sistema/norma/fala, da seguinte maneira:

O *sistema* funcional coincide com o conceito de *língua* de Saussure, no entanto, o que Coseriu chama *língua* é o *sistema* articulado com suas *normas*, ou seja, com suas variantes linguísticas. Assim, o conceito de língua, para Coseriu, abrange o *sistema*, que é do domínio de todos os falantes de uma mesma língua, e as *normas*, que, como variantes desse

sistema, são do domínio de grupos sociais, regionais, etc. (COSERIU *apud* PIETROFORTE, 2011, p. 92).

Após o reconhecimento desse novo sistema da língua, novos estudos foram surgindo. Conforme Coelho *et al* (2015), foi a partir de 1966, nos Estados Unidos, no simpósio “Direções para a Linguística Histórica”, organizado por Weinreich, Labov e Herzog, que os estudos sobre a variação e a mudança linguísticas fundamentaram uma teoria-metodológica, até porque os autores tinham

[...] razões convincentes para modificar essa posição nos fatos confirmados de que os desvios de um sistema homogêneo não são todos eles erros aleatórios de desempenho, mas são num alto grau codificados e parte de uma descrição realista da competência de um membro de uma comunidade de fala. (WEINREICH, LABOV, HERZOG, 2006, p. 60).

Nesse contexto, uma nova teoria linguística, denominada Sociolinguística⁶ entra em cena, e tem como princípio a noção de língua como um sistema inerentemente variável. A Teoria da variação e mudança tem como objeto de estudo a língua falada em um contexto social, ou seja, em situações reais de uso da língua.

Como a questão da mudança sempre instigou os estudos linguísticos, a Sociolinguística procurou estudar o fenômeno da mudança, enquanto um processo, que, segundo Weinreich, Labov e Herzog (2006, p. 122), ocorre em três etapas: “(1) à medida que um falante aprende uma forma alternativa, (2) durante o tempo em que as duas formas existem em contato dentro de sua competência, e (3) quando uma das formas se torna obsoleta”.

Para que ocorra uma mudança na língua, é necessário o surgimento da variação na língua por um longo período de tempo e que contenha duas formas variantes que irão concorrer e coocorrer, isto é, as duas variantes serão usadas ao mesmo tempo. Com isso, pode haver uma mudança em progresso ou a variação pode permanecer estável.

A mudança em progresso pode ser constatada quando na língua falada dos mais jovens, predomina a variante inovadora. Nesse caso, supõe-se que esses falantes, após um determinado período de tempo, quando mais velhos, continuarão a usar a variante inovadora, isto é, a forma que não é a padrão e, muitas vezes, é a forma estigmatizada pela comunidade linguística.

⁶ As principais obras (que são até hoje referências nos estudos sociolinguísticos) são: “Fundamentos empíricos para uma teoria da mudança linguística”, de 1968, publicada por Weinreich, Labov e Herzog, e “Padrões sociolinguísticos”, de 1972, publicada por Labov.

Na esteira social, o processo de mudança pode estar relacionado às novas gerações de dada comunidade de fala, porque a “transferência parece ocorrer entre grupos de pares de faixas etárias levemente diferentes” (Weinreich, Labov e Herzog, 2006, p. 122).

É importante esclarecer que, segundo Weinreich, Labov e Herzog (2006), nem toda variação e heterogeneidade na estrutura da língua podem levar à mudança, porém, toda mudança apresenta variação e heterogeneidade em sua estrutura. Evidentemente, um dos objetivos do presente estudo é verificar se há uma possível mudança em progresso na comunidade de Guarapuava da forma *presença de CN* para a forma *ausência de CN*.

Para tanto, à luz da Sociolinguística, nos embasaremos principalmente no pesquisador William Labov ([1972] 2008), considerado o grande nome da teoria variacionista. Para sistematizar, descrever e interpretar os fenômenos da variação linguística, Labov identificou fatores sociais como idade, sexo, escolaridade, etnia, que atuam como condicionadores do aparecimento das variações e das mudanças nas línguas. Por consequência, sua pesquisa conseguiu demonstrar como a avaliação dos dados linguísticos e a variação constatada, podem ser explicados por meio dos elementos sociais considerados.

Como a metodologia da pesquisa sociolinguística visa a analisar a língua falada de um determinado grupo de pessoas, é provável que o pesquisador se depare com a realidade da variação. A esse grupo a Sociolinguística denomina comunidade linguística ou comunidade de fala, que, pela definição de Labov ([1972] 2008, p. 188), a comunidade “não pode ser concebida como um grupo de falantes que usam todas as mesmas formas; ela é mais bem definida como um grupo que compartilha as mesmas normas a respeito da língua.”

Além disso, cada comunidade de fala compartilha um conjunto de regras de uso linguístico, porém, todos os membros de uma comunidade não falam da mesma maneira, no entanto, observamos que

[...] é comum muitos indivíduos manterem, principalmente na fala, a variante típica de seu grupo social para não serem “excluídos” e preservarem sua identificação com este grupo, mesmo que esta variante seja aquela considerada não padrão. (MARIANO, 2013, p. 17).

Tendo essas informações, passamos agora a apresentar dois dos primeiros trabalhos de Labov, relevantes para este estudo, a fim de demonstrar como a avaliação dos dados linguísticos e a variação constatada, podem ser explicadas por meio dos elementos sociais considerados e, por isso, justificam as conclusões apresentadas pelo autor.

Labov ([1972] 2008) realizou um estudo na ilha de *Martha’s Vineyard*, localizada em Nova Iorque, no estado de *Massachusetts*, com o enfoque em duas semivogais: a pronúncia

dos ditongos /ay/ e /aw/. O pesquisador analisou as diferenças na altura do primeiro elemento, o /a/ e verificou que esta vogal tende a ser centralizada em sua pronúncia, ou seja, há uma pronúncia inconsciente diferenciada da vogal que faz com que a pronúncia pareça mais próxima do /e/, como nas ocorrências [ey] e [ew] ou, ainda, [ay] e [aw].

Para explicar tal fato de variação fonética, Labov ([1972] 2008) recorre teoricamente aos elementos sociais, buscando correlacionar o traço linguístico – a centralização dos ditongos - com os traços sociais, por meio da distribuição geográfica da ilha, dos grupos sociais e da etnia. Uma suposição é de que o primeiro elemento do ditongo /ay/ era uma vogal centralizada na pronúncia utilizada nos Estados Unidos.

Os vineyardenses eram movidos por pressões econômicas, mas, mesmo diante das dificuldades em continuar morando na ilha, devido à alta taxa de desemprego e ao alto custo de vida, o estudo de Labov concluiu que os falantes com alta centralização de /ay/ e /aw/ e que possuíam uma pronúncia insular, queriam permanecer na ilha, como um gesto de resistência às invasões dos veranistas. Por essas razões, ocorria uma tentativa dos vineyardenses em manter sua identidade por intermédio da língua falada.

Outro estudo feito por Labov ([1972] 2008) em Nova Iorque analisou a presença ou a ausência do /r/ pós-vocálico em palavras como *car*, *card*, *four*, por exemplo. Para isso, o autor observou as práticas linguísticas dos empregados de três grandes lojas de departamentos de *status* social diferentes: *Saks* (*status* superior); *Macy's* (*status* médio); *Klein* (*status* inferior).

A metodologia de Labov consistia em pedir informações sobre um determinado departamento, por exemplo: “Por favor, onde ficam os sapatos femininos?”. Como ele já sabia que ficavam no quarto andar, os vendedores respondiam de acordo com a forma esperada: *fourth floor*, (quarto andar). Distante do informante, Labov anotava por escrito o dado obtido na fala e assim, registrava a pronúncia do /r/.

Labov ([1972] 2008) realizou um cruzamento dos dados entre a variável fonológica, o /r/ e o tipo da loja, conforme a diferença de *status* entre elas, considerando aspectos como a localização, os preços, a política de publicidade e as instalações físicas. Seus resultados registraram que 62% dos empregados da *Saks*, 51% dos empregados da *Macy's* e 20% dos empregados da *Klein* pronunciavam o /r/ de uma forma marcada e evidente.

De acordo com sua hipótese inicial, Labov ([1972] 2008) identificou que a diferença do emprego do /r/ está para a mesma ordem da escala das lojas, isto é, da estratificação social. O pesquisador percebeu que os empregados da *Saks* pronunciavam com maior ocorrência o /r/ final, pois esta forma quando empregada na cidade era considerada como variante de prestígio, por conseguinte, atuando como um diferenciador social.

A partir desses dados, podemos observar que quando admitimos a variedade de línguas no mundo, consideramos um rol de regras linguísticas que faz com que o PB seja diferente de outras línguas, e a variação se apresenta em diferentes níveis linguísticos: variação lexical; variação fonológica; variação morfológica; variação semântica; variação sintática; e variação discursiva. (Coelho *et al*, 2015).

As variações nas línguas estão relacionadas a diversos fatores linguísticos e aos fatores sociais, ou extralinguísticos. Com relação a este último, vale ressaltar que as diferenças entre comunidades e suas características sociais são de natureza quantitativa, isto é, podem ser mensuradas através dos resultados fornecidos pelo programa estatístico utilizado pelo pesquisador, como fez Labov, como fazem outros pesquisadores e como veremos também em nosso trabalho.

Além dos fatores sociais, temos os fatores linguísticos que podem influenciar a variação nas línguas. Na análise da CN, por exemplo, conforme nos apresenta Scherre (1988), temos: a saliência fônica, a tonicidade, a posição do elemento no SN, a classe gramatical do elemento (artigo, substantivo, adjetivo), e as marcas precedentes, dentre outros. No próximo capítulo, as variáveis mencionadas acima serão apresentadas por meio de estudos anteriores e, posteriormente, serão analisadas no presente estudo.

CAPÍTULO 2 - CONCORDÂNCIA NOMINAL E ESTUDOS ANTERIORES

Neste capítulo, apresentamos o nosso objeto de estudo, a CN de número plural, pela definição da gramática tradicional do PB, porém, considerando o fenômeno em variação, segundo a perspectiva da Sociolinguística variacionista. Na sequência, traremos um resumo de alguns estudos iniciais sobre o tema, que evidenciaram o papel das questões linguísticas e sociais para subsidiar a análise da CN que será efetuada em nosso estudo.

2.1. Concordância nominal

Conforme Scherre (1998), o PB prevê as regras gramaticais da concordância de gênero, de número e de pessoa. “A concordância de gênero (*a menina/o menino*) é especificamente nominal e a de pessoa (*nós estudamos/eles estudam*) é especificamente verbal. A concordância de número, por sua vez, pode ser nominal ou verbal.” (SCHERRE, 1998, p. 87). Este estudo trata da concordância de número plural entre os elementos do SN, que, por sua vez, é representada pelo acréscimo do –s, na língua portuguesa.

De acordo com a visão da gramática tradicional, Cegalla (2008, p. 438), destaca que “concordância é o princípio sintático segundo o qual as palavras dependentes se harmonizam, nas suas flexões, com as palavras de que dependem.”. Na perspectiva da Sociolinguística, a CN é a reiteração do mesmo conteúdo de um nome no determinante, quantificador ou adjetivo a ele inter-relacionado de forma sintática e semântica. (BRANDÃO, 2007).

Ainda de acordo com Brandão (2007), a concordância na categoria de número, mesmo pressupondo as regras tradicionais da língua portuguesa, está sujeita a variações, por isso, o uso das regras são variáveis, dependendo de fatores linguísticos e extralinguísticos, sem que se perca a inter-relação morfossintático-semântica entre os elementos. Como regra variável, a CN pode apresentar duas formas variantes: *ausência de CN* e *presença de CN*.

Talvez uma explicação para isso seja que “Nas mais recentes pesquisas, a perda de morfologia flexional e de regras de concordância vem sendo interpretada como decorrente de um processo de transmissão linguística irregular.” (BRANDÃO, 2007, p. 59).

A propósito da CN, a ocorrência de variedades de concordância pode estar relacionada aos vários constituintes que compõem o SN e aos fatores linguísticos e extralinguísticos que a condicionam, isto é, aos aspectos ligados à estrutura do SN, às alterações morfofonológicas do mecanismo de flexão, às características dos falantes como sexo, idade, escolaridade, ao contexto social, etc.

Conforme Brandão (2007), o SN é uma construção sintática que contém um elemento central (núcleo), que pode ser único ou acompanhado de outros constituintes como quantificadores ou determinantes. Os principais constituintes que apresentam a flexão de número ou indicam o plural são três classes: a. A que ocupa a posição à esquerda do núcleo (determinantes); b. A dos quantificadores; c. A dos adjetivos, à direita ou à esquerda do núcleo. Além disso, há fatores importantes que determinam o uso de variedades no SN, que são: a. A classe da palavra; b. A posição linear no sintagma; c. A saliência fônica, entre outros.

Para melhor entender a discussão apresentada até aqui, é necessário revisitar alguns dos trabalhos que trataram do tema. Na próxima seção, esses trabalhos são apresentados.

2.2. Revisão de estudos anteriores sobre CN

Nesta seção, apresentamos uma síntese de alguns estudos anteriores sobre a CN, nos quais iremos nos ater aos elementos relevantes de cada pesquisa, para que possam contribuir ao presente estudo em Guarapuava. Desse modo, não arrolamos toda a composição e o estudo de todas as variáveis consideradas pelos autores.

Os estudos elencados foram: Braga & Scherre (1976); Braga (1977); Poplack (1980); Scherre (1988; 1998); e Fernandes (1996). Ressaltamos, ainda, que o trabalho que tomamos como referência é o de Scherre (1988), por ser o grande nome no estudo do tema aqui no Brasil. Também apresentamos o estudo de Fernandes (1996), da região sul do estado, com o objetivo de realizarmos uma comparação entre os resultados obtidos pela autora com os de Guarapuava.

Apesar do processo da variação de CN ter sido estudado por vários pesquisadores, pressupomos que cada comunidade de fala possui características peculiares, que a diferenciam de outra comunidade. Por isso, consideramos fundamental o estudo acerca do mesmo fenômeno em diferentes regiões, a fim de verificar como esse fenômeno em variação é descrito e regulado. A partir daí buscamos analisar a CN em Guarapuava.

Braga & Scherre (1976)

Os primeiros estudos realizados sobre a análise da concordância entre os elementos do SN são de Braga e Scherre em 1976. Esta pesquisa foi realizada sobre o PB do Rio de Janeiro, com falantes de classes sociais e de procedência geográficas diferentes. A principal variável

trabalhada pelas autoras foi a saliência fônica⁷, que, segundo Scherre (1988), consiste em estabelecer que as formas mais salientes e mais perceptíveis, são mais propensas de serem marcadas que as menos salientes.

Para analisar esta variável na CN, Braga e Scherre apresentaram cinco níveis de diferenciação fônica:

- 1) inserção de -S e abertura vocálica ou plural metafônico (ovo/óvus);
 - 2) inserção de -S e mudança silábica (milhão/milhões; imóvel/imóveis);
 - 3) inserção de -ES em palavras em -R (flor/flores);
 - 4) inserção de -S em palavras de plural regular (livro/livros);
 - 5) inserção de -ES em palavras que terminam em -S (rapaz/rapazes).
- (SCHERRE, 1988, p. 66-67).

Dos níveis da saliência fônica considerados por Braga e Scherre como os mais salientes estão: o plural metafônico, os itens terminados em -ão ou -l e os itens terminados em -r. Logo abaixo na escala, os itens menos salientes foram: o plural regular e os itens terminados em -s.

Pelos resultados obtidos, Braga & Scherre chegaram à conclusão de que as formas mais salientes, como por exemplo, o plural metafônico, os itens terminados em -ão ou -l e os itens terminados em -r, são mais propensas a adquirir as marcas de plural que as formas menos salientes, neste caso, o plural regular e os itens terminados em -s. Além disso, que a escala da saliência fônica é mais recorrente nos falantes de classe média que nos falantes de classe baixa, mesmo havendo a ocorrência nas três classes distintas.

Braga (1977)

Em estudo desenvolvido por Braga em 1977, em continuidade ao estudo anterior (Braga & Scherre, 1976), a autora analisou a CN no PB com/o regra variável na língua falada de sete moradores do Triângulo Mineiro, com a faixa etária (15-20 anos), de classe média e baixa e com escolaridade dos níveis Ensino Fundamental e Ensino Médio.

A autora retomou a variável saliência fônica e trabalhou com as seguintes variáveis linguísticas: presença ou ausência de flexão no termo anterior ao elemento analisado; natureza fonológica do contexto seguinte; posição do elemento no SN; e grau de saliência fônica entre

⁷ O princípio da saliência fônica foi introduzido por Lemle e Naro (1976), em estudos sobre a concordância verbal do PB, e tem por objetivo determinar quais são os itens mais salientes e os menos salientes na formação de plural.

as formas (singular/plural). As variáveis sociais selecionadas para a análise foram: grau de formalidade da entrevista e classe social do falante.

Os resultados obtidos por Braga (1977), considerando as variáveis linguísticas e sociais, foram:

1. A probabilidade de um elemento do SN não receber a marca de plural é maior quando o elemento anterior não possui a marca, em casos com mais de dois elementos;
2. Quanto à natureza fonológica do contexto seguinte, Braga prefere considerar o contexto fonológico imediatamente anterior ao –s e a posição da sílaba tônica do elemento.
3. Os elementos favoráveis à *presença de CN* estão na primeira posição do SN;
4. O grau de formalidade se mostrou influente nos falantes de classe média;
5. Os falantes de classe média utilizam mais a forma *presença de CN* que os falantes de classe baixa, conforme o grau de saliência fônica, isto é, as formas mais salientes são mais utilizadas por eles.

Poplack (1980)

Segundo Scherre (1988), Poplack em 1980, realizou um estudo com o enfoque no cancelamento do –s plural no SN no Espanhol de Porto Rico. A autora analisou 6.349 dados retirados de uma amostra de 24 falantes do espanhol porto-riquenho (não com falantes do norte da Filadélfia), com idade acima de 21 anos.

A partir dos resultados, Poplack confirma sua hipótese, que a presença de marca de plural no elemento precedente favorece a supressão de marca no elemento considerado, de modo semelhante aos demonstrados em Braga (1977). Poplack conclui que o cancelamento da forma plural ocorre em decorrência das marcas precedentes, principalmente quando há duas marcas precedentes canceladas e quando o elemento precedente está na primeira posição.

Uma consideração relevante de Poplack para a variável classe gramatical: para um elemento que está posicionado à esquerda, o adjetivo desfavorece a aplicação da regra, enquanto que um determinante favorece a *presença de CN*.

Scherre (1988)

Scherre (1988), em sua tese de doutorado, intitulada *Reanálise da concordância nominal em português*, reanalisou a concordância de número no SN com uma amostra do banco de dados do *Corpus Censo* do Programa de Estudos sobre o Uso da Língua (PEUL),

grupo de pesquisa da Universidade Federal do Rio de Janeiro, com um total de 64 entrevistas, distribuídas por 48 adultos (15-71 anos) e 16 crianças (7-14 anos), sexo, faixa etária (7-14, 15-25, 26-49 e 50-71) e anos de escolarização (1-4, 5-8, 9-11).

Scherre (1988) retirou da amostra todos os SN's no plural, para verificar possíveis variações no uso dos elementos, considerando, por sua vez, dois casos de variação: pela presença (a variante marcada) e pela ausência (a variante não marcada) da forma do plural. Para isso, ela analisa os dados de duas formas: uma atomística, que analisa cada elemento flexionável do SN e outra não-atomística, que analisa o SN como um todo.

Assim, a autora propõe que há um controle sobre o uso da variação e isso vai depender de estruturas linguísticas e de situações sociais para o falante utilizar ou não as marcas de plural nos elementos do SN no PB. Ao evidenciar tal hipótese, ela apresenta os resultados obtidos em sua tese, que foram baseados apenas nos dados dos adultos.

Para melhor entendermos como Scherre (1988) desenvolveu suas análises, explanaremos brevemente como se deu o procedimento adotado pela autora em relação às variáveis linguísticas e sociais.

Dentre as variáveis linguísticas consideradas por Scherre estão: saliência fônica, tonicidade da sílaba, número de sílabas⁸, posição linear, classe gramatical do elemento no SN e marcas precedentes. Primeiramente, a autora testou o comportamento da saliência fônica e da tonicidade de forma isolada e verificou que os resultados dos pesos relativos não foram significativos. Daí Scherre realizou outra rodada com as variáveis saliência fônica e tonicidade aliadas, num mesmo grupo de fatores.

No entanto, Scherre (1988, p. 82), “Considerando que a diferenciação material fônica só ocorre basicamente nos substantivos, adjetivos e categorias substantivadas, uma vez que os demais itens lexicais são regulares”, efetuou essa análise somente com os substantivos, adjetivos e categorias substantivadas. Vejamos os resultados na tabela abaixo:

⁸ Na análise deste estudo não adotamos *número de sílabas* como variável, devido à sua irrelevância no estudo de Scherre (1988).

TABELA 1: Análise final de Processos e Tonicidade dos substantivos, adjetivos e categorias substantivadas dos adultos por escolarização

SALIÊNCIA FÔNICA		FALANTES			
		Todos	Primário	Ginasial	Colegial
Duplo (novo) (nóvos)	F.	$\frac{63}{68} = 93\%$	$\frac{22}{25} = 88\%$	$\frac{24}{26} = 92\%$	$\frac{17}{17} = 100\%$
	P.	0,80	0,81	0,82	+
-l (casal) (casais)	F.	$\frac{80}{93} = 86\%$	$\frac{21}{30} = 70\%$	$\frac{18}{19} = 95\%$	$\frac{41}{44} = 93\%$
	P.	0,69	0,64	0,85	0,68
-R (cor) (cores)	F.	$\frac{237}{268} = 88\%$	$\frac{77}{94} = 82\%$	$\frac{74}{83} = 89\%$	$\frac{85}{90} = 94\%$
	P.	0,65	0,67	0,64	0,71
-ão (dentão) (dentões)	F.	$\frac{173}{200} = 86\%$	$\frac{31}{43} = 72\%$	$\frac{61}{73} = 84\%$	$\frac{81}{84} = 96\%$
	P.	0,59	0,56	0,51	0,77
-S (país) (países)	F.	$\frac{219}{265} = 83\%$	$\frac{87}{113} = 78\%$	$\frac{68}{82} = 83\%$	$\frac{64}{70} = 91\%$
	P.	0,56	0,64	0,44	0,51
Regular oxítono e monossílabo tônico	F.	$\frac{134}{185} = 72\%$	$\frac{38}{67} = 57\%$	$\frac{51}{66} = 25\%$	$\frac{44}{51} = 86\%$
	P.	0,38	0,38	0,38	0,36
Regular proparoxítono	F.	$\frac{81}{156} = 52\%$	$\frac{15}{48} = 31\%$	$\frac{28}{60} = 47\%$	$\frac{38}{48} = 79\%$
	P.	0,21	0,16	0,19	0,32
Regular paroxítono	F.	$\frac{2604}{5242} = 50\%$	$\frac{600}{1875} = 32\%$	$\frac{913}{1750} = 52\%$	$\frac{1076}{1602} = 67\%$
	P.	0,17	0,19	0,16	0,17

Fonte: Scherre (1988, p. 139).

De acordo com os resultados, Scherre (1988) verificou que da análise com todos os dados, os itens mais salientes e favoráveis à marcação de plural são: o *plural duplo*, os *itens terminados em -l*, os *itens terminados -r*, os *itens terminados em -ão* e os *itens terminados em -s*.

Para Scherre (1988), a variável número de sílabas não se mostrou relevante, pois os três fatores, *monossílabo*, *dissílabo* e *mais de duas* sílabas apresentaram pesos relativos praticamente iguais e próximos do ponto neutro.

Scherre também realizou uma rodada com as variáveis posição linear e classe gramatical do elemento isoladas e chegou à conclusão de que ambas são mais relevantes se forem aliadas. A tabela a seguir apresenta os resultados:

TABELA 2: Classe gramatical e posição no SN: Análise C

CLASSE E POSIÇÃO	F A L A N T E S			
	Adultos		Crianças	
	Apl./Total	% Prob	Apl./Total	% Prob
Substantivo 1a pos.	153/161 = 95	(0,62)	16/18 = 89	(0,29)
Substantivo 2a pos.	2756/5175 = 53	(0,33)	419/1052 = 40	(0,28)
Substantivo 3a pos.	317/514 = 62	(0,20)	46/86 = 53	(0,28)
Cat. subst. 2a pos.	59/92 = 65	(0,43)	13/20 = 65	(0,45)
Cat. subst. 3a pos.	16/23 = 70	(0,28)	2/2 = 100	+
Pr. pessoal 1a pos.	12/12 = 100	+	Xxx	
Pr. pessoal 2a pos.	27/33 = 82	(0,80)	3/5 = 60	(0,59)
Adjetivo 1a pos.	47/48 = 98	(0,84)	4/4 = 100	+
Adjetivo 2a pos.	118/156 = 78	(0,48)	18/18 = 100	+
Adjetivo 3a pos.	125/307 = 31	(0,12)	16/40 = 40	(0,31)
Quant. 1a pos.	133/133 = 100	+	23/23 = 100	+
Quant. 2a pos.	3/7 = 43	(0,34)	xxx	
Quant. 3a pos.	22/108 = 20	(0,10)	1/5 = 20	(0,14)
Possessivo 1a pos.	184/184 = 100	+	42/43 = 98	(0,74)
Possessivo 2a pos.	55/60 = 92	(0,75)	12/13 = 92	(0,87)
Possessivo 3a pos.	2/12 = 17	(0,03)	2/4 = 50	(0,53)
Adjetivo 2 1a pos.	9/10 = 90	(0,46)	Xxx	
Adjetivo 2 2a pos.	21/28 = 78	(0,69)	1/1 = 100	+
Adjetivo 2 3a pos.	3/3 = 100	+	0/1 = 0	-
Indefinido 1a pos.	424/429 = 99	(0,92)	80/80 = 100	+
Indefinido 2a pos.	44/47 = 94	(0,92)	5/8 = 63	(0,64)
Indefinido 3a pos.	0/1 = 0	-	Xxx	
Art. e dem. 1a pos.	3268/3363 = 97	(0,83)	672/688 = 98	(0,83)
Art. e dem. 2a pos.	75/75 = 100	+	14/14 = 100	+

Fonte: Scherre (1988, p. 214).

De acordo com Scherre (1988), considerando a posição, podemos observar que todas as classes gramaticais quando ocupam a primeira posição, favorecem a CN. Na segunda posição, os *artigos* e *demonstrativos*, *indefinidos*, *adjetivos 2* e *possessivos* quando ocupam a segunda posição favorecem a CN.

Já os *adjetivos* tiveram queda nos pesos relativos nos dados dos adultos e totalmente marcados nos dados das crianças. A categoria substantivada também apresentou um baixo valor para a marcação de plural e o substantivo foi a classe gramatical menos marcada. Os pronomes pessoais se mostraram equivalentes aos determinantes.

Na terceira posição, todas as classes gramaticais apresentaram um número baixo para a aplicação de CN.

As variáveis sociais consideradas por Scherre (1988) foram: sexo, faixa etária e escolaridade. Os resultados obtidos mostraram que:

1. Os falantes do sexo feminino apresentaram um resultado mais favorável para o uso de CN que os falantes do sexo masculino;
2. Para a faixa etária, os resultados apresentaram uma distribuição curvilínea não acentuada, isto é, das três faixas etárias dos adultos (15-25, 26-49 e 50-71), os pesos relativos foram (0,49, 0,54 e 0,47, respectivamente). A faixa etária que mais apresentou concordância foi a de 26 a 49 anos;
3. Os anos de escolarização são proporcionais ao uso de CN, posto que, quanto maior o tempo de escolarização, maior será o aumento do uso da regra.

Essas variáveis trabalhadas por Scherre (1988) foram retomadas para análise no presente estudo, com a comunidade de Guarapuava, a fim de se estabelecer comparação entre os resultados acerca do tema, a CN. Referente à variável marcas precedentes, discorreremos no próximo estudo.

Scherre (1998)

Scherre (1998), em seu estudo “Sobre a influência de três variáveis relacionadas na concordância nominal em português”, ao discordar da correlação feita por Guy (1981), que os efeitos da classe gramatical e posição do elemento no SN são equivalentes, desenvolve um novo método de análise, no qual inclui três variáveis aliadas: posição linear do elemento no SN, classe gramatical e marcas precedentes. Sua amostra contava com 11.086 dados retirados de 48 entrevistas, do *Corpus Censo*, do Rio de Janeiro.

Após verificar a inter-relação entre as três variáveis, a autora as transforma em apenas duas, são elas: 1) relação entre elementos nucleares e não nucleares e posição linear dos elementos nucleares do SN e (2) marcas precedentes em função da posição.

Para explicar o efeito da classe gramatical, proposta por (Guy, 1981), em função da posição, Scherre (1998) apresenta uma tabela com os dados distribuídos que justificam a sua discordância com Guy (1981).

TABELA 3: Distribuição dos dados dos falantes adultos em função da classe gramatical e da posição dos elementos no SN

CLASSE	POSIÇÃO			TOTAL
	Primeira	Segunda	Terceira	
Substantivo	153/161= 95% Ex.: <u>problemas</u> assim maiores	2.777/5.196=53% Ex.: essas <u>carne</u> congelada	317/514= 62% Ex.: das <u>otas</u> famílias	3247/5871
Categoria substantivada	Não ocorre	60/93= 65% Ex.: trabalha uma, as <u>outras</u> fica	16/23= 70% Ex.: as coisas <u>boas</u> e as ruins	76/116
Pronome Pessoal 3ª pessoa	12/12= 100% Ex.: <u>delas</u> todas	27/33= 82% Ex.: de todos <u>eles</u>	Xxxx Ex.: todos <u>ele</u>	39/45
Adjetivo	47/48= 98% Ex.: <u>novas</u> escolas	125/163= 77% Ex.: as <u>boas</u> ações	125/307= 41% Ex.: uma coisinha <u>bonitinha</u>	297/518
Quantificador	133/133= 100% Ex.: <u>todos</u> os anos	3/7= 43% Ex.: eles <u>todo</u>	22/108= 20% Ex.: aquelas cruzinha <u>toda</u>	158/248
Possessivo	184/184= 100% Ex.: <u>suas</u> tias	130/135= 96% Ex.: nos <u>meus</u> filhos	2/12= 17% Ex.: uns colega <u>meu</u>	316/331
Demais classes não nucleares	3701/3802= 97%	141/151= 93%	3/4= 75%	3845/3953
Total	4230/4340= 97%	3263/5778= 56%	485/968= 50%	7978/11086

Fonte: Scherre (1998, p. 95).

Os resultados de Scherre (1998) apontaram que os SN's que ocupam a terceira posição não são somente adjetivos, mas também substantivos. Além disso, 77% dos adjetivos ocupam a segunda posição, o que revela um percentual bem significativo para ser desconsiderado, ainda que seja um número menor (163) comparado ao número de substantivos na mesma posição (5.196).

Os substantivos são marcados com maior frequência na terceira posição (62%), ao passo que os adjetivos são mais frequentes na segunda posição (77%). Quanto aos possessivos e os demais elementos não nucleares, também aparecem fortemente favoráveis na segunda posição (96% e 93%, respectivamente). O grande destaque é para qualquer classe gramatical que ocupe a primeira posição do SN por ser a mais marcada.

Com base nos resultados, a autora atestou que “tomar classe por posição ou posição por classe encobre regularidades lingüísticas importantes” (SCHERRE, 1998, p. 97). Por consequência, Scherre inclui em sua análise um novo método analítico que considere as duas variáveis relacionadas, posto que:

1. As classes gramaticais não nucleares na primeira posição são todas antepostas ao núcleo do SN;
2. As classes gramaticais da segunda posição são necessariamente antepostas ao núcleo do SN;
3. As classes não nucleares da terceira posição são necessariamente pospostas ao núcleo do SN.

Para exemplificar, segue abaixo a tabela apresentada por Scherre (1998) com as classes não nucleares distribuídas em função da posição linear no SN, e em função da relação de anteposição ou posposição ao núcleo do SN.

TABELA 4: Distribuição das classes gramaticais não nucleares em função da posição e da relação com o núcleo

CLASSE E POSIÇÃO	Anteposto	Posposto
Adjetivo na 1ª posição (<u>novas escolas</u>)	47/48= 98%	Não ocorre
Adjetivo na 2ª posição (as <u>boas</u> ações/coisas <u>lindas</u>)	36/39= 92%	89/124= 72%
Adjetivo na 3ª posição (uma casinha <u>bonitinha</u>)	Não ocorre	127/307= 40%
Quantificador na 1ª posição (<u>todos</u> os anos)	133/133= 100%	Não ocorre
Quantificador na 2ª posição (eles <u>todo</u>)	Não ocorre	3/7= 43%
Quantificador na 3ª posição (aquelas cruzinha <u>toda</u>)	Não ocorre	22/108= 20%
Possessivo na 1ª posição (<u>suas</u> tias)	184/184= 100%	Não ocorre
Possessivo na 2ª posição (os <u>meus</u> filhos/colegas <u>minhas</u>)	127/130= 98%	3/7= 43%
Possessivo na 3ª posição (uns colegas <u>meu</u>)	Não ocorre	2/12= 17%
Demais elementos não-nucl. 1ª pos. (<u>alguns</u> meses/ <u>determinados</u> pontos)	3791/3802= 97%	Não ocorre
Demais elementos não-nucl. 2ª pos. (<u>todos</u> os anos/eles <u>mesmo</u>)	141/148= 95%	0/3= 0%
Demais elementos não-nucl. 3ª pos. (do meus <u>próprios</u> filhos)	1/1= 100%	2/3= 66%

Fonte: Scherre (1998, p. 99).

Conforme os resultados referentes às classes não nucleares, Scherre (1998) obteve as seguintes conclusões: todos os elementos antepostos ao núcleo são mais marcados e todos os elementos pospostos ao núcleo são menos marcados. A relação entre classe gramatical e posição, referente aos elementos não nucleares, deve ser priorizada pela distribuição dos elementos em torno do núcleo, sem dar muita importância nem à classe nem à posição, mas à distribuição da classe não nuclear em relação ao centro do SN.

Ainda segundo Scherre (1998), há uma segunda questão importante: a posição linear mostra-se importante apenas com relação aos elementos não nucleares do SN, sendo pertinente tal observação tendo em vista que: as categorias substantivadas assemelham-se aos substantivos, quando apresentam maior marcação de plural na terceira posição; os pronomes pessoais apresentam um maior número de marcas na primeira posição.

Desse modo, a autora propõe aliar os dois fatores em questão numa única variável denominada relação entre os elementos do SN e destaca:

1. As classes não nucleares antepostas ao núcleo são mais marcadas do que as pospostas a ele, e destaca que não é o adjetivo que é pouco marcado, mas sim o adjetivo posposto; não é o quantificador que é muito marcado, mas sim o quantificador anteposto, o que justifica que nem sempre a primeira posição do SN é a mais marcada;
2. Os elementos não nucleares não são marcados da mesma forma em todas as posições, já que a primeira e a terceira posição são geralmente as mais marcadas do que a segunda posição.

Segundo Scherre (1998), as variáveis marcas precedentes e posição linear não foram trabalhadas adequadamente no estudo do Espanhol (Poplack, 1980) e para mostrar as semelhanças entre o Espanhol e o PB, Scherre retoma a variável marcas precedentes, selecionando nove fatores para análise, que seguem abaixo com os respectivos exemplos:

1. ausência – refere-se à ausência de qualquer constituinte do SN precedente ao elemento da primeira posição (OS fregueses, DO meus tio);
2. zero formal na primeira posição – refere-se ao elemento da primeira posição não marcado (DO meus tios, FILHOTINHO novos);
3. numerais da primeira posição (TRÊS capítulos, NOVE núcleos formados);
4. presença de marca formal na primeira posição – (OS fregueses, UNS troço);

5. núcleo semântico e presença de marca formal – refere-se aos casos: a) do núcleo nominal mais alto com marca de plural (MILHÕES de coisas) e b) do núcleo nominal com ou sem marca de plural, acompanhado de marca de plural no primeiro elemento (UM GRUPO DE CRIANÇAS abandonadas).

6. núcleo semântico e ausência de marca formal – referente aos casos: a) do núcleo nominal mais alto sem marca de plural (UMA PORÇÃO DE carro, UM GRUPO DE crianças abandonadas) e b) do núcleo nominal sem marca de plural acompanhado de ausência de marca de plural no primeiro elemento analisado. (UMA PORÇÃO DE COISA interessante).

7. presença de marcas formais a partir da primeira posição – quando há duas ou mais marcas de plural precedendo o elemento analisado na terceira, quarta ou quinta posição. (AS MAIORES privações, AS PARTIDAS TODAS iguais).

8. mistura de marcas – refere-se aos casos de presença de pelo menos uma marca de plural no elemento que precede o elemento analisado, não mediada por zero em elemento que contém marca (do MEUS tios).

9. zero formal a partir da primeira posição – entre a última marca e o elemento em questão deve ter zero em elemento que admite marca, mesmo mediado por um numeral ou quantificador (UMAS BORRACHA grande, DOIS RISCO verde).

A partir dos resultados obtidos por Scherre (1998), salvo a exceção da ausência fora do SN e zero na primeira posição, que, logicamente, favorecem a aplicação de marca de plural, a autora concluiu que os outros fatores corroboram a afirmação de que “marcas levam a marcas e zeros levam a zeros”.

Enfim, sua principal conclusão é que a variável posição, isolada das variáveis marcas precedentes e classe gramatical, não explicam e não descrevem adequadamente o fenômeno em estudo na sua totalidade, por isso, julga necessário analisar as três variáveis, posição, classe gramatical e marcas precedentes relacionadas e considerar seus fatores de forma cruzada.

Fernandes (1996)

Um dos estudos que trataram do tema sobre a CN na região sul, foi o de Fernandes (1996), intitulado “Concordância nominal na região sul”. O estudo da pesquisadora foi realizado com uma amostra composta por 47 informantes de Florianópolis, Chapecó, Irati e Panambi considerando a fala de situação informal e 19 informantes de diversas regiões da fala de situação formal, retirada do Projeto VARSUL (Variação Linguística Urbana da Região Sul), que teve início em 1989, baseado no projeto Censo.

As variáveis linguísticas que Fernandes (1996) considerou foram: posição linear dos elementos no SN; classe gramatical dos elementos; relação com o núcleo do SN; marcas precedentes; processos morfofonológicos de formação do plural; tonicidade dos itens; contexto seguinte; grau dos substantivos e adjetivos; e animacidade dos substantivos e adjetivos. As variáveis sociais foram: idade; níveis de escolarização; sexo; etnia; e níveis de formalidade.

De acordo com as postulações de Scherre (1988), Fernandes seguiu a mesma forma de análise das variáveis. A pesquisadora realizou uma rodada com as variáveis saliência fônica e tonicidade isoladas, e, verificando que as mesmas não se mostraram significativas, efetuou uma nova rodada com as variáveis aliadas, num mesmo conjunto de fatores. A tabela seguinte traz os resultados da variável saliência fônica e tonicidade aliadas, somente dos substantivos, adjetivos e categorias substantivadas:

TABELA 5: Processos e tonicidade dos substantivos, adjetivos e categorias substantivadas

Saliência Fônica	Apl./Total	%	PRel.
Duplo	25/35	81%	0,81
Terminado em -L	57/77	77%	0,77
Terminado em ãO (OE)	79/103	74%	0,74
Terminado em -R	85/116	73%	0,73
Terminado em -S	52/77	68%	0,58
Regular oxítono	108/192	56%	0,54
Regular paroxítono	1263/2487	51%	0,46
Regular proparoxítono	93/152	61%	0,45
Total	1762/3234	54%	

Fonte: Fernandes (1996, p. 88).

Pelos resultados de Fernandes, podemos observar que os itens mais favoráveis à CN foram: o *plural duplo*, os *itens terminados em -l*, os *itens terminados em ão (oe)* e os *itens terminados em -r*, de forma semelhante à análise de Scherre (1988).

Fernandes também realizou uma análise com a posição linear e classe gramatical de forma isolada e constatou que os resultados foram irrelevantes, assim como os de Scherre e, então, Fernandes decidiu aliar as variáveis posição linear e classe gramatical numa única variável. Seguem abaixo os resultados:

TABELA 6: Distribuição das classes gramaticais do SN em função da posição e da relação com o núcleo

Fatores	Apl./Total	%	PRel
Classe não nuclear 1' posição	1929/1981	97%	0,85
Classe não nuclear anteposta T posição	181/203	89%	0,72
Classe nuclear 1* posição	113/120	94%	0,67
Classe nuclear 2@ posição	1320/2554	52%	0,23
Classe nuclear nas demais posições	140/268	52%	0,22
Classe não nuclear posposta T posição	75/103	73%	0,32
Classe não nuclear posposta demais pos.	70/146	48%	0,25
Total	3829/5375	71%	

Fonte: Fernandes (1996, p. 57).

Em suma, as principais conclusões obtidas por Fernandes (1996) acerca dos resultados da pesquisa foram:

1. A presença de CN é fortemente influenciada pela posição do elemento do SN em relação ao núcleo, e os elementos que recebem a marcação de plural geralmente estão posicionados antes do núcleo;
2. O substantivo pode estar em todas as posições do SN e o que irá determinar a marcação de plural vai depender da sua relação com o núcleo. O mesmo vale para os adjetivos e determinantes.
3. Sobre a saliência fônica, a autora também atestou a mesma conclusão obtida por outros estudiosos, ou seja, que as formas mais salientes são mais perceptíveis, e, por isso, são mais marcadas. Também concluiu que a análise com a saliência fônica e tonicidade num mesmo conjunto de fatores é mais significativa.

4. A escolaridade é o fator que mais influencia a aplicação de concordância e quanto maior o nível de escolaridade o falante possuir, maior será o uso da regra.

5. As mulheres mais jovens tendem a utilizar mais a CN que as mulheres mais velhas. No geral, as mulheres aplicam a CN com maior frequência que os homens.

A partir do que foi arrolado até aqui, iremos nos embasar para desenvolvermos nosso trabalho. No próximo capítulo, serão apresentados os procedimentos metodológicos desta pesquisa.

CAPÍTULO 3 – PROCEDIMENTOS METODOLÓGICOS

O presente estudo adotou os procedimentos teórico-metodológicos da Sociolinguística variacionista, área proposta por William Labov. Neste capítulo, discorremos sobre a comunidade de Guarapuava, Paraná. Na sequência, apresentamos as etapas da pesquisa realizada: a constituição da amostra, a coleta de dados e o programa estatístico GoldVarbX.

Para finalizar, apresentamos as variáveis linguísticas - dependentes e independentes - e as variáveis sociais ou extralinguísticas analisadas nesta pesquisa.

3.1. A comunidade de Guarapuava/Paraná

Segundo Izidoro (1976) e Marcondes (2010), a cidade de Guarapuava está situada na região sul do Brasil, no Terceiro planalto, chamado Planalto de Guarapuava, precisamente, no centro-oeste do estado do Paraná, no trajeto entre as cidades de Curitiba e Foz do Iguaçu, nas margens da BR-277.

Guarapuava está localizada a 246 km de Curitiba (capital do Paraná), 362 km do Porto de Paranaguá, 405 km de Foz do Iguaçu. Guarapuava conta com quatro distritos rurais: Palmeirinha, Guará, Guairacá e Entre-Rios. De acordo com o CENSO de 2015, o município de Guarapuava apresenta 170.000 habitantes⁹.

Historicamente, antes da descoberta de Guarapuava pelos portugueses, a cidade era ocupada pela povoação indígena tupi-guaraní, uma vasta área silvestre, onde havia muitos animais, entre eles, lobos guará. Daí a origem do nome de Guarapuava, que significa: GUARA- lobo, cão selvagem; PUAVA – bravo, arisco; GUARAPUAVA – lobo bravo.

Conforme Marcondes (2010), pouco antes da descoberta do Brasil no final do século XIV, houve um acordo entre Portugal e Espanha pelo Tratado de Tordesilhas, no qual ficou determinado a partir da linha imaginária a 370 léguas de Cabo Verde, que as terras descobertas situadas ao Leste fossem de Portugal e as do Oeste pertenceriam à Espanha. Geograficamente, o território de Guarapuava pertencia ao domínio espanhol.

Após quatro décadas do Tratado de Tordesilhas, a expedição espanhola nomeia governador Dom Alvár Nuñez Cabeza de Vaca, para tomar posse e ocupar os campos do Paraná e de Santa Catarina, incluindo Guarapuava, que em 1541, denomina essas terras de Província de Vera.

⁹ Disponível em: <<http://cidades.ibge.gov.br/xtras/perfil.php?codmun=410940>>. Acesso em: 12/04/2016.

Com as invasões, destruições e extermínios dos índios pelos portugueses, vindos de São Paulo, a cidade permaneceu praticamente abandonada. Aos poucos, outros indígenas de diversas origens vieram habitá-la. Após um século e meio, houve um acordo entre Espanha e Portugal pelo Tratado de Madri, que delimitou as fronteiras ocupadas pelos portugueses.

De acordo com Marcondes (2010), no início do século XIX, houve outra tentativa de conquista do território de Guarapuava pela coroa portuguesa, numa expedição comandada por Diogo Pinto Azevedo Portugal e dois missionários, o Padre Francisco das Chagas Lima e Frei Nolasco, para catequizar os índios. Essa expedição chegou a Guarapuava em 1810 e fundou a povoação de Atalaia.

Os pioneiros para cá trouxeram as suas famílias, os seus recursos, seus escravos, seu gado; transplantaram a cultura portuguesa com sua língua, seus costumes, seu folclore, a religião católica, a arquitetura da Vila Portuguesa com casarões geminados no estilo barroco colonial. (MARCONDES, 2010, p. 129).

Em 1819, agravou-se o estado de saúde de Azevedo Portugal, que por sua vez, foi afastado e substituído por Rocha Loures. O município de Guarapuava foi fundado em 09 de dezembro de 1819, povoado pela Freguesia de Nossa Senhora de Belém, por influência do Padre Chagas Lima, primeiro Vigário da cidade. Em 1852, a Freguesia foi elevada à Vila, em 1859, foi criada a Comarca de Guarapuava, mas somente em 1871, passou à categoria de Cidade de Guarapuava.

No ano de 1841, Padre Chagas Lima auxiliou na fundação da Igreja Matriz de Nossa Senhora de Belém, atualmente Catedral. A santa padroeira da cidade de Guarapuava é Nossa Senhora do Belém. Apresentamos abaixo a imagem da igreja Catedral:

FIGURA 1: Catedral Nossa Senhora de Belém – Guarapuava/Paraná.



Fonte: www.google.com.br/imagens¹⁰.

De acordo com Izidoro (1976) e Marcondes (2010), em 1896, ocorreu a fundação da Colônia de Prudentópolis, que mais tarde, passou a município, formada por camponeses poloneses e ucranianos. No entanto, no início dos anos 50, imigrantes alemães conhecidos como os Suábios do Danúbio, chegaram ao município.

No total, quinhentas famílias instalaram-se em cinco vilas: Vitória, Jordãozinho, Cachoeira, Samambaia e Socorro, e fundaram a Colônia de Entre Rios, assim nomeada porque se situava entre os rios Jordão e Pinhão. A população alemã trabalha na agricultura, principalmente com o trigo, o milho e a soja, amparada pela Cooperativa Agrária Mista Entre Rios Ltda.

Enfim, com o passar do tempo, imigrantes de diversas etnias fixaram-se em Guarapuava, dentre eles índios, portugueses, espanhóis, africanos, franceses, ucranianos, poloneses, alemães, austríacos, italianos, árabes, sírio-libaneses e japoneses, que contribuíram à formação da cultura de Guarapuava.

Apresentamos abaixo o mapa com a localização de Guarapuava.

¹⁰ Disponível em:

<https://www.google.com.br/?gfe_rd=cr&ei=TDdUV8j9M8aU8Qff1a2gBA&gws_rd=ssl#q=imagens+de+guarapuava>. Acesso em: 05/06/2016.

FIGURA 2: Mapa de Guarapuava/Paraná e região

Fonte: Centro Estadual de Educação Profissional Arlindo Ribeiro – CEEPAR¹¹

Na próxima seção, discutiremos sobre a constituição da amostra de Guarapuava, com a distribuição dos informantes e apresentaremos um quadro com as características sociais dos mesmos.

3.2. A constituição da amostra

O *corpus* de análise do nosso estudo é constituído por uma amostra do banco VARLINGUA, criado por pesquisadoras da Universidade Estadual do Centro-Oeste – Unicentro, Loremi Loregian-Penkal e Lucelene Teresinha Franceschini, do Programa de Pós-

¹¹ Disponível em:

<<http://www.grpceeparlindoribeiro.seed.pr.gov.br/modules/conteudo/conteudo.php?conteudo=11>>. Acesso em: 05/02/2016.

Graduação em Letras. O banco de dados apoia-se teoricamente na Sociolinguística Variacionista e assemelha-se ao do Projeto VARSUL.

O banco VARLINGUA em seu estágio atual contém trinta e duas 32 entrevistas, distribuídas por sexo (*feminino* e *masculino*), faixa etária (*25 a 45 anos* e *50 anos ou mais*) e escolaridade (*Fundamental I*, *Fundamental II*, *Ensino Médio* e *Ensino Superior*). Dessa amostra, vinte e quatro 24 entrevistas serão analisadas nesta pesquisa.

Os vinte e quatro 24 informantes foram assim distribuídos: temos 12 com idades de *25 a 45 anos* e 12 com idades de *50 anos ou mais*, sendo 12 do sexo *feminino* e 12 do sexo *masculino* de cada faixa etária. A amostra contém 8 informantes com *Fundamental I*, 8 informantes com *Fundamental II* e 8 informantes com *Ensino Médio*.

Para a seleção dos informantes, foram estabelecidos alguns critérios, são eles:

1. Ser natural de Guarapuava (da zona urbana ou rural);
2. Não ter residido em outro município;
3. Ter envolvimento com a comunidade local.

No quadro abaixo, apresentamos os dados sociais de 24 informantes pertencentes ao banco VARLINGUA e incluídos na análise desta pesquisa.

QUADRO 1: Estratos sociais dos falantes de Guarapuava

IDENTIFICAÇÃO	SEXO	FAIXA ETÁRIA	ESCOLARIDADE
Inf1 F1p	F	I	Fundamental I
Inf2 M1p	M	I	Fundamental I
Inf3 F1g	F	I	Fundamental II
Inf4 M1g	M	I	Fundamental II
Inf5 F1c	F	I	Ensino Médio
Inf6 M1c	M	I	Ensino Médio
Inf7 F2p	F	II	Fundamental I
Inf8 M2p	M	II	Fundamental I
Inf9 F2g	F	II	Fundamental II
Inf10 M2g	M	II	Fundamental II
Inf11 F2c	F	II	Ensino Médio
Inf12 M2c	M	II	Ensino Médio
Inf13 F1p	F	I	Fundamental I
Inf14 M1p	M	I	Fundamental I
Inf15 F1g	F	I	Fundamental II
Inf16 M1g	M	I	Fundamental II
Inf17 F1c	F	I	Ensino Médio
Inf18 M1c	M	I	Ensino Médio
Inf19 F2p	F	II	Fundamental I
Inf20 M2p	M	II	Fundamental I
Inf21 F2g	F	II	Fundamental II
Inf22 M2g	M	II	Fundamental II
Inf23 F2c	F	II	Ensino Médio
Inf24 M2c	M	II	Ensino Médio

Fonte: Dados do VARLINGUA 2015

No subitem 3.3, a seguir, apresentamos como se deu o processo da coleta de dados.

3.3. A coleta de dados

A coleta de dados do banco VARLINGUA ocorreu entre os anos de 2014 e 2015, na comunidade de Guarapuava. O procedimento¹² de coleta foi feito por meio de entrevistas sociolinguísticas gravadas, no mínimo com quarenta minutos de duração cada. Como especificado no item anterior, a escolha dos informantes não foi aleatória, pois eles foram selecionados e distribuídos por sua estratificação social.

A finalidade da pesquisa sociolinguística é atingir o vernáculo¹³ do falante, para identificar quais fatores são mais condicionadores ou inibidores da variação em estudo. Na amostra do VARLINGUA, os informantes foram selecionados por sexo (*feminino* e *masculino*), faixa etária (*25 a 45 anos* e *50 anos ou mais*) e escolaridade (*Fundamental I*, *Fundamental II* e *Ensino Médio*).

Para que o vernáculo do falante seja alcançado, conforme propõe Labov ([1972] 2008), durante a entrevista, o informante expõe sua experiência pessoal ao pesquisador, que pergunta sobre temas ou situações vividas como: infância, relacionamentos, família, filhos, casamento, perigos de morte, sonhos, entre outros. Dessa forma, o informante tende a desviar o foco da forma como fala e age com naturalidade.

Segundo Labov ([1972] 2008), o pesquisador ao monitorar a fala do informante assume a posição de observador, e, com isso, ele encontra dificuldades na coleta de dados, pois a sua participação na gravação pode interferir na naturalidade da conversa, o que o autor denomina de paradoxo do observador:

O objetivo da pesquisa linguística na comunidade deve ser descobrir como as pessoas falam quando não estão sendo sistematicamente observadas – no entanto, só podemos obter tais dados por meio da observação sistemática. (LABOV, [1972] 2008, p. 244).

Logo, cabe ao pesquisador saber envolver o informante para conseguir o seu objetivo, o vernáculo. Vale ressaltar que o informante assina o Termo de Consentimento Livre e Esclarecido (TCLE) e a ficha social¹⁴ antes da gravação. Para esta pesquisa, analisamos vinte e quatro (24) entrevistas do banco VARLINGUA.

Após a gravação das entrevistas, estas foram transcritas. Esta etapa visa anotar os dados a serem analisados, conforme o seu objeto de estudo, neste caso, anotamos todas as ocorrências que tinham SN's, com a *presença de CN* ou a *ausência de CN*. Todas as

¹² O processo de coleta dos dados do VARLINGUA foi realizado com a participação das autoras deste estudo e também com a participação de Vanessa Deon.

¹³ Segundo Labov ([1972] 2008, p. 244), vernáculo é “o estilo em que se presta o mínimo de atenção ao monitoramento da fala”, ou seja, a fala espontânea, numa situação informal do uso da língua.

¹⁴ A ficha social e o modelo de TCLE encontram-se em anexo nas páginas 95 e 96.

entrevistas foram transcritas e codificadas manualmente, conforme os códigos que demos à variável CN e às variantes¹⁵, para inseri-las no programa estatístico.

Por fim, para a obtenção dos resultados, a codificação dos dados foi submetida à análise quantitativa do programa estatístico GoldVarbX, que será abordado na próxima seção.

3.4. O programa estatístico GoldVarbX

De acordo com a sociolinguística, a variação linguística não é aleatória, mas condicionada por fatores linguísticos e sociais, que favorecem ou desfavorecem o uso de uma variante, que, por sua vez, se correlaciona à variável posta em análise, neste caso, a CN. Assim, a partir dos dados obtidos será feita uma análise quantitativa dos dados pelo programa estatístico escolhido.

A análise quantitativa “possibilita o estudo da variação linguística, permitindo ao pesquisador apreender sua sistematicidade, seu encaixamento linguístico e social e sua eventual relação com a mudança linguística.” (GUY e ZILLES, 2007, p. 73).

O programa estatístico irá executar os cálculos e fornecer os resultados. Nesse cálculo, são comparadas as porcentagens com que os dados se distribuem pelos diversos fatores. Porém, são os pesos relativos que irão precisar tal distribuição. Os pesos relativos são valores que vão de zero a um e que indicam matematicamente o peso com que um fator (linguístico ou extralinguístico) influencia o uso de uma variante em relação a todos os outros fatores considerados a partir de um fenômeno linguístico.

Segundo Mollica (2013), o peso relativo cujo valor é superior a 0.50, indica que o fator favorece o uso da variante. Se o valor calculado for igual a 0.50, é neutro, ou seja, não favorece nem desfavorece a variante em análise, embora um valor aproximado do ponto neutro possa ser considerado como indicador, mesmo sem valor significativo. E quando o valor do peso relativo for inferior a 0.50, indica que maior será a probabilidade de o fator desfavorecer o uso da variante em análise. Contudo, leva-se em consideração também a diferença entre os números.

No entanto, é importante salientar, que a ferramenta estatística fornece os resultados numéricos e nos auxilia a entender melhor o funcionamento de um fenômeno linguístico, mas os dados estatísticos serão interpretados pelo linguista que irá comprovar ou refutar as hipóteses levantadas.

¹⁵ O documento utilizado na escolha e na codificação das variantes encontra-se em anexo na página 95.

Nesta pesquisa, para o tratamento dos dados, utilizamos o programa GoldVarbX, de David Sankoff, Sali Tagliamonte e Eric Smith (2005), do pacote de programas VARBRUL¹⁶, uma versão projetada para ambiente Windows. Nas palavras de Guy e Zilles (2007, p. 105):

O Varbrul é um conjunto de programas computacionais de análise multivariada, especificamente estruturado para acomodar dados de variação sociolingüística. A análise se chama ‘multivariada’ porque permite investigar situações em que a variável lingüística em estudo é influenciada por vários elementos do contexto, ou seja, múltiplas variáveis independentes.

Portanto, o pesquisador irá interpretar os resultados obtidos pelo programa estatístico para confirmar ou não suas hipóteses. No próximo subitem, são apresentadas as variáveis independentes lingüísticas e sociais analisadas, neste estudo, para que possamos verificar os efeitos que elas causam ou não sobre a variável dependente, a CN.

3.5. As variáveis e os grupos de fatores

Nesta seção, apresentamos as variáveis lingüísticas e sociais adotadas para esta pesquisa. As variáveis lingüísticas estão divididas em: variável dependente CN; e as variáveis independentes que são: saliência fônica, tonicidade, posição do elemento no SN, classe gramatical do elemento e marcas precedentes. As variáveis sociais são: sexo, faixa etária e escolaridade.

3.5.1. A variável dependente CN:

Denomina-se variável dependente a variável que posta em análise, pode conter mais de uma forma, isto é, pode apresentar a presença do fenômeno ou a ausência do mesmo. De acordo com Guy e Zilles (2007, p. 141):

A variável dependente é tratada em termos das probabilidades e percentuais de acontecimento de determinada alternativa, oposta à ausência dessa alternativa. Tal ausência pode incluir apenas uma alternativa, ou várias, mas, no caso de várias, uma análise binária trata todas elas em um só conjunto.

Neste estudo, a variável dependente CN apresenta duas variantes:

¹⁶ O VARBRUL foi criado em 1971, por Sankoff & Rousseau.

0. Ausência de CN.

(3) Minha infância foi boa né, bem melhor do que a infância de hoje, *das criança*. (VARLINGUA Inf3 F1g).

(4) Hoje im dia *as criança* quérим ficá só no computador. (VARLINGUA Inf15 F1g).

1. Presença de CN.

(5) *Os meus pais* erum pobres, eu não tinha nada de tecnologia. (VARLINGUA Inf5 F1c).

(6) O que eu aprendi com *meus pais* eu acho qui eu continuo até agora né. (VARLINGUA Inf9 F2g).

3.5.2. As variáveis independentes:

No tocante às variáveis independentes, segundo Guy e Zilles (2007, p. 137), elas são “tratadas no Varbrul como ‘grupos de fatores’, formam a parte central do sistema analítico. Cada uma delas representa uma hipótese de possíveis efeitos sobre a variável dependente (a ‘aplicação da regra’)

Assim, neste estudo, poderemos verificar se as variáveis independentes são relevantes ou não para a aplicação da regra de CN em Guarapuava. As variáveis consideradas para a análise são: saliência fônica, tonicidade, posição do elemento no SN, classe gramatical do elemento no SN e marcas precedentes.

3.5.2.1. Saliência fônica

A variável saliência fônica refere-se aos processos morfofonológicos de formação de plural, que representam os níveis de diferenciação do material fônico na forma singular/plural, apresentando, respectivamente, o maior e o menor grau de saliência fônica. Esta variável divide-se em seis fatores (ou níveis) que seguem na tabela abaixo¹⁷.

¹⁷ Scherre (1988) não utilizou tabela para apresentar os níveis da saliência fônica, mas inserimos uma para melhor visualização da variável e de suas variantes. O mesmo ocorre com a tabela 7, na sequência.

TABELA 7: Graus da variável saliência fônica

1. vocábulos com plural duplo ¹⁸	novo/novos
2. vocábulos terminados em -l	casal/casais
3. vocábulos terminados em -ão	seleção/seleções
4. vocábulos terminados em -r	cor/cores
5. vocábulos terminados em -s	freguês/fregueses
6. plural regular	lei/leis

Fonte: SCHERRE (1988, p. 75-77).

Para Scherre (1988), os resultados apresentados para a *saliência fônica* mostraram que a probabilidade de ocorrer mais marcas de CN está nos itens lexicais com maior diferença fônica entre a forma singular/plural, como no *plural duplo*, portanto, mais saliente.

Em nosso estudo, o grupo de fatores considerado na análise da *saliência fônica* divide-se em seis fatores (níveis de formação de plural), são eles:

1. *Plural duplo*. Ex: avô/avós.

(7) No tempo *dos avós*, eles não se conheciam, eram os pais que arrumavam o casamento né. (VARLINGUA Inf4 M1g).

2. *Itens terminados em -l*. Ex: animal/animais.

(8) Vamos sair por aí conhecê *os animais*, conhecê as criações. (VARLINGUA Inf10 M2g).

3. *Itens terminados em -r*. Ex: lugar/lugares.

(9) Eu acho que o comportamento do guarapuavano é diferente *dos outros lugares*. (VARLINGUA Inf3 F1g).

4. *Itens terminados em -ão*. Ex: oração/orações.

(10) Eu sozinha né, a gente pega, faz *as orações* da gente í tal. (VARLINGUA Inf9 F2g).

5. *Itens terminados em -s*. Ex: mês/meses.

¹⁸ Entendemos por *plural duplo* (ou plural metafônico), o termo que na passagem de singular para plural tenha inserção de -s e também sofra alteração de abertura vocálica, como em (ovo/óvus), segundo Scherre (1988).

(11) Eu fiquei *três mesis* só no convento. (VARLINGUA Inf3 F1g).

6. Plural regular. Ex: igreja/igrejas.

(12) Naquele tempo eu trabalhava muito *nas igreja*. (VARLINGUA Inf8 M2p).

Considerando o que diz Scherre (1988), a nossa hipótese para a variável saliência fônica é que a incidência de ocorrer o favorecimento da variante *presença de CN* de número esteja nos elementos lexicais mais salientes (como o *plural duplo* e os *itens terminados em -l* e os *itens terminados em -ão*).

3.5.2.2. Tonicidade

A variável tonicidade refere-se à ênfase dada a determinadas sílabas, e, por isso, tem por objetivo analisar os elementos que condicionam a *presença de CN* de acordo com a sílaba tônica. Scherre (1988) analisou a variável tonicidade da sílaba e verificou que os elementos oxítonos singulares ou monossílabos tônicos favorecem a marca de plural, posto que os paroxítonos e os proparoxítonos desfavorecem a aplicação da regra.

TABELA 8: Constituição da variável tonicidade

palavras oxítonas e monossílabos tônicos	professor/cor
palavras paroxítonas e monossílabos átonos	tribo/o/um
palavras proparoxítonas	Indígena

Fonte: SCHERRE (1988, p. 77).

Em suma, os resultados apresentados por Scherre (1988) em relação à variável saliência fônica e seus desdobramentos, tais como os processos morfofonológicos de formação de plural e tonicidade da sílaba, revelam que ambos são fatores relevantes na influência da concordância de número no SN.

Nesta seção, para a variável *tonicidade* do item singular, iremos considerar quatro fatores:

1. *Monossílabos átonos*.

(13) Você tem que agendá *us horário*. (VARLINGUA Inf11 F2c).

2. *Palavras oxítonas e monossílabos tônicos*.

(14) Nós tomava reguada *nas mão* se não soubesse fazê a tarefa. (VARLINGUA Inf3 F1g).

3. *Palavras paroxítonas*.

(15) Agora é diferente, a gente não pode confiá *nas pessoa*. (VARLINGUA Inf7 F2p).

4. *Palavras proparoxítonas*.

(16) A mãe ficô doente bem num congresso *dus médico* do coração. (VARLINGUA Inf1 F1p).

Seguindo Scherre (1988), a nossa hipótese para essa variável é que haja maior ocorrência de concordância se a sílaba do elemento singular for marcada, se ela for oxítona. Caso contrário, se a sílaba do elemento singular for paroxítona ou proparoxítona, não apresentando a sílaba final marcada, a ocorrência de marcas de concordância é menor.

3.5.2.3. Posição do elemento no SN

A variável posição do elemento no SN refere-se à posição que ocupa o elemento em análise no SN. Scherre (1994) analisou a posição linear de cada integrante flexional do SN e verificou que a maior ocorrência das marcas de plural aparece em todos os elementos nominais determinantes antepostos ao núcleo, como nos exemplos que ela nos oferece:

(17) (3)¹⁹ **novas** escolas/**aquelas** cruzinha toda. (SCHERRE, 1994, p. 4).

Já os casos em que aparecem menos marcas de plural estão nos elementos nominais determinantes pospostos ao núcleo:

(18) (4) essas estradas **nova**/dez senhoras lá **sentada**. (SCHERRE, 1994, p. 4).

¹⁹ A segunda numeração entre parênteses segue a numeração de Scherre (1994), assim como o destaque dos exemplos.

Geralmente, os elementos determinantes ou o núcleo são claramente mais marcados quando ocupam a primeira posição do SN e menos marcados os que ocupam a segunda posição do SN:

(19) (5) **coisas** lindas/**eles** todo. (SCHERRE, 1994, p. 5).

(20) (6) os **menino**/ dois **ano**. (SCHERRE, 1994, p. 5).

Desse modo, os elementos que ocupam a terceira, quarta ou quinta posição recebem com maior frequência a marca do plural que os elementos da segunda posição:

(21) (7) os outros **colégios**/ os meus ainda velhos **amigos**. (SCHERRE, 1994, p. 5).

Assim, podemos perceber que os elementos que estão à esquerda do núcleo, que são de primeira posição, são mais propensos à marcação explícita de plural que os elementos posicionados à direita do núcleo, ou também, da segunda e das outras posições. E, ainda, Scherre (1994) revela o forte aumento do uso da marca de plural no núcleo que ocupa a terceira, quarta ou quinta posição.

O grupo de fatores considerado na análise da variável *posição do elemento* no SN é composto por quatro fatores:

1. *Primeira posição.*

(22) **Os** casamento duravum a vida intera. (VARLINGUA Inf4 M1g).

2. *Segunda posição.*

(23) A criação depende muito **dus pais** né. (VARLINGUA Inf10 M2g).

3. *Terceira posição.*

(24) Eu já não tenho participação **nus otrus setor**, daí eu não sei como que é a convivência delis lá. (VARLINGUA Inf15 F1g).

4. *Quarta ou demais posições.*

(25) Eu acho que, muita gente não conhece i tem **umas história super legal**. (VARLINGUA Inf2 M1p).

Conforme foi proposto por Scherre (1988), nossa hipótese para essa variável é que os elementos que estão na primeira posição, geralmente à esquerda do núcleo, serão mais favoráveis à marcação de plural. Os elementos posicionados na segunda e nas demais posições serão desfavoráveis à marcação de plural.

3.5.2.4. Classe gramatical do elemento

No tocante ao fator classe gramatical do elemento no SN, Scherre (1994) argumenta que classe gramatical e posição do elemento são inter-relacionados, tendo em vista a relação que se estabelece entre os determinantes e o núcleo no SN, pois acredita que o isolamento dos fatores não dá conta de explicar o fato da variação na CN, e, por isso, a autora trabalhou com os dois fatores de forma cruzada.

Scherre (1994) analisou a posição do elemento e classe gramatical do elemento na perspectiva atomística, ou seja, cada elemento flexionável do SN a partir de dois grupos de fatores: localização do SN na oração e configuração sintagmática do SN. A autora verificou que o sintagma localizado à esquerda da oração geralmente carrega todas as marcas explícitas de plural, diferentemente do sintagma localizado à direita da oração ou em posição indistinta.

Em sua análise quantitativa, Scherre analisou 948 sintagmas nominais (SNs). Com relação à influência da localização do SN na oração, analisou da seguinte forma: localização à esquerda; localização à direita; localização indistinta. Novamente, a partir da análise feita com as duas perspectivas, a atomística e a não-atomística, Scherre conclui que a posição à esquerda favorece mais o uso das marcas explícitas de plural, considerando a hipótese de que esta influência se deve à posição de tópico à esquerda.

Scherre (1994) percebeu que a influência da configuração sintagmática do SN com a presença das marcas de plural ocorre mais em casos em que há um determinante definido ou um quantificador na primeira posição do SN, seguida de um substantivo na última posição do SN, como nos exemplos:

(26) (8) **todos os dias/os meus filhos.** (SCHERRE, 1994, p. 7).

Seguem, abaixo, as conclusões de Scherre (1994) assim resumidas:

“1) A presença de artigo definido ou do quantificador, na primeira posição do SN, e a presença de substantivo ou categoria substantivada, na última posição do SN, são duas forças poderosas no sentido de provocarem a

presença de todas as marcas de plural em todos os elementos flexionáveis do SN.

2) A presença de qualquer outra classe gramatical que não seja artigo definido ou quantificador, na primeira posição do SN, e a presença do adjetivo, do possessivo ou do quantificador, na última posição do SN, são igualmente duas forças poderosas no sentido de desfavorecer a presença de SNs com todas as marcas.

3) O cruzamento das categorias mencionadas em 1 e 2 mostra o equilíbrio das duas forças no sentido de influenciar a presença de SNs com todas as marcas.

4) A estrutura sintagmática composta desfavorece a existência de SNs totalmente marcados". (SCHERRE, 1994, p. 7).

Apesar de Scherre (1988), já ter demonstrado que o isolamento das variáveis classe gramatical e posição do elemento não dá conta de explicar satisfatoriamente a variação na CN, em nossa análise consideramos, inicialmente, as variáveis classe gramatical e posição do elemento isoladamente para, assim, observarmos os pesos relativos de cada fator.

O grupo de fatores considerado na análise da variável *classe gramatical do elemento no SN* divide-se em oito fatores:

1. Artigo.

(27) *As pessoa* se respeitavam mais. (VARLINGUA Inf7 F2p).

2. Substantivo.

(28) Você imitava o que *os pais* faziam né, o trabalho dos pais. (VARLINGUA Inf9 F2g).

3. Adjetivo.

(29) intão, num tive assim, *coisas triste* no passado. (VARLINGUA Inf4 M1g).

4. Demonstrativo.

(30) Lá no sítio era o que nós fazia, era caçá passarinho, *essas coisa* de muleque. (VARLINGUA Inf2 M1p).

5. Possessivo.

(31) *Meus pais* erum, erum agricultoris né. (VARLINGUA Inf12 M2c).

6. *Quantificador.*

(32) Ela tirô **todus** *us medo* dele, *todus*. (VARLINGUA Inf3 F1g).

7. *Indefinido.*

(33) Eu sei **algumas** *palavras* assim, mais não de sabê falá, dialogá im otra língua. (VARLINGUA Inf5 F1c).

8. *Categoria substantivada*²⁰.

(34) Veio só eu, **os** *otrus* ficarum pra lá daí. (VARLINGUA Inf14 M1p).

A partir dos postulados de Scherre (1988, 1994), construímos nossa hipótese para essa variável: acreditamos que os *substantivos*, *adjetivos* e *categorias substantivadas* desfavoreçam a CN na amostra de Guarapuava, tendo em vista que, na maioria das ocorrências, essas classes gramaticais estão na segunda ou nas demais posições do SN.

3.5.2.5. Marcas precedentes

A variável marcas precedentes já foi atestada em diversos estudos. Essa variável analisa a influência do elemento precedente ao elemento analisado para a *presença de CN* ou a *ausência de CN*. O objetivo dessa variável é confirmar se “marcas levam a marcas e zeros levam a zeros”, entre os elementos do SN, segundo Scherre (1988).

Por exemplo, se um elemento do SN recebe a marca –s de plural, este elemento condicionaria o aparecimento da marca de plural no elemento subsequente. O mesmo ocorreria caso fosse um elemento com ausência da marca de plural, como ilustram os exemplos abaixo, de Brandão (2007, p. 65):

(35) (12a) TodoS oS meuS alunoS leram o livro.

(36) (12b) Comentei oS livroØ didáticoØ.

Desse modo, esta análise contém nove variantes, que serão analisadas, são elas:

²⁰ Entendemos por *categoria substantivada* aquela que “envolve todas as classes gramaticais (adjetivos, possessivos e indefinidos) que ocorrem em SNs que não têm um substantivo ou um pronome pessoal como núcleo” (SCHERRE, 1988, p. 151).

1. *Ausência de elemento precedente.* (Quando não há nenhum elemento precedente ao elemento em análise, geralmente, ocorre em casos de elemento que ocupa a primeira posição do SN).

(37) O pai na parte da casa ele até que não si envolvia, era a mãe que tomava *as* **decisão**. (VARLINGUA Inf4 M1g).

2. *Elemento precedente é um numeral.* (Quando há um numeral precedente ao elemento em análise).

(38) Hoje eu tenho *três* **filho** né. (VARLINGUA Inf10 M2g).

3. *Ausência de marca morfológica no elemento precedente.* (Quando o elemento precedente ao elemento em análise não possui a marcação de plural).

(39) A *minhas* **tias** também custuravum, a gente sabia um poquinho né. (VARLINGUA Inf7 F2p).

4. *Apenas uma marca (em -s) precedente.* (Quando há um elemento precedente ao elemento em análise com a marca de plural).

(40) Intão hoje *us* **casamento** são bem melhor. (VARLINGUA Inf3 F1g).

5. *Dois ou mais marcas (em -s) precedente.* (Quando há dois ou mais elementos precedentes com a marca de plural).

(41) Ela mi ajudava lavando *as* **minhas** **ropas**. (VARLINGUA Inf5 F1c).

6. *Mistura de marcas com zero precedente à 3ª e 4ª posição.* (Ocorre entre os elementos anteriores ao elemento em análise, ou seja, o primeiro elemento com a marca de plural e o segundo elemento sem a marca de plural).

(42) Cada um tinha *suas* **tarefa** **doméstica**. (VARLINGUA Inf9 F2g).

7. *Mistura de marcas com -s precedente à 3ª e 4ª posição.* (Ocorre entre os elementos anteriores ao elemento em análise, ou seja, o primeiro elemento sem a marca de plural e o segundo elemento com a marca de plural).

(43) Olha, eu sei porque eu tenho *o meus subrinho*. (VARLINGUA Inf21 F2g).

8. *Sintagma preposicionado precedente*. (Quando há um sintagma preposicionado precedente ao elemento em análise).

(44) Eu, tiu, tia, *um monte de primos*. (VARLINGUA Inf2 M1p).

9. *Termo precedente é invariável*. (Quando o elemento precedente ao elemento em análise não possui variação).

(45) No mercado eu fui repositor, fui cartazista, fui, controlador de istoque, tive *várias função* dentro do mercado. (VARLINGUA Inf18 M1c).

Para a variável marcas precedentes, o nosso objetivo é verificar se os resultados de Guarapuava assemelham-se aos de Scherre (1988 e 1998) e Fernandes (1996), ou seja, verificar se os fatores linguísticos e sociais também em nossos dados favorecem a CN.

Na sequência, serão apresentadas as variáveis sociais consideradas neste estudo.

3.6. As variáveis sociais

Nesta seção, apresentamos as variáveis sociais ou extralinguísticas selecionadas para análise, que são: *sexo, faixa etária e escolaridade*. Por meio dessas variáveis, pretendemos identificar quais fatores condicionam o uso da variável CN em Guarapuava.

3.6.1. A variável sexo

A variável *sexo* na questão da variação e mudança linguísticas tem papel fundamental, visto que, na proporção da ocorrência de algumas variantes com relação à forma padrão e não-padrão, segundo Labov ([1972] 2008), as mulheres são mais propensas a utilizarem as formas linguísticas de maior prestígio que os homens, por serem mais sensíveis aos valores sociolinguísticos.

Nesse sentido, sobre o processo de mudança,

[...] as mulheres tendem a assumir a liderança da mudança. Ao contrário, quando se trata de implementar uma forma socialmente desprestigiada, as

mulheres assumem uma atitude conservadora e os homens tomam a liderança do processo. (MOLLICA, 2013, p. 36).

Desse modo, podemos observar que as mulheres são mais dispostas a assumir a forma considerada padrão na língua portuguesa e tomam posição à frente dos homens no processo de mudança.

A partir do que foi desenvolvido por Scherre (1988), a nossa hipótese para este estudo é que as mulheres utilizem com maior frequência a variante *presença de CN*, enquanto forma padrão, que os homens em Guarapuava. Para isso, nesta seção, iremos trabalhar com a variável sexo considerando dois fatores:

F. *Feminino*;

M. *Masculino*.

3.6.2. A variável faixa etária

A significativa distribuição dos falantes pela faixa etária está comumente relacionada ao que Labov ([1972] 2008) definiu como “mudança em tempo aparente”, uma concepção baseada na hipótese clássica da aquisição da linguagem (Bailey 2003), o que, teoricamente, significa fazer uma comparação da língua falada pelas diferentes faixas etárias, e com isso indicar uma mudança em progresso na língua.

Ainda segundo o autor, um processo de mudança linguística pode ocorrer quando uma determinada variante é mais frequente na fala da faixa etária mais jovem e menos frequente na dos mais idosos. Caso a forma variante em longo prazo fosse adotada pelos falantes mais idosos, seria permitido falar em mudança em progresso. Nas palavras de Labov,

Gerações sucessivas de falantes dentro do mesmo subgrupo, reagindo às mesmas pressões sociais, fazem a variável linguística avançar mais um passo no processo de mudança, para além do modelo estabelecido por seus pais. (LABOV, [1972] 2008, p. 210).

Após constatada a mudança baseada nos dados do tempo aparente, conforme Tarallo (2001, p. 71), “é fundamental que o investigador procure dar à variável a dimensão histórica do tempo real”, ou seja, realizar um estudo no tempo real para completar o percurso histórico da variável analisada.

Para verificarmos nesta pesquisa se a *faixa etária* é um fator importante na escolha linguística dos informantes de Guarapuava e, por consequência, se há indícios de uma mudança em progresso nessa comunidade, elegemos duas faixas etárias:

1. 25 a 45 anos;
2. 50 anos ou mais.

3.6.3. A variável escolaridade

A variável escolaridade tem exercido grande influência nos estudos sociolinguísticos, no sentido de que a escola tenta preservar as formas de prestígio, tidas como referências do bom uso da língua e não aceita as variações. Por exemplo, um grupo de falantes com menos tempo de escolaridade, pode utilizar mais a variante *ausência de CN*, isso porque a tendência geral é que quanto mais escolaridade, maior é o uso das normas gramaticais trabalhadas no ambiente escolar.

Sobre os casos de concordância, Mollica (2013, p. 53), afirma:

As regras de concordância nominal e verbal e os estudos de regência verbal ilustram bem a ênfase que se dá ao padrão culto. Ao justificar e implantar na literatura e no uso das pessoas cultas as formas focalizadas, o ensino contribui para provocar adesão dos alunos a essas formas, como marcas de prestígio social.

Para analisar essa variável, o presente estudo considera três fatores:

- p. *Fundamental I*;
- g. *Fundamental II*;
- c. *Ensino Médio*.

A nossa hipótese, tendo como base a pesquisa de Scherre (1988) é verificar se a variante *presença de CN* predomina nos falantes com maior grau de escolaridade que nos falantes menos escolarizados.

No próximo capítulo, serão apresentados os resultados obtidos na análise dos dados da CN em Guarapuava.

CAPÍTULO 4 – RESULTADOS DA CONCORDÂNCIA NOMINAL EM GUARAPUAVA

Neste capítulo, são apresentados os resultados da análise da CN, fornecidos pelo programa estatístico GoldVarbX, a partir dos dados dos 24 informantes da comunidade de Guarapuava.

Os resultados gerais da análise quantitativa dos dados dos 24 informantes totalizaram 4.213 ocorrências, das quais foram registrados 1.684 casos de *ausência de CN*, e 2.529 casos de *presença de CN*, o equivalente a 40% dos casos para a ausência e 60% dos casos para a presença do fenômeno analisado.

No entanto, considerando os diferentes constituintes do SN, verificamos que os determinantes apresentam a marca de plural em quase 100% das ocorrências. Já nos elementos nucleares (substantivos e categorias substantivadas) e nos adjetivos predomina a ausência de concordância, ou seja, há uma alta frequência de cancelamento da marca (67% e 70%, respectivamente).

Em uma rodada que realizamos sem os determinantes, isto é, somente com *substantivos*, *adjetivos* e *categorias substantivadas*, obtivemos um total de 2.422 ocorrências, 32% (782 ocorrências) com *presença de CN* e 68% (1.640 ocorrências) com *ausência de CN*, o que comprova o elevado índice de *ausência de CN* nesses constituintes do SN.

Para a análise estatística dos dados, as variáveis consideradas foram:

1. *Saliência fônica*;
2. *Tonicidade*;
3. *Posição do elemento no SN* (doravante *posição*);
4. *Classe gramatical do elemento* (doravante *classe gramatical*);
5. *Marcas precedentes*;
6. *Sexo*;
7. *Faixa etária*;
8. *Escolaridade*.

Realizamos algumas rodadas com essas variáveis isoladamente e algumas delas depois foram aliadas²¹. Realizamos, no total, cinco rodadas:

1. Rodada específica com as variáveis *saliência fônica* e *tonicidade* isoladas;
2. Rodada específica com as variáveis *posição* e *classe gramatical* isoladas;

²¹ Consideramos como aliadas duas variáveis analisadas em um mesmo grupo de fatores, ou seja, em conjunto.

3. Rodada definitiva com a variável *saliência e tonicidade aliadas* com a exclusão do fator *monossílabos átonos*;
4. Rodada definitiva com a variável *posição e classe gramatical aliadas*;
5. Rodada final.

Nesta última rodada, foram incluídas todas as variáveis linguísticas (já com as variáveis aliadas) e as variáveis sociais, com seus respectivos fatores. Primeiramente, serão apresentados os resultados gerais da rodada final da amostra de Guarapuava. Em seguida, discutiremos cada uma das variáveis selecionadas e apresentaremos as análises realizadas para chegarmos ao resultado final. Além disso, apresentaremos a análise da rodada efetuada somente com os *substantivos*, os *adjetivos* e as *categorias substantivadas*.

4.1. Resultados gerais da CN em Guarapuava

Para a rodada final, foram consideradas as seguintes variáveis: 1) *saliência e tonicidade aliadas*; 2) *posição e classe gramatical aliadas*; 3) *marcas precedentes*; 4) *sexo*; 5) *faixa etária*; e 6) *escolaridade*.

Foram selecionadas pelo programa estatístico GoldVarbX, as variáveis consideradas significativas, nesta ordem: 1) *posição e classe gramatical aliadas*; 2) *escolaridade*; 3) *marcas precedentes*; 4) *saliência e tonicidade aliadas*; e 5) *sexo*. Na tabela abaixo, apresentamos os resultados probabilísticos da rodada final:

TABELA 9: Resultados da CN em Guarapuava/PR- rodada final

Grupo de fatores	Aplicação /Total	%	P.R.
1. Posição e classe gramatical aliadas			
b. Determinantes antepostos ao núcleo na 2ª posição	113/115	98%	0.98
i. Determinantes pospostos ao núcleo	3/4	75%	0.78
a. Determinantes antepostos ao núcleo na 1ª posição	1634/1676	97%	0.73
g. Categoria substantivada como núcleo na 2ª posição	24/46	52%	0.57
j. Adjetivo posposto ao núcleo na 2ª posição	6/18	33%	0.40
e. Substantivo como núcleo na 2ª posição	625/1974	32%	0.29
c. Adjetivo anteposto ao núcleo	9/25	36%	0.23
k. Adjetivo posposto ao núcleo nas demais posições	16/61	26%	0.20
f. Substantivo como núcleo na 3ª e 4ª posição	80/259	31%	0.20
h. Categoria substantivada como núcleo na 3ª e 4ª posição	3/19	16%	0.06
2. Escolaridade			
c. Ensino médio	791/1141	69%	0.68
g. Fundamental II	1009/1723	59%	0.46
p. Fundamental I	729/1349	54%	0.38
3. Marcas precedentes			
1. ausência de elemento precedente	1653/1695	97%	0.82
3. ausência de marca morfológica no elemento precedente	38/41	93%	0.64
2. elemento precedente é um numeral	305/708	43%	0.40
5. duas ou mais marcas (em -s) precedente	37/95	39%	0.39
9. termo precedente é invariável	12/40	30%	0.36
8. sintagma preposicionado precedente	3/15	20%	0.26
7. mistura de marcas com -s precedente à 3ª e 4ª posição	9/41	22%	0.22
4. apenas uma marca (em -s) precedente	469/1541	30%	0.20
6. mistura de marcas com zero precedente à 3ª e 4ª posição	3/37	8%	0.05
4. Saliência e tonicidade aliadas			
D. Plural duplo	15/19	79%	0.92
L. Itens terminados em -l	26/51	51%	0.72
S. Itens terminados em -s	47/84	56%	0.65
R. Itens terminados em -r	41/97	42%	0.61
O. Regular oxítono	46/112	41%	0.59
X. Regular proparoxítono	21/67	31%	0.56
A. Itens terminados em -ão	16/61	26%	0.51
P. Regular paroxítono	1106/2477	45%	0.47
5. Sexo			
F. Feminino	1321/2125	62%	0.53
M. Masculino	1208/2088	58%	0.46
Total	2529/4213	60%	
<i>Input: 0.74</i>			

Fonte: Dados do VARLINGUA 2015

Conforme os resultados da tabela 9, considerando todos os elementos do SN, obtivemos em nossos dados uma porcentagem de 60% e um *input*²² de 0.74 para a aplicação de CN. No entanto, conforme podemos observar na tabela 9, os *determinantes* apresentam elevada marcação de plural, ao contrário das demais classes gramaticais. Isso parece mostrar que a CN propriamente dita, que se revela principalmente nos *substantivos*, *adjetivos* e *categorias substantivadas*, é bastante reduzida em nossos dados.

Esse resultado também foi comprovado em uma rodada que realizamos apenas com os *substantivos*, *adjetivos* e *categorias substantivadas*, pois, nessa rodada obtivemos somente 32% para a aplicação da regra (*input*: 0.31), o que confirma o baixo índice de concordância nos dados de Guarapuava.

Dadas essas constatações, podemos confirmar a hipótese geral desta pesquisa: na língua falada da comunidade de Guarapuava predomina um maior número de ocorrências constituídas pela *ausência de CN* que pelas formas marcadas pela *presença de CN*.

Em relação às variáveis selecionadas, como podemos observar na tabela 9, todas as variáveis linguísticas foram selecionadas pelo GoldVarbX, e isso comprova a importância das análises realizadas com as variáveis aliadas, num mesmo conjunto de fatores. Além disso, das variáveis sociais, a *escolaridade* foi a primeira selecionada pelo GoldVarbX e, provavelmente, a mais significativa para o condicionamento da CN em Guarapuava.

Na próxima seção, apresentamos os resultados das variáveis linguísticas, e, na seção posterior, as variáveis sociais selecionadas como significativas em nossa amostra.

4.2. Resultados das variáveis linguísticas em Guarapuava

A partir da totalidade dos dados, das 3 (três) variáveis linguísticas adotadas para a rodada final, todas se mostraram significativas. Pela ordem de seleção do programa estatístico, são elas: *posição e classe gramatical aliadas*; *marcas precedentes*; e *saliência e tonicidade aliadas*.

4.2.1. Posição e classe gramatical aliadas

A variável *posição e classe gramatical aliadas*, isto é, a variável que considera a *posição* do elemento no SN e sua *classe gramatical* em um mesmo grupo de fatores, foi a

²² O *input* refere-se ao “nível geral de uso de determinado valor da variável dependente” (GUY e ZILLES, 2007, p. 238).

primeira variável selecionada, ou seja, foi considerada a mais significativa pelo programa estatístico. Para um melhor entendimento, esclarecemos na sequência como realizamos a análise para chegarmos a essa variável.

Tomando por base outros estudos, especialmente o embasamento teórico de Scherre (1988), realizamos primeiramente uma rodada com as variáveis *posição* e *classe gramatical* isoladas. O objetivo dessa rodada foi o de observar o peso relativo de cada fator isoladamente.

Os resultados obtidos seguem nas tabelas 10 e 11, abaixo:

TABELA 10: Resultados da posição na presença de CN

Fatores	Aplicação/Total	%	P.R.
Primeira posição	1653/1695	97%	0.88
Segunda posição	775/2172	36%	0.21
Terceira posição	91/314	29%	0.17
Quarta ou demais posições	10/32	31%	0.17
Total	2529/4213	60%	

Fonte: Dados do VARLINGUA 2015

Como podemos observar na tabela 10, os resultados são altamente significativos, sendo a *primeira posição* (0.88) favorável à *presença de CN*, já na *segunda posição* (0.21), assim como na *terceira* e na *quarta ou demais posições*, com o mesmo peso relativo (0.17), predominou a *ausência de CN*. Constatamos, portanto, que todas as posições, com exceção da primeira, desfavoreceram a concordância, ocorrendo uma queda brusca na marcação de plural em relação à *primeira posição*.

Observando os resultados de Fernandes (1996, p. 48), para a *primeira posição* (0.77), a *segunda posição* (0.33) e a *terceira e quarta posição* (0.27), podemos observar que os resultados se mostraram semelhantes aos de Guarapuava, pois a *primeira posição* também foi a mais marcada, e as demais posições, obtiveram um peso relativo baixo desfavorecendo a concordância.

Os estudos de Scherre (1996), citados no subitem 2.2, também já demonstraram essa tendência à maior marcação de plural na *primeira posição* do SN:

A primeira posição do SN é o fator que mais favorece a inserção da marca formal de plural, ocorrendo uma queda brusca em relação às demais posições, as quais tendem a desfavorecer de forma decrescente a presença de marca formal de plural no SN. (SCHERRE, 1996, p. 45).

Sobre isso, Scherre (1988) argumenta que esse fato se dá em decorrência do fenômeno da redundância de marcas de plural, característica da língua portuguesa. A marca em apenas um elemento do SN, já é suficiente para indicar a pluralidade, em outras palavras, apenas uma marca é necessária para estabelecer a informação de plural.

Assim, a nossa hipótese de que o primeiro elemento do SN é o mais marcado se confirmou nos dados de Guarapuava. A seguir, apresentamos os resultados da variável *classe gramatical*:

TABELA 11: Resultados da classe gramatical do elemento na presença de CN

Fatores	Aplicação/Total	%	P.R.
Indefinido	212/213	99%	0.97
Possessivo	158/160	99%	0.96
Artigo	1110/1144	97%	0.63
Categoria substantivada	27/65	41%	0.38
Quantificador	74/81	91%	0.32
Substantivo	723/2252	32%	0.31
Adjetivo	32/105	30%	0.27
Total	2336/4020	58%	

Fonte: Dados do VARLINGUA 2015

A variável *classe gramatical* apresentou um nocaute²³ para o fator *demonstrativo*, com 192 ocorrências, e todas elas com a aplicação da regra de CN, ou seja, o uso da concordância foi categórico (100%). O pronome *demonstrativo*, portanto, embora tenha sido retirado da nossa análise, mostrou-se altamente favorável à marcação de plural, como em Scherre (1988). Na maioria dos estudos mencionados (Scherre, 1988, 1998), Fernandes (1996) e Poplack (1980), os *demonstrativos* também favoreceram a marcação de plural.

Como podemos verificar nos resultados, obtivemos como fatores mais favorecedores da CN: o pronome *indefinido* (0.97) e o *possessivo* (0.96). O *artigo*, com um número elevado de ocorrências (1110), apresentou um peso relativo de (0.63), também favorecendo a concordância.

²³ A terminologia *nocaute* ou (*KnockOut*) representa um fator “que, num dado momento de análise correspondente a uma frequência de 0% ou 100% para um dos valores da variável dependente.” (GUY e ZILLES, 2007, p. 158). Em outras palavras, o *nocaute* ocorre quando um dos fatores registra 0% ou 100% para a variável dependente, ou seja, não há variação, portanto, esse fator não precisa ser analisado.

Logo abaixo na tabela, as classes gramaticais: *categoria substantivada* (0.38), *quantificador* (0.32), *substantivo* (0.31) e *adjetivo* (0.27), apresentaram pesos relativos desfavorecedores da CN em nossos dados. A classe gramatical que somou o maior número de ocorrências foi o *substantivo* (2252), ocupando, principalmente, a segunda ou demais posições do SN.

Na análise de Fernandes (1996, p. 51), o *possessivo* foi o fator que mais apresentou concordância (0.94), seguido do *artigo*²⁴ (0.77), e do *indefinido* (0.76), fatores que também favoreceram a CN nos dados de Guarapuava, embora não na mesma ordem. As demais classes gramaticais desfavoreceram a aplicação da regra, tanto em Guarapuava como no estudo da autora. Por exemplo, a *categoria substantivada* (0.38 e 0.45), o *quantificador* (0.32 e 0.44), o *substantivo* (0.31 e 0.29), e por último o *adjetivo* (0.27 e 0.44, respectivamente).

Os resultados obtidos nos dados de Guarapuava confirmam nossa hipótese para a variável classe gramatical, pois os *substantivos*, *adjetivos* e *categorias substantivadas*, geralmente em segunda ou demais posições desfavoreceram a CN.

De acordo com Scherre (1988), a melhor forma de analisar a CN é aliar as variáveis *posição* e *classe gramatical*, ou seja, analisá-las em conjunto, a fim de obter resultados mais significativos. Com base nos estudos da autora, realizamos uma rodada com as variáveis *posição* e *classe gramatical* aliadas.

Nesta análise, assim como Scherre (1988) em sua tese, aliamos a variável *posição* e *classe gramatical*, incluindo, também, a posição relativa, que se refere à distribuição dos elementos em relação ao núcleo. Para esta análise, os fatores considerados foram:

a. *Determinantes antepostos ao núcleo na 1ª posição.*

(46) *As brincadeira*, eu sempre penso qui é muito diferente de agora né. (VARLINGUA Inf7 F2p).

b. *Determinantes antepostos ao núcleo na 2ª posição.*

(47) Eu tenho mais ligação com a minha sogra né, i com *us meus cunhado*. (VARLINGUA Inf24 M2c).

c. *Adjetivo anteposto ao núcleo.*

²⁴ Fernandes (1996) considerou os artigos e os demonstrativos num único fator.

(48) Ela i mais um foram *us melhores aluno* né, ela vai muito bem, bem istudiosa. (VARLINGUA Inf1 F1p).

d. *Substantivo como núcleo na 1ª posição.*

(49) Não vô dizê qui eu sei fazê *pratus diferente* assim né, mais o básico. (VARLINGUA Inf10 M2g).

e. *Substantivo como núcleo na 2ª posição.*

(50) Tenho *minhas amizade*, saio, passeio. (VARLINGUA Inf21 F2g).

f. *Substantivo como núcleo na 3ª e 4ª posição.*

(51) *As maior reclamação* que a gente iscuta hoje im dia é questão de saúde i questão de segurança pública, im Guarapuava. (VARLINGUA Inf18 M1c).

g. *Categoria substantivada como núcleo na 2ª posição.*

(52) Eu adotei ela como mãe, eu acho que *us otrus* também. (VARLINGUA Inf13 F1p).

h. *Categoria substantivada como núcleo na 3ª e 4ª posição.*

(53) *Os mais velho* gostavum de caçá passarinho. (VARLINGUA Inf12 M2c).

i. *Determinantes pospostos ao núcleo.*

(54) *Os professoris meus* erum daquele sistema assim de, ainda do puxão de orelha. (VARLINGUA Inf5 F1c).

j. *Adjetivo posposto ao núcleo na 2ª posição.*

(55) Não precisa sê *grandis obras* né, *grandis feitos* né, são *coisas simplis*. (VARLINGUA Inf4 M1g).

k. *Adjetivo posposto ao núcleo nas demais posições.*

(56) Nós temo muito cachorro, aí eu digo que são *meus filho adotivo*. (VARLINGUA Inf15 F1g).

Tendo como base Scherre (1988, 1996), nossa hipótese para essa variável é que os determinantes, geralmente em *primeira posição*, favoreçam a *presença de CN*, já nas demais classes gramaticais, substantivos, adjetivos e categorias substantivadas, predomina a *ausência de CN*. Na tabela seguinte, apresentamos os resultados para a variável *posição e classe gramatical aliadas*:

TABELA 12: Resultados da posição e classe gramatical aliadas na presença de CN

Fatores	Aplicação/Total	%	P.R.
b. determinantes antepostos ao núcleo na 2ª posição Ex: todus us dias (VARLINGUA Inf1 F1p).	113/115	98%	0.98
i. determinantes pospostos ao núcleo Ex: treis filhus meu (VARLINGUA Inf10 M2g).	3/4	75%	0.78
a. determinantes antepostos ao núcleo na 1ª posição Ex: us passarinho (VARLINGUA Inf6 M1c).	1634/1676	97%	0.73
g. categoria substantivada como núcleo na 2ª posição Ex: us guarapuavanus (VARLINGUA Inf4 M1g).	24/46	52%	0.57
j. adjetivo posposto ao núcleo na 2ª posição Ex: cavalus criolo (VARLINGUA Inf14 M1p).	6/18	33%	0.40
e. substantivo como núcleo na 2ª posição Ex: us professor (VARLINGUA Inf11 F2c).	625/1974	32%	0.29
c. adjetivo anteposto ao núcleo Ex: grandes obras (VARLINGUA Inf12 M2c).	9/25	36%	0.23
k. adjetivo posposto ao núcleo nas demais posições Ex: uns prato diferente (VARLINGUA Inf5 F1c).	16/61	26%	0.20
f. substantivo como núcleo na 3ª e 4ª posição Ex: os meus filho (VARLINGUA Inf10 M1g).	80/259	31%	0.20
h. categoria substantivada como núcleo na 3ª e 4ª posição Ex: us mais antigo (VARLINGUA Inf16 F1c).	3/19	16%	0.06
Total	2513/4197	60%	

Fonte: Dados do VARLINGUA 2015

A variável *posição e classe gramatical aliadas* apresentou um nocaute para o fator *substantivo como núcleo na primeira posição*, com 16 ocorrências, e todas elas com a aplicação da regra de concordância.

Pela tabela, é possível notar que os *determinantes (artigos, possessivos, indefinidos e quantificadores) antepostos ao núcleo na 2ª posição* apresentaram um peso relativo quase categórico para a *presença de CN* (0.98). Em seguida, seguem os *determinantes pospostos ao*

núcleo como favorecedores da concordância (0.78), os *determinantes antepostos ao núcleo na 1ª posição* (0.73), e a *categoria substantivada como núcleo na 2ª posição* (0.57).

Por outro lado, desfavoreceram a concordância, o *adjetivo posposto ao núcleo na 2ª posição* (0.40), o *substantivo como núcleo na 2ª posição* (0.29), o *adjetivo anteposto ao núcleo* (0.23), o *adjetivo posposto ao núcleo nas demais posições* (0.20), e o *substantivo como núcleo na 3ª e 4ª posição* (0.20). E com uma probabilidade ainda maior, a concordância também foi desfavorecida na *categoria substantivada como núcleo na 3ª e 4ª posição* (0.06).

Verifica-se, portanto, que em nossos dados, a *primeira posição*, ocupada principalmente pelos determinantes, favoreceu a CN. Lembramos que os *substantivos*, com um uso categórico da concordância nessa posição, foram retirados da análise. Na *segunda posição*, os determinantes e a *categoria substantivada* favoreceram a CN e os *adjetivos* e *substantivos* a desfavoreceram. Já na terceira e demais posições, todas as classes gramaticais desfavoreceram a aplicação da regra.

Scherre (1988, p. 214), observando seus dados a partir da posição, verifica que, para todas as classes gramaticais, é a *primeira posição* que mais favorece o índice de marcas formais de concordância. Na *segunda posição*, a autora constata que a CN mostra diferenças mais acentuadas em função da classe gramatical.

Seus resultados mostram que os *artigos*, os *demonstrativos* e os *indefinidos* apresentam um favorecimento quase categórico de concordância. Os *adjetivos* apresentam um resultado próximo ao ponto neutro (0.48), a *categoria substantivada* (0.43) e o *substantivo* (0.33) desfavorecem a CN, sendo esta última a classe menos marcada. Na *terceira posição*, Scherre (1988) constata que todas as classes gramaticais apresentam um baixo índice de marcas de plural.

Fernandes (1996), em sua análise, constatou que na *primeira posição* os determinantes e também a classe nuclear (*substantivo* e *categoria substantivada*) apresentam um peso relativo bastante alto de concordância (0.85 e 0.67, respectivamente). Na *segunda posição*, seus resultados mostram que a classe não nuclear anteposta ao núcleo favorece a concordância (0.72) e que as demais classes a desfavorecem (classe não nuclear: 0.32; classe nuclear: 0.23). Já na *terceira posição*, a probabilidade de *ausência de CN* é bastante elevada em todas as classes gramaticais.

Portanto, a partir dessas análises da CN, podemos dizer que os resultados de Guarapuava se mostraram semelhantes aos de Scherre (1988, p. 213) e Fernandes (1996, p. 57), pois em todos esses estudos os determinantes antepostos ao núcleo na primeira e segunda posição e as classes nucleares (*substantivos* e *categorias substantivadas*) na *primeira posição*

favoreceram a CN. Também em todos esses estudos, os *substantivos* (classe nuclear), na segunda, terceira e quarta posição, desfavoreceram a CN.

A maior diferença entre os resultados está nos *determinantes pospostos ao núcleo*, que apresentaram peso relativo desfavorável à CN (0.25) nos dados de Fernandes, e resultado favorável à CN em nossa amostra (0.78). Verificamos, no entanto, que nossos dados apresentaram somente quatro ocorrências de *determinantes pospostos ao núcleo*, e esse número, bastante reduzido de ocorrências, torna inviável maiores generalizações a respeito desse fator.

Em síntese, pelos resultados apresentados acima, podemos concluir que:

1. Os determinantes, que geralmente ocupam a primeira e a *segunda posição* do SN, apresentam a marca de plural;
2. Os *substantivos*, na segunda e na *terceira posição* desfavoreceram a CN. Já os substantivos na *primeira posição*, apesar do número reduzido de ocorrências, apresentaram 100% de CN;
3. Os *adjetivos*, em todas as posições, desfavoreceram a *presença de CN*.
4. A *categoria substantivada* na *segunda posição* favoreceu a CN e na terceira e na *quarta posição* a desfavoreceu.

Esses resultados confirmam a nossa hipótese para a variável *posição e classe gramatical aliadas*, pois os determinantes, geralmente em *primeira posição*, favoreceram a marcação de plural, enquanto as demais classes gramaticais, *substantivos*, *adjetivos* e *categorias substantivadas*, na maioria das vezes em segunda ou demais posições, a desfavoreceram.

4.2.2. Marcas precedentes

A variável *marcas precedentes* foi a terceira selecionada em termos de significância na rodada final do banco VARLINGUA. Segundo Scherre (1988), essa variável analisa a influência do elemento precedente ao elemento analisado para a presença ou a ausência da marca de plural.

Os resultados obtidos da variável *marcas precedentes* estão na tabela 13, adiante:

TABELA 13: Resultados de marcas precedentes na presença de CN

Fatores	Aplicação/Total	%	P.R.
1. ausência de elemento precedente Ex: <i>as coisas</i> (VARLINGUA Inf12 M2c).	1653/1695	97%	0.82
3. ausência de marca morfológica no elemento precedente Ex: <i>do meus filho</i> (VARLINGUA Inf17 F1c).	38/41	93%	0.64
2. elemento precedente é um numeral Ex: setenta i três anos. (VARLINGUA Inf13 F1p).	305/708	43%	0.40
5. duas ou mais marcas (em –s) precedente Ex: us meus materiais (VARLINGUA Inf10 M2g).	37/95	39%	0.39
9. termo precedente é invariável Ex: vários professoris . (VARLINGUA Inf18 M1c).	12/40	30%	0.36
8. sintagma preposicionado precedente Ex: um monte de tijolo (VARLINGUA Inf21 F2g).	3/15	20%	0.26
7. mistura de marcas com –s precedente à 3ª e 4ª posição Ex: da tuas metas né. (VARLINGUA Inf11F2c).	9/41	22%	0.22
4. apenas uma marca (em –s) precedente Ex: as coisa (VARLINGUA Inf15 F1g).	469/1541	30%	0.20
6. mistura de marcas com zero precedente à 3ª e 4ª posição Ex: as minina piquena . (VARLINGUA Inf20 F2p).	3/37	8%	0.05
Total	2529/4213	60%	

Fonte: Dados do VARLINGUA 2015

Pelos resultados apresentados, podemos observar que o fator *ausência de elemento precedente* apresentou um peso relativo de 0.82 para a *presença de CN*. Este resultado se mostrou coerente, pois todos esses casos são de elementos que ocupam a *primeira posição* do SN, e, por isso, recebem a marca de plural. Também a *ausência de marca morfológica no elemento precedente*, favoreceu a aplicação da concordância (0.64).

Em nossos dados, todos os demais fatores desfavoreceram a CN. O fator *elemento precedente é um numeral* apresentou um peso relativo de 0.40. Para os fatores *duas ou mais*

marcas (em -s) precedente (0.39) e *termo precedente é invariável* (0.36), também desfavoreceram a CN em nossa amostra.

O fator *sintagma preposicionado precedente* ao elemento em análise também desfavoreceu a *presença de CN* (0.26). O fator *mistura de marcas com -s precedente à 3ª e 4ª posição*, entre os elementos anteriores ao elemento analisado, isto é, o primeiro elemento sem a marca de plural e o segundo elemento com a marca, também desfavoreceu a concordância (0.22).

Na sequência, o mesmo ocorreu com o fator *apenas uma marca (em -s) precedente* com peso relativo de (0.20), e um número elevado de ocorrências (1541). Esse resultado mostra que o fato de o elemento precedente já indicar a pluralidade leva a não utilização da marca de plural no elemento seguinte, evitando, assim, a redundância.

Por último, o fator *mistura de marcas com zero precedente à 3ª e 4ª posição* entre os elementos anteriores ao elemento em análise, ou seja, o primeiro elemento com a marca de plural e o segundo elemento sem a marca, foi o fator que mais desfavoreceu a *presença de CN*, com peso relativo de (0.05).

Verificamos, assim, que nos dados de Guarapuava somente a *ausência de elemento precedente* e a *ausência de marca morfológica no elemento precedente* favoreceram a marcação de plural. Isso significa que geralmente o primeiro elemento é marcado, e quando isso não ocorre – em 38 ocorrências – é o elemento seguinte que recebe a marca de plural. Em todos os demais casos, a *ausência de CN* predomina.

Nos resultados de Scherre (1988, p. 205), o fator *ausência de elemento precedente* favoreceu a CN (0.92) e a *ausência de marca morfológica no elemento precedente* apresentou 100% de uso da marca no elemento seguinte. Os fatores *duas ou mais marcas (em -s) precedente* (0.61) e o *elemento precedente é um numeral* também favoreceram a CN (0.56).

Já o *sintagma preposicionado precedente* desfavoreceu a CN (0.34). Tal fato explica-se, segundo a autora, pela carga semântica de plural que já carregam esses elementos preposicionados, o que interfere na marcação de plural no elemento subsequente. Também o fator *mistura de marcas com zero precedente à 3ª e 4ª posição* desfavoreceu a CN (0.04).

Fernandes (1996) não analisou o primeiro elemento do SN (*ausência de elemento precedente*), pois realizou sua análise, a partir da *segunda posição* do elemento no SN. O fator *numeral* (0.64) foi o que mais favoreceu a CN, seguido do fator *duas ou mais marcas (em -s) precedente* (0.61). Já os fatores *mistura de marcas* (0.47) e *zero formal a partir da primeira posição* (0.10) desfavoreceram a CN.

Em suma, podemos verificar que os resultados de Guarapuava corroboram parcialmente os de Scherre (1988) e Fernandes (1996):

1. O fator *ausência de marca morfológica no elemento precedente* em Guarapuava favoreceu a CN. Nos estudos de Scherre (1988) e Fernandes (1996), verificamos o uso categórico da concordância para esse fator. Também o fator *ausência de elemento precedente* em nossos dados e nos de Scherre favoreceu a CN.
2. Já os fatores *duas ou mais marcas (em -s) precedente e elemento precedente é um numeral* em nossos dados desfavoreceram a CN, enquanto que nos dados de Scherre e Fernandes a favoreceram.
3. Quanto aos demais fatores, tanto em nossos dados como nos de Scherre e Fernandes desfavoreceram a CN.

4.2.3. Saliência e tonicidade aliadas

A variável *saliência e tonicidade aliadas*²⁵ foi a quinta e última selecionada pelo programa estatístico em nossa rodada final dos dados. Inicialmente, porém, com o objetivo de observarmos o peso de cada fator isoladamente, efetuamos algumas rodadas alternativas.

Na primeira rodada, incluímos somente as variáveis *saliência fônica* e *tonicidade* isoladas. Nesta etapa, a *tonicidade* foi selecionada pelo programa GoldVarbX como a variável mais significativa estatisticamente e a *saliência fônica* como a terceira variável mais relevante. Os resultados para cada variável são apresentados nas tabelas 14 e 15, na sequência:

²⁵Assim como na variável *posição e classe gramatical aliadas*, apresentamos, primeiramente, os resultados das variáveis *saliência fônica* e *tonicidade* isoladas e, depois, os resultados da variável *saliência e tonicidade aliadas* (em um mesmo grupo de fatores).

TABELA 14: Resultados da escala da saliência fônica na presença de CN

Fatores	Aplicação/Total	%	P.R.
Plural duplo	15/19	79%	0.84
Itens terminados em –s	47/84	56%	0.65
Itens terminados em –l	26/51	51%	0.60
Plural regular	2385/3902	61%	0.49
Itens terminados em –r	40/96	42%	0.48
Itens terminados em –ão	16/61	26%	0.35
Total	2529/4213	60%	

Fonte: Dados do VARLINGUA 2015

Na análise da *saliência fônica*, os resultados apresentaram favorecimento para a *presença de CN* para os três primeiros fatores: *plural duplo* (0.84), *itens terminados em –s* (0.65), e os *itens terminados em –l* (0.60).

Para o *plural regular*, que apresentou um elevado número de ocorrências (3902) em comparação aos demais itens, o resultado desfavoreceu a concordância, embora esse peso relativo esteja próximo do ponto neutro (0.49). Os *itens terminados em –r* e os *itens terminados em –ão* (0.48 e 0.35) também desfavoreceram a CN em nossa amostra.

Nos resultados de Scherre (1988, p. 79), os fatores que mais favoreceram a CN foram o *plural duplo* (0.86) e os *itens terminados em –l* (0.56), assim como em nossos dados (0.84 e 0.60, respectivamente).

O *plural regular*, os *itens terminados em –r* e os *itens terminados em –ão* desfavoreceram a CN (0.49, 0.48 e 0.35, respectivamente) em Guarapuava, e nos resultados de Scherre, com pesos relativos de 0.24, 0.48 e 0.42, respectivamente. A maior diferença entre os resultados está nos itens terminados em –s (0.65), que favoreceram a aplicação da regra em nossa amostra, e nos dados de Scherre a desfavoreceram (0.38).

Nos resultados de Fernandes (1996, p. 76), o *plural duplo*, os *itens terminados em –l*, os *itens terminados em –s* (0.81, 0.78 e 0.59, respectivamente) favoreceram a CN como em nossos dados. Com pesos relativos semelhantes e próximos do ponto neutro, em Guarapuava e no estudo de Fernandes, o *plural regular* desfavoreceu a CN (0.49 e 0.48, respectivamente).

Já os resultados para os *itens terminados em –r* e os *itens terminados em –ão* (0.48 e 0.35, respectivamente) em nossos dados desfavoreceram a CN, enquanto nos dados de Fernandes, esses itens favoreceram a marcação de plural (0.74 e 0.75, respectivamente).

Por meio dos resultados da tabela 10, constatamos que a nossa hipótese para a variável *saliência fônica* não se confirmou totalmente nessa primeira rodada dos dados de Guarapuava, tendo em vista que alguns itens mais salientes (terminados em *-r* e *-ão*) apresentaram menos concordância que o *plural regular* (0.49).

A seguir, apresentamos os resultados obtidos para a variável *tonicidade*:

TABELA 15: Resultados da tonicidade na presença de CN

Fatores	Aplicação/Total	%	P.R.
Monossílabos átonos	1211/1245	97%	0.93
Paroxítonos	1109/2480	45%	0.24
Oxítonos e monossílabos tônicos	188/421	45%	0.21
Proparoxítonos	21/67	31%	0.15
Total	2529/4213	60%	

Fonte: Dados do VARLINGUA 2015

Para a variável *tonicidade*, os *monossílabos átonos* favoreceram altamente a concordância (0.93), ao contrário dos outros fatores, que apresentaram uma elevada probabilidade de *ausência de CN*. Na escala, os *paroxítonos* apresentaram um peso relativo somente de 0.24 para a concordância, com o maior número de ocorrências (2480) em nossos dados.

Assemelhando-se aos *paroxítonos*, os *oxítonos e monossílabos tônicos* (0.21) e por último, os *proparoxítonos* (0.15) também desfavoreceram a aplicação da regra de concordância neste estudo.

Os resultados de Scherre (1988, p. 80), para esses fatores foram: *oxítonos e monossílabos tônicos* (0.66), *proparoxítonos* (0.44) e *paroxítonos* (0.39). Já os resultados de Fernandes²⁶ (1996, p. 86) para esses fatores foram: *oxítonos e monossílabos tônicos* (0.58), *paroxítonos* (0.48) e *proparoxítonos* (0.47), estes dois últimos com pesos relativos próximos ao ponto neutro.

Comparando os resultados de Scherre (1988) e Fernandes (1996) com os de Guarapuava, houve um desfavorecimento bem maior da CN em nossos dados. Os resultados

²⁶ É importante esclarecer que Fernandes (1996) realizou essa análise somente com os substantivos, adjetivos e categorias substantivadas, pois seus resultados mostraram-se mais significativos em relação à primeira análise efetuada com todos os dados. No estudo de Scherre (1988), para esta variável, os resultados permaneceram inalterados, nas duas análises: a primeira com todos os dados e a segunda, somente com os substantivos, adjetivos e categorias substantivadas.

das autoras apresentaram maiores semelhanças: Scherre e Fernandes (*oxítonos e monossílabos tônicos*: 0.66 e 0.58, respectivamente); (*proparoxítonos*: 0.44 e 0.47, respectivamente) e (*paroxítonos*: 0.39 e 0.48, respectivamente). Já em Guarapuava, obtivemos para esses fatores pesos de 0.21, 0.15 e 0.24, respectivamente, o que parece indicar que a *ausência de CN* é bem mais significativa em nossos dados.

Segundo Scherre (1988), há uma relação de dependência entre as variáveis *saliência fônica* e *tonicidade* e ao analisar essas variáveis individualmente, a autora verificou que ocorria uma sobreposição entre ambas, o que causava diferenças probabilísticas nos resultados, com destaque para os itens regulares e não regulares.

Neste estudo, de forma semelhante à análise da autora, decidimos refazer a rodada e redefinir as variáveis de nossa análise, ou seja, iremos aliar (analisar em um mesmo grupo de fatores) as variáveis *saliência fônica* e *tonicidade*, “ignorando a relação oxítono e paroxítono para os itens não regulares e mantendo-a, acrescida do fator proparoxítono, para os itens regulares” (Scherre, 1988, p. 136). Desse modo, consideramos em nossa análise os mesmos fatores da autora:

1. *Plural duplo* (oxítono ou paroxítono).

(57) No caso, eu era o segundo mais velho, daí tinha *dois mais novo* né. (VARLINGUA Inf6 M1c).

2. *Itens terminados em -l* (oxítono ou paroxítono).

(58) Como eu não fui bem atendida na minha cidade aqui, *nus hospital* aqui da minha cidade, eu procurei um hospital fora. (VARLINGUA Inf5 F1c).

3. *Itens terminados em -r* (oxítono ou paroxítono).

(59) As *mulher* tão se desvalorizando demais porque elas não tem noção do que elas conseguiram. (VARLINGUA Inf11 F2c).

4. *Itens terminados em -ão* (oxítono irregular).

(60) Graças a Deus eu, tenho *minhas devoção*, saio tranquilo. (VARLINGUA Inf20 M2p).

5. *Itens terminados em -s* (oxítono).

(61) Feiz *dois mesis* onte, que ele feis a cirurgia. (VARLINGUA Inf3 F1g).

6. Regular oxítono.

(62) Uma minina de oito anus, agora *as mães* não dexum nem chegá perto do fogão. (VARLINGUA Inf9 F2g).

7. Regular paroxítono.

(63) Meu pai era um hóme muito assim, como é que eu te digo, amigo, *dos filho* sabe. (VARLINGUA Inf21 F2g).

8. Regular proparoxítono.

(64) Quando eu fiz essa cirurgia *us médico* me derum cinco ano só de vida. (VARLINGUA Inf8 M2p).

Desse modo, assim como Scherre (1988), a nossa hipótese para as duas variáveis aliadas é que os elementos mais salientes favoreçam a *presença de CN* na língua falada em Guarapuava. Diante desses resultados, com o objetivo de evitar uma sobreposição dos fatores *monossílabos átonos* com o *artigo* analisado na variável *classe gramatical do elemento*, retiramos os *monossílabos átonos* e efetuamos uma nova e última rodada contemplando as variáveis *saliência fônica* e *tonicidade* aliadas. Os resultados seguem na tabela 16:

TABELA 16: Resultados da saliência e tonicidade aliadas na presença de CN

Fatores	Aplicação/Total	%	P.R.
Plural duplo Ex: us avós dele (VARLINGUA Inf5F1c).	15/19	79%	0.92
Itens terminados em -l Ex: seis hospitais (VARLINGUA Inf18 M1c).	26/51	51%	0.72
Itens terminados em -s Ex: duas vez (VARLINGUA Inf21 F2g).	47/84	56%	0.65
Itens terminados em -r Ex: cos menores (VARLINGUA Inf12M2).	41/97	42%	0.61
Regulares oxítonos Ex: dois irmãos (VARLINGUA Inf21 F2g).	46/112	41%	0.59
Regulares proparoxítonos Ex: alguns médicos (VARLINGUA Inf13 F1p).	21/67	31%	0.56
Itens terminados em -ão Ex: todas as função . (VARLINGUA Inf4 M1g).	16/61	26%	0.51
Regulares paroxítonos Ex: as amiguinha (VARLINGUA Inf22 M2g).	1106/2477	45%	0.47
Total	1318/2968	44%	

Fonte: Dados do VARLINGUA 2015

Os resultados da tabela acima para a *presença de CN* apresentaram peso relativo de 0.92 para o fator *plural duplo* e um peso relativo de 0.72 para os *itens terminados em -l*, seguidos dos *itens terminados em -s* (0.65), itens que mais favoreceram a concordância em nossa amostra.

Na sequência, também favorecendo a concordância estão os *itens terminados em -r* (0.61), os *regulares oxítonos* (0.59), e os *regulares proparoxítonos* (0.56). Logo abaixo, com resultado próximo ao ponto neutro, estão os *itens terminados em -ão* (0.51). Por último, na escala da *saliência fônica* estão os *regulares paroxítonos* (0.47) com resultado desfavorável à aplicação de concordância.

Conforme proposto por Scherre (1988), realizamos também uma análise da *saliência e tonicidade aliadas* considerando somente os *substantivos, adjetivos e categorias substantivadas*, ou seja, retirando os determinantes (*artigos, possessivos, demonstrativos, indefinidos e quantificadores*), a fim de observarmos se os resultados atribuídos aos diferentes fatores sofrem alterações.

Como afirma a autora, a “nossa intuição de linguista nos diz que a segunda análise condiz melhor com a realidade dos fatos, porque a variável Processos²⁷ é a que se aplica especificamente ao fenômeno da concordância.” (SCHERRE, 1988, p. 85). Para Scherre, essa variável refere-se aos processos morfofonológicos de formação de plural, que atende ao princípio da *saliência fônica*, no qual os itens mais salientes tendem a favorecer mais marcas de plural.

Nessa análise, obtivemos um total de 2422 ocorrências, 782 ocorrências (32%) com *presença de CN* e 1640 (68%) com *ausência de CN*. A tabela abaixo apresenta os resultados da *saliência e tonicidade aliadas* na rodada sem os determinantes.

²⁷ Processos refere-se aos processos de formação de plural, relacionado à variável *saliência fônica*.

TABELA 17: Resultados da saliência e tonicidade aliadas - Substantivos, Adjetivos e Categorias Substantivadas

Fatores	Aplicação/Total	%	P.R.
Plural duplo	15/19	79%	0.89
Itens terminados em –s	47/84	56%	0.75
Itens terminados em –l	26/51	51%	0.69
Itens terminados em –r	41/97	42%	0.58
Regulares oxítonos	44/110	40%	0.58
Regulares proparoxítonos	21/67	31%	0.49
Regulares paroxítonos	572/1933	30%	0.47
Itens terminados em –ão	16/61	26%	0.44
Total <i>Input: 0.31</i>	782/2422	32%	

Fonte: Dados do VARLINGUA 2015

Os resultados obtidos para o *plural duplo*, apontado por Scherre (1988) como o item mais saliente (0.89) se mostrou favorável à marcação de plural em nossos dados, seguido dos *itens terminados em –s* (0.75) e dos *itens terminados em –l* (0.69).

Também favoreceram a aplicação da regra de concordância os *itens terminados em –r* e os itens *regulares oxítonos* (0.58), com o mesmo peso relativo. Já os *regulares proparoxítonos* desfavoreceram a CN (0.49), com um resultado próximo do ponto neutro, assim com os *regulares paroxítonos* (0.47) e os *itens terminados em –ão* (0.44).

Nos resultados de Scherre (1988, p. 139) favoreceram a *presença de CN*: o *plural duplo* (0.80), os *itens terminados em –l* (0.69), os *itens terminados em –r* (0.65), os *itens terminados em –ão* (0.59) e os *itens terminados em –s* (0.56). Já os itens regulares desfavoreceram a CN: os *regulares oxítonos* (0.38), os *regulares proparoxítonos* (0.21) e os *regulares paroxítonos* (0.17).

Do mesmo modo, Fernandes (1996, p. 88) também realizou essa análise. Dos resultados obtidos pela pesquisadora, o *plural duplo* (0.81), os *itens terminados em –l* (0.77), os *itens terminados em –ão* (0.74), os *itens terminados em –r* (0.73), os *itens terminados em –s* (0.58) e os *regulares oxítonos* (0.54) favoreceram a *presença de CN*. Já os *regulares paroxítonos* (0.46) e os *regulares proparoxítonos* (0.45) desfavoreceram a aplicação da regra nos dados de Fernandes.

Comparando os nossos resultados aos de Scherre²⁸ (1988) e Fernandes (1996), embora não na mesma escala da *saliência fônica*, podemos verificar que os resultados de vários fatores apresentaram convergência com a análise das pesquisadoras. Os itens considerados os mais salientes pelas autoras como o *plural duplo*, os *itens terminados em -l*, os *itens terminados em -r* e os *itens terminados em -s*, também favoreceram a CN em nossa amostra.

Outra semelhança entre esses resultados com o de Guarapuava estão nos regulares *proparoxítonos* e *paroxítonos*, que desfavoreceram a CN. No entanto, verificamos algumas diferenças entre os três resultados: nos *regulares oxítonos*, que favoreceram a CN em Guarapuava e no estudo de Fernandes, nos dados de Scherre, desfavoreceram a concordância. E os *itens terminados em -ão*, que desfavoreceram a aplicação da regra em nossos dados, nos dados de Scherre e Fernandes favoreceram a CN.

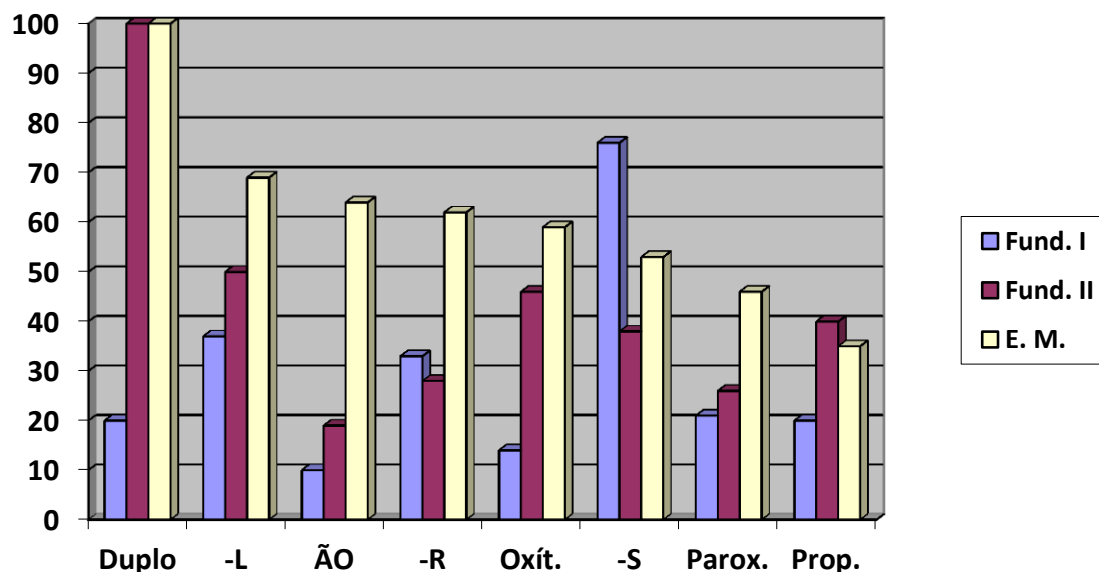
Assim, a nossa hipótese para a variável *saliência e tonicidade aliadas* foi parcialmente confirmada em Guarapuava, pois o *plural duplo*, os *itens terminados em -s*, os *itens terminados em -l* e os *itens terminados em -r*, considerados os mais salientes, favoreceram a CN. Já os *itens terminados em -ão*, também tidos como um dos mais salientes por Scherre, não favoreceram a concordância em nossos dados.

Verificamos, ainda, que os resultados para esta variável permaneceram praticamente inalterados, nas duas análises: a primeira com todos os dados e a segunda, somente com os *substantivos, adjetivos e categorias substantivadas*.

Para analisarmos a CN mais detalhadamente, realizamos também o cruzamento da variável *saliência e tonicidade aliadas* com a variável *escolaridade*. O objetivo desse cruzamento é verificar se a CN nos itens mais e menos salientes está relacionada ao nível de escolaridade do falante. Os resultados obtidos são apresentados a seguir:

²⁸ Estes resultados de Scherre foram obtidos da análise somente com os substantivos, adjetivos e categorias substantivadas.

GRÁFICO 1: Cruzamento da saliência e tonicidade aliadas com a escolaridade – Substantivos, adjetivos e categorias substantivadas



Fonte: Dados do VARLINGUA 2015

Como podemos visualizar no gráfico acima, conforme o esperado, em praticamente todos os níveis de saliência fônica, os falantes com *ensino médio* foram os que mais realizaram a CN, especialmente nos itens mais salientes (com exceção dos *itens terminados em -s*): *plural duplo* (100% de CN), *itens terminados em -l* (69%), *itens terminados em -ão* (64%), *itens terminados em -r* (62%) e os *itens terminados em -s* (53%). Nos itens regulares, os falantes com *ensino médio* usaram 59% de concordância nos *oxítonos*, 46% nos *paroxítonos* e 35% nos *proparoxítonos*.

Para os falantes com *fundamental II*, o resultado da CN no *plural duplo* também foi categórico (100%) e os *itens terminados em -l* apresentaram 50% de concordância. Em todos os demais níveis de saliência predominou a *ausência de CN*: *itens terminados em -ão* (19%), *itens terminados em -r* (28%), *itens terminados em -s* (38%). Nos itens regulares, a concordância foi de 46% nos *oxítonos*, 40% nos *paroxítonos* e 26% nos *proparoxítonos*.

Quanto aos falantes com *fundamental I*, com exceção dos *itens terminados em -s*, com 76% de CN, a queda nos índices de CN foi mais acentuada, com percentagens ainda menores: *plural duplo* (20%), os *itens terminados em -l* (37%), os *itens terminados em -r* (33%) e os *itens terminados em -ão* (10%). Os itens regulares apresentaram 21% de concordância nos *paroxítonos*, 20% nos *proparoxítonos* e 14% nos *oxítonos*.

Como podemos observar, os falantes com *ensino médio* tendem a realizar a CN com maior frequência que os falantes com *fundamental II* e *fundamental I*, pois apresentam percentuais mais elevados de concordância em praticamente todos os fatores, especialmente os mais salientes.

Esse resultado nos permite dizer que a *escolaridade* é um fator social relevante no uso da CN, visto que os falantes mais escolarizados tendem a usar mais marcas de plural na língua falada nos diferentes níveis de formação de plural que os falantes menos escolarizados. Em nossa rodada final dos dados de Guarapuava, dos fatores sociais, a *escolaridade* foi selecionada pelo programa estatístico como a variável mais significativa.

A seguir, são apresentados os resultados das variáveis sociais na análise da CN em Guarapuava.

4.3. Resultados das variáveis sociais em Guarapuava

Apresentamos, a seguir, os resultados das variáveis sociais *escolaridade* e *sexo*, selecionadas em segunda e quinta posição na análise da CN em Guarapuava. A variável *faixa etária* não foi selecionada pelo programa estatístico GoldVarbX.

4.3.1. Escolaridade

A *escolaridade* foi selecionada como a segunda variável mais significativa pelo programa estatístico GoldVarbX nos dados de Guarapuava. Assim como atesta a maioria dos estudos realizados no Brasil acerca do fenômeno da variação na CN de número, a nossa hipótese é que a realização da regra de concordância será maior na fala dos informantes mais escolarizados.

Seguem os resultados, na tabela abaixo:

TABELA 18: Resultados da escolaridade na presença de CN

Fatores	Aplicação/Total	%	P.R.
c. Ensino médio	791/1141	69%	0.68
g. Fundamental II	1009/1723	59%	0.46
p. Fundamental I	729/1349	54%	0.38
Total	2529/4213	60%	

Fonte: Dados do VARLINGUA 2015

Os resultados apresentados neste estudo para a *escolaridade* mostraram que os falantes de Guarapuava mais escolarizados realizam a CN com maior frequência que os falantes menos escolarizados, confirmando nossa hipótese. Os falantes com *ensino médio* apresentaram peso relativo de 0.68, favorecedor da concordância.

Em seguida, para os falantes com *fundamental II*, o peso relativo (0.46) desfavoreceu a concordância, com um resultado que apontou para o ponto da neutralidade, ou seja, esses falantes quase permaneceram entre as duas variáveis, *presença de CN* e *ausência de CN*. Já os falantes menos escolarizados, com *fundamental I*, apresentaram um peso relativo de (0.38), o que indica um desfavorecimento da aplicação de concordância, conforme o esperado.

Comparando os nossos resultados com os de Fernandes (1996, p. 106): *ensino médio* (0.68 e 0.66, respectivamente), *fundamental II* (0.46 e 0.48, respectivamente) e *fundamental I* (0.38 e 0.32, respectivamente), observamos uma significativa proximidade nos resultados, os quais apontam que a aplicação da CN predomina nos falantes com *ensino médio*, em detrimento aos falantes menos escolarizados, com *fundamental II* e *fundamental I*. Como argumenta Scherre:

Inquestionavelmente, as pessoas mais escolarizadas, mais sensíveis às marcas de prestígio e que exercem profissões de trato público tendem a fazer mais concordância e, se não a fazem, são criticadas por nós, que também deixamos de fazer concordâncias verbais e nominais, de forma regular, quer queiramos ou não, quer reconhecamos ou não. (SCHERRE, 2002, p. 236).

Pelas palavras de Scherre (2002) e pelos resultados obtidos em nossos dados, podemos observar a importância do fator *escolaridade* na análise da variável CN, visto que, em alguns contextos, as pessoas deixam de realizar a concordância na língua falada, todavia, quanto maior o nível de escolaridade, maior é a frequência de uso da forma padrão.

4.3.2. Sexo

A variável sexo foi a quinta selecionada em nossa amostra. De acordo com Labov ([1972] 2008), as mulheres são mais sensíveis à forma padrão, na medida em que percebem essa forma associada ao prestígio social. Em relação à concordância, estudos anteriores (Scherre, 1988; 1998), mostram que as mulheres utilizam com maior frequência a concordância que os homens.

Os resultados são apresentados na tabela 19, a seguir:

TABELA 19: Resultados do sexo na presença de CN

Fatores	Aplicação/Total	%	P.R.
Feminino	1321/2125	62%	0.53
Masculino	1208/2088	58%	0.46
Total	2529/4213	60%	

Fonte: Dados do VARLINGUA 2015

Embora nossos resultados da variável *sexo* encontrem-se próximos ao ponto neutro, conforme a tabela acima, podemos observar que os falantes do *sexo feminino* utilizam mais a regra de concordância (0.53) que os falantes do *sexo masculino* (0.46) em Guarapuava, confirmando a nossa hipótese.

Os resultados de Fernandes (1996, p. 114) apresentaram os mesmos pesos relativos que os obtidos em nossos dados, tanto para os falantes do *sexo feminino* (0.53), quanto para os do *sexo masculino* (0.46). Diante disso, esses resultados vão de encontro com a postulação laboviana acerca dessa variável.

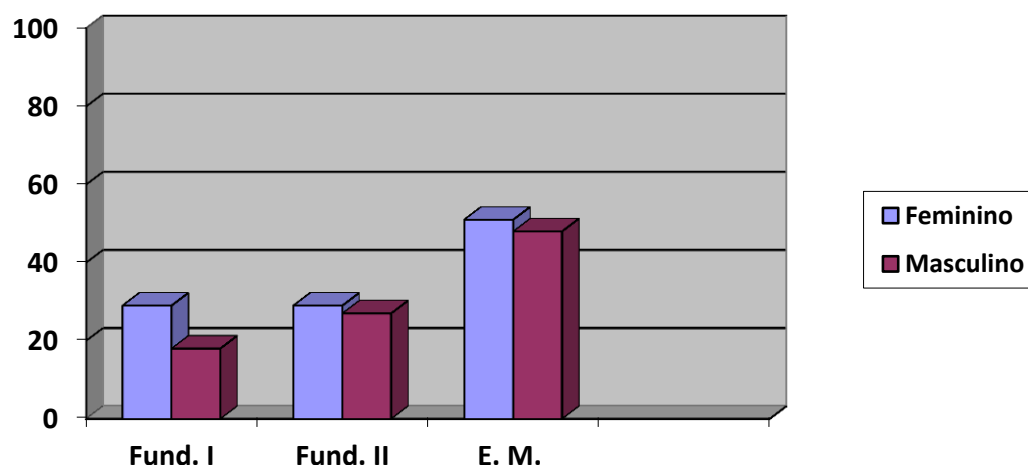
No entanto, para obtermos informações mais detalhadas sobre a influência dos fatores sociais no uso da concordância, serão realizados alguns cruzamentos entre as variáveis sociais: *sexo*, *faixa etária* e *escolaridade*. Os resultados dos cruzamentos são apresentados na próxima seção.

4.4. Resultados dos cruzamentos das variáveis sociais em Guarapuava

Para analisarmos melhor a influência dos fatores sociais no condicionamento da língua falada em Guarapuava, realizamos alguns cruzamentos entre eles, que são: 1) *escolaridade* e *sexo*; 2) *escolaridade* e *faixa etária*; e 3) *faixa etária* e *sexo*. A seguir, são apresentados os resultados.

4.4.1. Cruzamento: escolaridade e sexo

Esse cruzamento foi realizado com o objetivo de verificarmos mais detalhadamente o uso da CN na língua falada de homens e mulheres, de diferentes níveis de *escolaridade*, em nossa amostra. Os resultados obtidos seguem abaixo:

GRÁFICO 2: Escolaridade e sexo

Fonte: Dados do VARLINGUA 2015

Pelos resultados do gráfico 2, referente ao cruzamento das variáveis *escolaridade* e *sexo*, podemos verificar que as mulheres de Guarapuava com *ensino médio* utilizam com maior frequência (51%) a CN na língua falada do que as mulheres com níveis *fundamental II* e *fundamental I*, com a mesma percentagem (29%).

O resultado para os homens assemelham-se ao das mulheres, ou seja, os homens com *ensino médio* (48%) utilizam mais a regra de concordância que os homens com *fundamental II* e *fundamental I* (27% e 18%, respectivamente). Considerando, no entanto, os níveis fundamental I e II, nota-se que as mulheres apresentam o mesmo resultado (29%), com um aumento da CN somente no *ensino médio* (51%). Já na fala dos homens há uma diferença no uso da concordância em todos os níveis de ensino (*fundamental I*: 18%, *fundamental II*: 27%, *ensino médio*: 48%), ou seja, à medida que aumenta o nível de *escolaridade*, há também um aumento no uso da CN.

Os resultados demonstram ainda que a *ausência de CN* seja uma tendência na fala dos homens e das mulheres com nível *fundamental I* e *II* em nossa amostra. No *ensino médio*, mulheres e homens utilizam a ausência e a presença da CN praticamente na mesma proporção (51% e 48%, respectivamente), com uma leve diferença de uso entre os sexos.

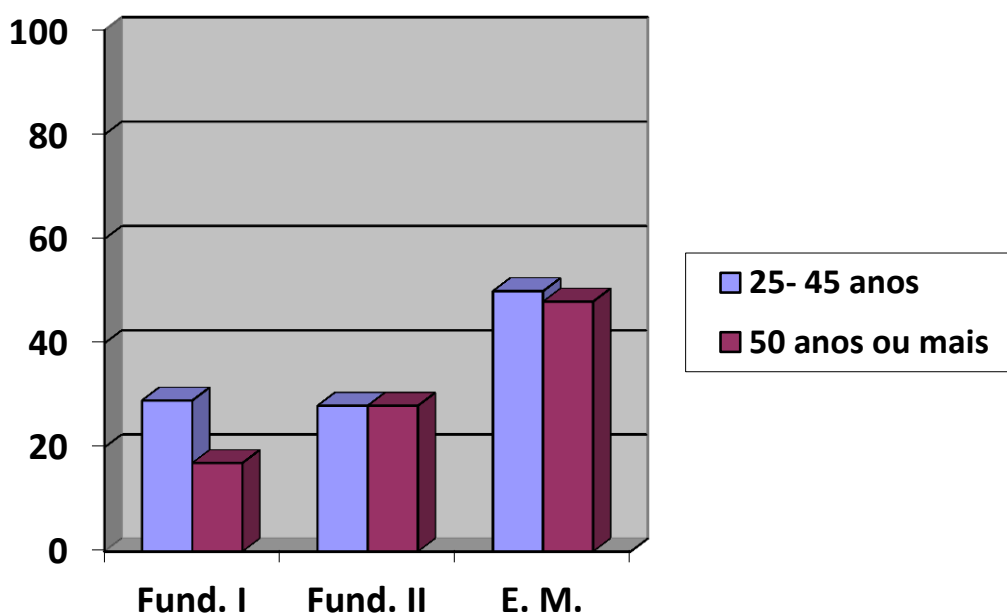
Apesar da pouca probabilidade, o percentual de uso da regra de concordância predominou nos falantes do *sexo feminino*. Como podemos perceber, esses resultados do cruzamento das variáveis *escolaridade* e *sexo* corroboraram os resultados apresentados na rodada final dos dados de Guarapuava, com pesos relativos, nos quais as mulheres utilizam mais a concordância que os homens, assim como os falantes do nível *ensino médio*.

4.4.2. Cruzamento: Escolaridade e faixa etária

A variável *faixa etária* não foi selecionada pelo programa estatístico GoldVarbX na rodada final dos dados de Guarapuava, porém, com o objetivo de verificar a importância dessa variável na língua falada dessa comunidade, realizamos o cruzamento da *faixa etária* com a *escolaridade*.

Seguem os resultados no gráfico 3:

GRÁFICO 3: Escolaridade e faixa etária



Fonte: Dados do VARLINGUA 2015

Os resultados apresentados do cruzamento das variáveis *escolaridade* e *faixa etária* indicaram que 50% da *faixa etária* mais jovem e com *ensino médio* utiliza a CN e 50% não a utiliza em Guarapuava. Os resultados para a mesma faixa etária com níveis *fundamental II* e *fundamental I* (28% e 29%, respectivamente) indicaram que a *ausência de CN* predomina entre os falantes mais jovens de Guarapuava.

Com relação à *faixa etária* de *50 anos ou mais*, os resultados mostraram que a *ausência de CN* predomina em todos os níveis de escolaridade, *ensino médio* (48%), *fundamental II* (28%) e *fundamental I* (17%).

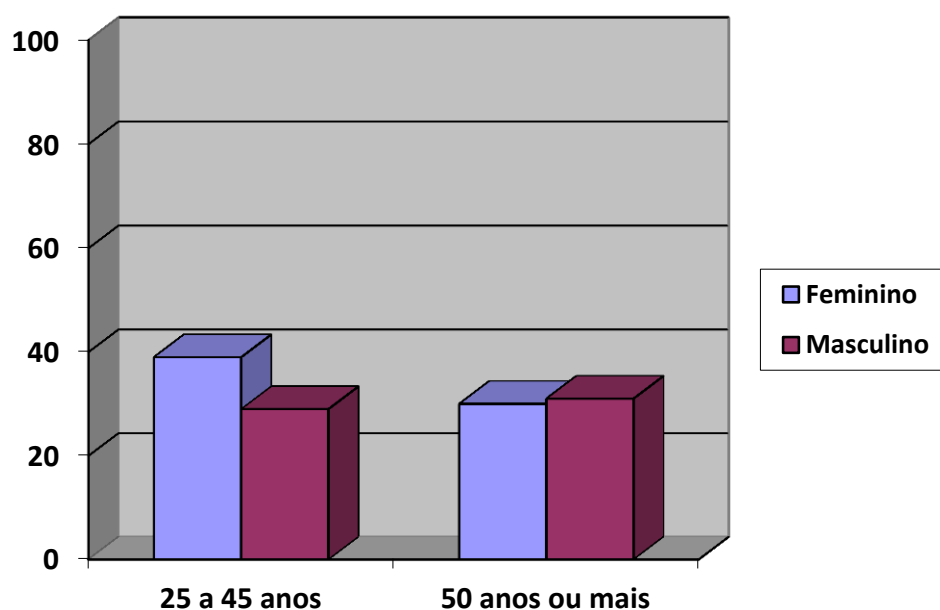
Logo, observando as duas faixas etárias, a de *25 a 45 anos* e a de *50 anos ou mais*, podemos verificar que o resultado percentual indica que tanto os falantes mais jovens quanto os falantes mais velhos, com *ensino fundamental I e II* fazem pouco uso da CN na língua falada. Os falantes com *ensino médio*, das duas faixas etárias utilizam a *presença* e a *ausência* de concordância praticamente na mesma proporção.

Esse cruzamento parece comprovar que a *escolaridade* é que condiciona a CN, pois os resultados para as duas faixas etárias são semelhantes, com exceção do *ensino fundamental I*, em que os falantes mais velhos usam menos concordância que os mais jovens.

4.4.3. Cruzamento: Faixa etária e sexo

Esse cruzamento foi realizado com o objetivo de verificarmos o uso da CN na fala de homens e mulheres de diferentes faixas etárias em nossa amostra. Os resultados são apresentados na sequência:

GRÁFICO 4: Faixa etária e sexo



Fonte: Dados do VARLINGUA 2015

Pelo cruzamento das variáveis *faixa etária* e *sexo*, verificamos que as mulheres de Guarapuava da *faixa etária* mais jovem e da *faixa etária* mais velha utilizam com pouca frequência a CN (39% e 30%, respectivamente). O resultado para os homens assemelhou-se

ao das mulheres. Os homens da *faixa etária* de 25 a 45 anos obtiveram 29% de *presença de CN* na língua falada e os homens de 50 anos ou mais apresentaram 31% de concordância.

Desse modo, podemos observar que mesmo as percentagens sendo próximas, as mulheres mais jovens (39%) utilizam mais a concordância que as mulheres mais velhas (30%). De forma contrária, mas com pouquíssima diferença nos percentuais, os homens mais velhos de Guarapuava utilizam mais a CN que os homens mais jovens (31% e 29%, respectivamente).

Por meio dos resultados apresentados acima, com todos os cruzamentos das variáveis sociais, podemos constatar que a *presença de CN*, apesar da baixa percentagem de uso, predomina no *ensino médio* e no *sexo feminino*, resultados que corroboram os obtidos em peso relativo.

Na sequência, são apresentadas as considerações finais de nosso estudo sobre a CN em Guarapuava.

CONSIDERAÇÕES FINAIS

O principal objetivo deste estudo foi analisar o uso da CN constituída pelas formas *presença de CN* e *ausência de CN*, entre os elementos do SN, na comunidade de Guarapuava-PR. Para a análise dos dados, foram consideradas 24 entrevistas do banco VARLINGUA, distribuídas por sexo, faixa etária e escolaridade e os resultados foram obtidos por meio do programa estatístico GoldVarbX.

As variáveis consideradas foram: 1. Saliência fônica; 2. Tonicidade; 3. Posição do elemento no SN; 4. Classe gramatical do elemento; 5. Marcas precedentes; 6. Sexo; 7. Faixa etária; e 8. Escolaridade.

Primeiramente, efetuamos duas rodadas estatísticas no Programa GolVarbX, com algumas variáveis isoladas, são elas: saliência fônica e tonicidade, posição do elemento no SN e classe gramatical. Após verificarmos que essas variáveis seriam mais significativas se fossem aliadas, isto é, analisadas num mesmo grupo de fatores, tomando por base o estudo de Scherre (1988) e Fernandes (1996), decidimos refazer as rodadas.

Efetuamos mais duas rodadas com a variável *saliência e tonicidade aliadas*, e com a variável *posição e classe gramatical aliadas*. Definidas as variáveis, por último, realizamos a rodada final com as variáveis linguísticas e as variáveis sociais, incluindo todos os dados.

As variáveis selecionadas pelo programa estatístico GoldVarbX foram, nesta ordem de significância: 1. *posição e classe gramatical aliadas*; 2. *escolaridade*; 3. *marcas precedentes*; 4. *saliência e tonicidade aliadas*; e 5. *sexo*. Por meio da análise dessas variáveis pudemos observar a atuação de diferentes fatores no condicionamento da variação do fenômeno da CN.

Na análise da variável *posição e classe gramatical aliadas*, os fatores linguísticos que favoreceram a *presença de CN* foram: os *determinantes* (artigos, possessivos, indefinidos e quantificadores) *antepostos ao núcleo na 2ª posição*, pois apresentaram um peso relativo quase categórico para a *presença de CN* (0,98). Em seguida, seguem como favorecedores da concordância os *determinantes antepostos ao núcleo na 1ª posição* (0,73) e a categoria substantivada como núcleo na 2ª posição com 0,57. Já a *categoria substantivada em 3ª ou demais posições* e os *substantivos e adjetivos*, em todas as posições, desfavoreceram a CN.

A nossa hipótese para a variável *posição e classe gramatical aliadas* se confirmou nos dados de Guarapuava, pois os determinantes, geralmente em *primeira posição*, favoreceram a *presença de CN*, enquanto as demais classes gramaticais, *substantivos, adjetivos e categorias substantivadas*, na maioria das vezes em segunda ou demais posições, a desfavoreceram.

A variável *marcas precedentes* apontou como fatores favorecedores da CN a *ausência de elemento precedente* (0,82) e a *ausência de marca morfológica no elemento precedente* (0,64). O resultado do fator *ausência de elemento precedente* indica que o primeiro elemento do SN geralmente é um termo marcado por não ter um elemento anterior a ele. Por consequência, quando há *ausência de marca morfológica no elemento precedente*, geralmente o elemento em análise apresenta a marca de plural, em segunda posição. Todos os outros fatores analisados desfavoreceram a CN em Guarapuava.

Verificamos, portanto, que os nossos resultados corroboram parcialmente os de Scherre (1988) e Fernandes (1996), pois o fator *ausência de marca morfológica no elemento precedente* em Guarapuava favoreceu a CN, e nos dados de Scherre (1988) e Fernandes (1996), constatamos o uso categórico da CN para esse fator.

Também o fator *ausência de elemento precedente* em Guarapuava e nos de Scherre favoreceu a CN. Já os fatores *duas ou mais marcas (em -s) precedente e elemento precedente é um numeral* em nossos dados desfavoreceram a CN, enquanto que nos dados de Scherre e Fernandes a favoreceram. Quanto aos demais fatores, tanto em Guarapuava como nos de Scherre e Fernandes desfavoreceram a CN.

Em relação à variável *saliência e tonicidade aliadas*, o *plural duplo* (0,92), os *itens terminados em -l* (0,72), seguidos dos *itens terminados em -s* (0,65), foram os itens que mais favoreceram a concordância em nossa amostra. Na escala da *saliência fônica*, esses itens são considerados os mais salientes, o que, segundo Scherre (1988), explicaria o maior uso da marca de plural.

A nossa hipótese para a variável *saliência e tonicidade aliadas* foi parcialmente confirmada em nossos dados, pois o *plural duplo*, os *itens terminados em -l* e os *itens terminados em -s*, considerados os mais salientes, favoreceram a CN. Já os *itens terminados em -ão*, também um dos mais salientes, não favoreceram a CN em Guarapuava.

Dentre os fatores sociais, destaca-se a influência da *escolaridade*, selecionada em segunda posição. Conforme já esperado, o aumento da *escolaridade* leva a um maior uso da concordância. Os informantes com *ensino médio* favoreceram a *presença de CN* (0,68), aqueles com *fundamental II* apresentaram um uso próximo ao ponto neutro (0,46) e no nível *fundamental I* os falantes desfavoreceram a CN (0,38).

Os falantes de Guarapuava com *ensino médio* foram os que mais utilizaram a CN, isto é, os mais escolarizados, confirmando a nossa hipótese. Verificamos, assim, que o aumento da concordância é proporcional ao aumento da *escolaridade*. Diante disso, podemos constatar que a *escolaridade* é um fator relevante para a realização da regra de CN.

Em relação ao *sexo*, última variável selecionada, verificamos que as mulheres, geralmente mais sensíveis à forma padrão, apresentaram maior CN (0,53) do que os homens (0,46) em Guarapuava. Nota-se, no entanto, que os pesos relativos encontram-se próximos ao ponto neutro e bastante próximos entre si.

A *presença de CN* predominou nos falantes do *sexo feminino*, em comparação aos do *sexo masculino*, confirmando a nossa hipótese, apesar do resultado próximo ao ponto neutro. Por essa razão, acreditamos que esta variável não está fazendo a diferença na análise efetuada e, por isso, foi a última variável selecionada pelo GoldVarbX.

Em relação à *faixa etária*, verificamos que essa variável não foi selecionada pelo programa estatístico, o que indica que a idade dos falantes não está condicionando o uso da CN em nossos dados. Assim, não confirmamos a nossa hipótese no tocante à questão da mudança em progresso em Guarapuava.

Os resultados gerais da análise totalizaram 4.213 ocorrências, das quais foram registrados 1.684 casos de *ausência de CN* (40%) e 2.529 casos de *presença de CN* (60%). No entanto, considerando os constituintes do SN, verificamos que os determinantes apresentam a marca de plural em quase 100% das ocorrências. Já nos substantivos, categorias substantivadas e adjetivos a CN é de somente 32%, ou seja, há uma elevada probabilidade de *ausência de CN* (68%).

A partir dos resultados obtidos com todos os dados da amostra, ao observarmos a atuação dos diferentes fatores, verificamos que a aplicação da CN ocorre principalmente nos determinantes, elementos que apresentam a marcação de plural em quase 100% das ocorrências, mas que não indicam necessariamente a ‘concordância’ no SN.

Por isso, realizamos uma análise somente com os *substantivos*, *adjetivos* e *categorias substantivadas*, e nessa rodada obtivemos um *input* de 0,31 e uma percentagem de 32% para a *presença de CN*. Os resultados dessa rodada mostraram, portanto, uma reduzida aplicação da regra de CN nos dados de Guarapuava.

Como em nosso estudo procuramos seguir as postulações de Scherre (1988), testamos os mesmos fatores que a autora, e constatamos que os fatores que condicionam ou não a CN não se diferenciam pela região, mas, sim, pela característica própria do PB, pois, a maioria dos fatores que condicionaram a *presença de CN* em nosso estudo também apresentaram a aplicação da regra nos estudos de Scherre (1988).

Portanto, por meio desses resultados, pudemos confirmar a nossa hipótese geral, que os falantes de Guarapuava utilizam com maior frequência a *ausência de CN* na língua falada. Além disso, o perfil do falante dessa comunidade que mais utiliza a regra de CN é o do sexo

feminino, da faixa etária mais jovem e com o *ensino médio*, logo, o falante que possui mais escolaridade. Assim, com esta pesquisa, esperamos contribuir para a descrição da CN de número plural e para os estudos de variação linguística na região Sul do Brasil.

REFERÊNCIAS

- BRAGA, Maria Luiza. (1977). *A concordância de número no sintagma nominal no triângulo mineiro*. PUC, Rio de Janeiro, 88p. Dissertação de mestrado, inédito.
- CALVET, Louis Jean. *Sociolinguística: uma introdução crítica*. São Paulo: Parábola, 2002.
- CEGALLA, Domingos Paschoal. *Novíssima gramática da língua portuguesa*. 48. ed. São Paulo: Companhia Editora Nacional, 2008.
- COELHO, Izete L. et al. *Para conhecer sociolinguística*. São Paulo: Contexto, 2015.
- FERNANDES, Marisa. *Concordância nominal na região sul*. Dissertação de mestrado. Florianópolis: Universidade Federal de Santa Catarina, 1996.
- GUY, Gregory Riordan; ZILLES, Ana. *Sociolinguística Quantitativa – instrumental de análise*. São Paulo: Parábola Editorial, 2007.
- IZIDORO, Heitor Francisco. *Guarapuava: das sesmarias a Itaipu*. Curitiba: Vicentina, 1976.
- LABOV, William. *Padrões sociolinguísticos*. Tradução de Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008. 392 p. Tradução de: Sociolinguistic patterns.
- MARCONDES, Gracita Gruber. *Duzentos anos de uma caminhada histórica: 1810-2010*. Guarapuava: O Autor, 2010.
- MARIANO, Mara Pereira. *O fenômeno da concordância nominal em redações escolares*. Dissertação de mestrado. Rio de Janeiro: UFRJ/FL, 2013.
- MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). *Introdução à sociolinguística: o tratamento da variação*. 4. ed. São Paulo: Contexto, 2013.
- PETTER, Margarida. Linguagem, língua, linguística. In: FIORIN, José Luiz (org.). *Introdução à Linguística – I. Objetos Teóricos*. São Paulo: Contexto, 2011, p. 11-24.
- PIETROFORTE, Antonio Vicente. A língua como objeto da linguística. In: FIORIN, José Luiz (org.). *Introdução à Linguística – I. Objetos Teóricos*. São Paulo: Contexto, 2011, p. 75-94.
- SAUSSURE, Ferdinand. *Curso de lingüística geral*. São Paulo: Cultrix, 2008.
- SCHERRE, Maria Marta Pereira. *Reanálise da concordância nominal em português*. Tese de Doutorado em Linguística. Rio de Janeiro, Faculdade de Letras, UFRJ, 1988. Em dois volumes, com 555p. mimeo.

_____. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa* (RILP)- Norma e Variação do Português. Associação das Universidades de Língua Portuguesa. 12: 37-49. dez. de 1994.

_____. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: OLIVEIRA E SILVA, SCHERRE, M. M. P. (orgs.) In: *Padrões sociolinguísticos*. Análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo brasileiro, 1998.

_____. A norma do imperativo e o imperativo da norma – uma reflexão sociolinguística sobre o conceito do erro. In: BAGNO, Marcos. *Linguística da Norma*. [Org] São Paulo, SP: Loyola, 2002.

TARALLO, Fernando. *A pesquisa sociolinguística*. São Paulo: Editora Ática, 2001.

VIEIRA, S. R; BRANDÃO, S. F. *Ensino de gramática: descrição e uso*. São Paulo: Contexto, 2007.

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

ANEXOS

ANEXO I**FICHA SOCIAL**

Nome:

Endereço:

Sexo:

Idade:

Etnia:

Profissão:

Fala quantas línguas? Quais? Como as aprendeu?

Profissão dos pais:

Pai:

Mãe:

Origem/etnia dos pais:

Pai:

Mãe:

Escolaridade – estuda e/ou até que série estudou:

Local de nascimento:

Se não nasceu em _____, com que idade se mudou para essa cidade?

Qual o passatempo favorito?

Sobre quais assuntos mais gosta de conversar?

O que costuma fazer nos dias de folga?

Dia e horário disponíveis para a entrevista:

ANEXO II**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO – TCLE**

O banco de dados linguísticos Variação Linguística de Guarapuava – VARLINGUA, da UNICENTRO, possui como objetivo registrar em áudio a língua falada por falantes nascidos, criados e residentes em Guarapuava, Paraná, e é coordenado pelas professoras Loremi Loregian-Penkhal e Lucelene Franceschini. O áudio será posteriormente utilizado para pesquisas e descrições (socio)linguísticas diversas. Os dados analisados poderão ser divulgados por meio de literatura especializada e eventos científicos e os sujeitos não serão identificados. Esta pesquisa possui caráter acadêmico e não representa qualquer dolo para o informante que participa voluntariamente com a gravação de sua fala.

Assim, eu, _____ portador(a) da cédula de identidade nº _____ e do CPF nº _____ estou ciente e de acordo com os termos da realização desta pesquisa. Dessa forma, aceito participar voluntariamente e autorizo a divulgação de dados relacionados à minha oralidade.

Assinatura do informante

Loremi Loregian-Penkhal - pesquisadora da UNICENTRO

Lucelene Franceschini - pesquisadora da UNICENTRO

Guarapuava, _____ de _____ de 2016.

ANEXO III

ROTEIRO DE PERGUNTAS

1. Qual a sua **idade**?
2. Onde você **nasceu**?
3. Qual a sua descendência? Por parte de pai e/ou de mãe?
4. Fale sobre a tua **infância**: como eram as brincadeiras, os brinquedos, as coisas que você fazia naquela época.
5. Na tua opinião, mudou muita coisa da tua **infância** para hoje em dia? Em que sentido?
6. Como era na época em que você frequentou a **escola**? E os professores?
7. Tem alguma **história engraçada ou triste** dessa época que você sempre lembra?
8. Como era a tua **família** quando você era criança?
9. Fale sobre a **colonização** da tua cidade. Em que época a tua família veio pra cá?
10. Como é comemorado o **NATAL** na tua família? Mudou muito de quando você era criança pra hoje?
10. E a festa de **PÁSCOA**?
11. O que você costuma fazer nos finais de semana?
12. Qual o tipo de **lazer** que os jovens tinham antigamente? E hoje?
13. Como eram os **casamentos** de antigamente? E hoje, mudou muito?
14. Você segue as **tradições e costumes** de seus pais? Quais? Por quê?
15. Que **comidas** você conhece e gosta de fazer? Explique como se faz.
16. Qual a sua **religião**? Você é praticante?
17. Como eram os **costumes religiosos** de antigamente? E hoje, mudaram muito?
18. Você (ou alguém conhecido) já passou por algum **perigo de morte**? Como foi?
19. Como é o teu **trabalho**?
20. Você já saiu daqui? Que cidades conhece? Como foi a **viagem**?
21. Tem **filhos**? Se sim, de que idade?

22. Você tem algum **sonho** que gostaria ou pretende realizar?
23. Você tem **animais** de estimação? Quais?
24. O que você mais gosta na sua **comunidade**? Na sua opinião, o que poderia melhorar?
- a) O que você acha do sistema de saúde (hospitais, postos de saúde...) da cidade?
- b) Já teve algum problema sério de saúde na família? Se sim, como foi a assistência médica e/ou hospitalar?
- c) O que você acha do sistema escolar da cidade?
- d) Como você avalia a infraestrutura da cidade? (água, coleta de lixo, esgoto, estradas...)
25. Se não nasceu em Guarapuava, com que idade se mudou para essa cidade?
26. Na sua opinião, o que é ser guarapuavano? Basta nascer em Guarapuava ou não?
27. É fácil de identificar um guarapuavano? Por quê?
28. Você gosta de ser guarapuavano? Por quê?
29. Você se mudaria de Guarapuava? Por quê?
30. Você fala outra língua? Qual? Como aprendeu?
31. Você percebe alguma característica especial na fala (pronúncia, palavras, expressões típicas) das pessoas de Guarapuava?

ANEXO IV

VARIÁVEIS CN - GUARAPUAVA

Variável dependente:

1: presença de CN

0: ausência de CN

Variáveis independentes:**1. Saliência fônica**

d: plural duplo

l: singular em – l

a: singular em – ão (pl. irregular)

r: singular em – r

s: singular em – s

p: plural regular

2. Tonicidade

y: monossílabos átonos

t: oxítonos e monossílabos tônicos

x: paroxítonos

z: proparoxítonos

3. Saliência fônica e TonicidadeD: plural duplo – *ovo/ovos*L: singular terminado em –l (oxítono ou paroxítono) – *papel/papéis*A: singular em –ão (oxítonos irregulares) – *televisão/televisões*R: singular em –r (oxítonos ou paroxítonos) – *mulher/mulheres*S: singular em –s (oxítono) – *freguês/fregueses*O: regular oxítono – *café/cafés, mão/mãos*P: regular paroxítono – *campo/campos*X: regular proparoxítono – *último/últimos*

N: monossílabos átonos

4. Posição do elemento no SN

1 – primeira posição

2 – segunda posição

3 – terceira posição

4 – quarta ou demais posições

5. Classe Gramatical

- A – artigo
- S – substantivo
- E – adjetivo
- P – possessivo
- D – demonstrativo
- C – categoria substantivada
- I – indefinido (outro, uns, alguns)
- Q – quantificador (todos, bastante, pouco)

6. Posição e classe gramatical do elemento

- a – determinantes antepostos ao núcleo na 1ª posição - tá **nos** café
- b – determinantes antepostos ao núcleo na 2ª posição – *todos os* amigos
- c – adjetivo anteposto ao núcleo -
- d – substantivo como núcleo na 1ª posição
- e – substantivo como núcleo na 2ª posição
- f – substantivo como núcleo na 3ª posição e 4ª posição
- g – categoria substantivada como núcleo na 2ª posição
- h – categoria substantivada como núcleo na 3ª posição e 4ª posição
- i – determinantes pospostos ao núcleo
- j – adjetivo posposto ao núcleo na 2ª posição
- k – adjetivo posposto ao núcleo nas demais posições
(*Determinantes*: artigos, pronomes e quantificadores)

7. Marcas precedentes

- 1 – ausência de elemento precedente - *os filhos*
- 2 – elemento precedente é um numeral – cinco anos
- 3 – ausência de marca morfológica no elemento precedente – no lugares que eu tava
- 4 – apenas uma marca (em –s) precedente – os livros
- 5 – duas ou mais marcas (em –s) precedente – os meus livros
- 6 – mistura de marcas com **zero precedente** à 3ª e 4ª posição - uns menino bagunceiro

7 – mistura de marcas com **–s precedente** à 3ª e 4ª posição – a minhas irmã jogavam bola

8 – sintagma preposicionado precedente – uma porção de fotos, um monte de galho

9 – termo precedente é invariável – *várias* vezes, *mais* velhos.

Variáveis sociais:

8. Sexo

M: masculino;

F: feminino.

9. Faixa etária

1: 25 a 49 anos;

2: 50 anos ou mais.

10. Escolaridade

p: Fundamental I;

g: Fundamental II;

c: Ensino médio.