



UNIVERSIDADE
ESTADUAL DE LONDRINA

ANA MARIA MATTOS NAVARRO

**A VARIAÇÃO NA CONCORDÂNCIA DE GÊNERO NA
VARIEDADE LINGÜÍSTICA DO PARANÁ TRADICIONAL
RURAL**

Londrina
2005

ANA MARIA MATTOS NAVARRO

**A VARIAÇÃO NA CONCORDÂNCIA DE GÊNERO NA
VARIEDADE LINGÜÍSTICA DO PARANÁ TRADICIONAL
RURAL**

Dissertação apresentada ao
Programa de Pós-Graduação, em
Estudos da Linguagem da
Universidade Estadual de Londrina,
como requisito parcial à obtenção do
título de Mestre.

Orientadora: Prof^a. Dr^a. Vanderci de
Andrade Aguilera

Londrina
2005

**Catálogo na publicação elaborada pela Divisão de Processos Técnicos da
Biblioteca Central da Universidade Estadual de Londrina.**

Dados Internacionais de Catalogação-na-Publicação (CIP)

N322v Navarro, Ana Maria Mattos.

A variação na concordância de gênero na variedade lingüística do Paraná tradicional rural / Ana Maria Mattos Navarro. – Londrina, 2005.
125 f.

Orientadora : Vanderci de Andrade Aguilera.

Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, 2005.

Bibliografia : f. 121-125.

1. Sociolingüística – Teses. 2. Língua portuguesa – Concordância de gênero – Teses. 3. Morfossintaxe – Língua portuguesa – Teses. I. Aguilera, Vanderci de Andrade. II. Universidade Estadual de Londrina. III. Título.

CDU 806.90-5

ANA MARIA MATTOS NAVARRO

**A VARIAÇÃO NA CONCORDÂNCIA DE GÊNERO NA VARIEDADE
LINGÜÍSTICA DO PARANÁ TRADICIONAL RURAL**

Dissertação apresentada ao Programa de Pós-Graduação, em Estudos da Linguagem da Universidade Estadual de Londrina, como requisito parcial à obtenção do título de Mestre.

Orientadora: Prof^a. Dr^a. Vanderci de Andrade Aguilera

COMISSÃO EXAMINADORA

Prof. Dr. Ludoviko Carnasciali dos Santos
Universidade Estadual de Londrina

Prof^a. Dr^a. Maria Marta Pereira Scherre
Universidade de Brasília

Prof^a. Dr^a. Vanderci de Andrade Aguilera
Universidade Estadual de Londrina

Londrina, 7 de março de 2005

A Deus, à minha família, e a todos
que desejaram meu sucesso nesta
jornada.

AGRADECIMENTOS

Reconhecer que um trabalho como este resulta de um longo e árduo processo de reflexão, aprendizado, amadurecimento investimento e esforço pessoal, não me desobriga do reconhecimento das inúmeras e valiosas contribuições que o tornaram possível. Sendo assim, agradeço de coração:

À Professora Vanderci, pela dedicação e disponibilidade, pela orientação segura, pela confiança em mim depositada ao acreditar no meu trabalho e, sobretudo, por ter me incentivado a vencer meus próprios medos.

Ao professor Ismael que despertou em mim o interesse pela Sociolingüística.

Aos professores Sônia e Ludoviko pelas sugestões enriquecedoras dadas no Exame de Qualificação.

Às tias Vera e Homilda e à minha mãe, pelo apoio, ajuda, compreensão e respeito que demonstraram ao longo de todo o meu Curso de Mestrado.

A Elen, principal responsável pela minha decisão de ingressar no Mestrado, por seu apoio e atitude positiva, sempre me motivando a seguir em frente.

Ao Gabriel, que, à sua maneira, também me ajudou nesta jornada.

À Elaine, pelo companheirismo, pela paciência de ter me ouvido durante todo esse tempo e pelas trocas de idéias em relação ao discurso científico.

A Ednéia, pelas sugestões e, principalmente, pela ajuda num dos momentos mais difíceis que passei ao final do Curso.

À Alessandra, pela ajuda no uso do programa computacional e pelas sugestões relativas à análise de dados.

Aos mestres de toda a vida, desde os do curso primário, e tantos outros, principalmente minhas professoras de língua portuguesa no ensino fundamental e médio.

Aos professores do Mestrado, pelo profissionalismo sempre presente em todas as atividades desenvolvidas.

A Deus, que tem cuidado dos meus caminhos, tem me orientado e abençoado, permitindo-me alcançar mais uma conquista.

NAVARRO, Ana Maria Mattos. **A variação na concordância de gênero na variedade lingüística do Paraná Tradicional rural**. 2005.113 f. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina.

RESUMO

Este trabalho tem como objetivo analisar a variação na concordância de gênero entre o sujeito e o predicativo na variedade do português falada na região rural do Estado do Paraná, colonizada entre os séculos XVII e XIX, denominada Paraná Tradicional. O construto teórico-metodológico da Sociolingüística Variacionista dá suporte à pesquisa que tomou por base a aplicação da regra: *presença da concordância de gênero entre o sujeito e o predicativo*. O corpus utilizado neste estudo foi constituído a partir das entrevistas coletadas durante a pesquisa de campo para a elaboração do Atlas Lingüístico do Paraná – ALPR - (AGUILERA, 1994) que se apresentam em três versões: *Atlas Lingüístico do Paraná* (AGUILERA, 1990a) - banco de dados sonoro, inédito; *Atlas Lingüístico do Paraná* (AGUILERA, 1990b) - versão digitada em 13 disquetes; *Atlas Lingüístico do Paraná* (AGUILERA, 1990c) - versão impressa, 13 volumes. Os informantes, em número de quarenta, são estratificados por sexo, idade (30 a 46 anos e 47 a 65 anos) e escolaridade (analfabeto e escolarizado). Os dados analisados receberam tratamento estatístico a fim de determinar os fatores condicionantes mais significativos para a produção do fenômeno pesquisado. O programa estatístico empregado na pesquisa é o VARBRUL, em uma versão mais recente, o GOLDVARB 2001 (ROBINSON; LAWRENCE; TAGLIAMONTE, 2001).

Palavras-chave: Concordância de gênero. Sociolingüística. Variáveis. Variação lingüística. Morfossintaxe. Predicativo.

NAVARRO, Ana Maria Mattos. **Variation in gender concordance in the linguistic variety of rural Traditional Paraná**. 2005.113 f. Dissertation (Master Program in Language Studies) – Universidade Estadual de Londrina.

ABSTRACT

This work was meant to analyze the variation of gender concordance between the subject and the predicate in the variety of the Portuguese language spoken in the rural area of State of Paraná, colonized between Centuries XVII and XIX, called Traditional Paraná. The theoretical and methodological construct from the Variationist Sociolinguistics gives support to the research work based on the use of the rule: *presence of gender concordance between the subject and the predicate*. The *corpus* used in this study was obtained from interviews made during the field research for the elaboration of the linguistic atlas of Paraná - Atlas Lingüístico do Paraná – ALPR - (AGUILERA, 1994) and these are presented in three versions: *Atlas Lingüístico do Paraná* (AGUILERA, 1990a) – sound databank, original; *Atlas Lingüístico do Paraná* (AGUILERA, 1990b) – typed version in 13 disks; *Atlas Lingüístico do Paraná* (AGUILERA, 1990c) – printed version, 13 volumes. The subjects, a total of 40, are grouped by sex, age (30 to 46 and 47 to 65 years old) and educational background (illiterate and educated). The analyzed data were treated statistically in order to determine the most significant conditioning factors for the production of the phenomenon under research. The statistics program used in the research was VARBRUL, in a more recent version, GOLDVARB 2001 (ROBINSON, LAWRENCE & TAGLIAMONTE, 2001).

Key words: Gender concordance. Sociolinguistics. Variables. Linguistic variation. Morphosyntax. Predicate.

LISTA DE ILUSTRAÇÕES

FIGURA 1: Os 20 pontos lingüísticos selecionados para a pesquisa.....	65
QUADRO 1: Ponto lingüístico de origem do informante e as variáveis sociais idade, escolaridade e sexo	67
QUADRO 2: Distribuição dos informantes por faixa etária e sexo	68
QUADRO 3: Distribuição dos informantes por escolaridade e sexo.....	69

LISTA DE TABELAS

TABELA 1: Número de Ocorrências e Percentuais das Variantes da CG.....	81
TABELA 2: Efeito da Variável Núcleo do Predicativo.....	83
TABELA 3: Efeito da Variável Ordem dos Elementos na Construção.....	85
TABELA 4: Efeito da Variável Característica Formal do Sujeito.....	87
TABELA 5: Efeito da Variável Tipo de Referência do Sujeito.....	90
TABELA 6: Efeito da Variável Verbo Utilizado	92
TABELA 7: Efeito da Variável Elementos Intervenientes entre o Verbo e o Predicativo.....	95
TABELA 8: Efeito da Variável Sexo	97
TABELA 9: Efeito da Variável Grau de Escolaridade	98
TABELA 10: Efeito da Variável Faixa Etária.....	100

SUMÁRIO

INTRODUÇÃO	12
1 DESCRIÇÃO E DELIMITAÇÃO DO OBJETO DE ESTUDO	19
1.1 O Paraná dos Séculos XVII, XVIII E XIX	19
1.2 O Gênero Gramatical na Variedade Lingüística do Paraná Tradicional Rural.....	21
2 PRESSUPOSTOS TEÓRICOS	29
2.1 Fundamentos da Sociolingüística.....	29
2.2 Teoria da Variação	31
2.3 A Análise Quantitativa	36
3 O GÊNERO NO PORTUGUÊS DO BRASIL	38
3.1 O Gênero como Categoria Gramatical	38
3.2 Flexão ou Derivação.....	40
3.3 A Concordância Nominal de Gênero na Relação Sujeito/Predicativo	43
3.4 A Variação do Gênero nas Variedades Populares do Português do Brasil	44
3.5 Estudos Variacionista sobre a Concordância de Gênero no Português do Brasil	49
3.5.1 Lucchesi (2000): A variação na concordância de gênero em uma comunidade de fala afro-brasileira: novos elementos sobre a formação do português popular do Brasil	49
3.5.2 Dettoni (2003): A concordância de gênero na anáfora pronominal: variação e mudança lingüística no dialeto da baixada cuiabana – Mato Grosso	55
3.5.3 Souza (2000): A concordância de gênero entre o sujeito e o predicativo na fala da comunidade quilombola da Caçandoca.....	58

3.5.4 Karim (2004): A variação na concordância de gênero no falar da comunidade de Cáceres – MT	60
4 PROCEDIMENTOS METODOLÓGICOS	63
4.1 Os Pontos Lingüísticos.....	64
4.2 Os Informantes.....	66
4.3 A Variável Dependente.....	71
4.4 As Variáveis Independentes.....	72
4.5 Codificação e Quantificação dos Dados.....	75
5 ANÁLISE DOS DADOS	79
5.1 Resultados da Análise Quantitativa.....	79
5.1.1 Variáveis lingüísticas significativas estatisticamente	81
5.1.2 Variáveis lingüísticas não significativas estatisticamente	91
5.1.3 Variáveis Sociais	100
CONSIDERAÇÕES FINAIS	103
REFERÊNCIAS	107

INTRODUÇÃO

Cada fase pela qual passa uma determinada sociedade é assinalada pela linguagem dos grupos que a compõem. A língua é sempre responsável pela oferta, em qualquer época, de elementos que contribuem para a leitura, caracterização e entendimento dos grupos sociais, caracterizando-se como um produto social e um instrumento essencial à interação do homem com sua comunidade.

É fato comprovado, em todas as áreas da linguagem, que a língua é dinâmica, pois sua renovação ocorre a todo instante em que é usada, levando-a a sofrer constantes transformações: cada falante é, ao mesmo tempo, usuário e transformador de sua língua e tem uma maneira própria de se expressar, a qual atua em conjunto com fatores lingüísticos, sociais (idade, sexo, escolaridade, situação sócio-econômica) ou de origem geográfica. Assim, é natural que a fala de qualquer comunidade seja suscetível às variações.

Em razão disso, e para um verdadeiro conhecimento da cultura de determinada comunidade, é necessário vivenciar e observar a forma através da qual os falantes se manifestam por meio da língua e entender a realidade que a circunda. Desse modo, justifica-se o papel de uma ciência como a Sociolingüística que, ao elevar as variações lingüísticas ao *status* de objeto de estudo, se dedica a estudar a relação entre fatos lingüísticos e fatos sociais.

A linguagem praticada por falantes de muitas regiões do Brasil oferece desafios aos pesquisadores do campo da Sociolingüística. Em um país de dimensões continentais, identificado pelo privilégio de reunir uma vasta

multiplicidade em níveis tão diferentes, seja do ponto de vista geográfico, climático ou cultural, seja em muitos outros níveis, é impossível ignorar a presença do intenso fenômeno de trocas culturais diversas. Decorre, a partir daí, então, a abertura de espaços para a convivência de diferentes variedades lingüísticas, algumas ainda inexploradas pela pesquisa lingüística.

Nesse contexto, a idealização desta pesquisa sobre a variedade lingüística falada na região rural do estado do Paraná surgiu, a princípio, de uma reflexão nascida a partir de um grande interesse despertado pelas disciplinas de Sociolingüística e Dialetoлогия que cursamos no início do mestrado. Ambas nos possibilitaram um maior contato com obras de autores consagrados na área, bem como com pesquisas, conceitos, descrições, análises e interpretações das variações lingüísticas que ocorrem na variante brasileira da língua portuguesa, envolvendo os diversos dialetos das regiões brasileiras.

Após algumas leituras, investigações e estudos, verificamos que a variação da concordância de gênero nos falares rurais do Paraná se constitui num amplo objeto de estudo e análise pouco explorado, o que nos motivou a elaborar um projeto de pesquisa que contribuísse para a descrição de uma variedade específica do português brasileiro, focalizando a variação na concordância nominal de gênero.

Temos conhecimento de algumas pesquisas já divulgadas sobre outras variedades do português do Brasil que atestam a presença da realização variável da concordância de gênero. Citamos, entre elas, a tese de doutorado de Lucchesi (2000), que têm como objeto de estudo o dialeto falado em Helvécia, comunidade localizada no Extremo Sul do Estado da Bahia e a tese de Dettoni (2003) sobre a concordância de gênero na anáfora pronominal no dialeto da Baixada Cuiabana, no Estado de Mato Grosso; a dissertação de mestrado de Souza (2000), envolvendo a

variação na concordância de gênero entre o sujeito e o predicativo na fala da comunidade da Caçandoca, no Estado de São Paulo e a dissertação de Karim (2004) sobre a variação na concordância de gênero no falar da comunidade de Cáceres, no Estado de Mato Grosso.

Apesar da importância de trabalhos como os citados acima, ainda há muito - e sempre haverá - a ser feito para que se obtenha um maior conhecimento acerca da variação na concordância nominal de gênero nas diversas variedades da língua portuguesa no Brasil.

Neste trabalho, propomo-nos focalizar a morfossintaxe, componente ainda não explorado dessa variedade. De maneira mais específica, este estudo é focado na análise da variação na concordância de gênero entre o sujeito e o predicativo¹, fenômeno morfossintático bastante produtivo na variedade do português falada na região rural do Estado do Paraná.

Ressaltamos que, freqüentemente, recorreremos à expressão *Paraná Tradicional rural* como referência à área geográfica à qual pertencem os informantes que tiveram seus falares analisados nesta pesquisa. De acordo com divisão proposta por Cardoso e Westphalen (1986), o Paraná Tradicional surgiu no século XVII em decorrência do ciclo do ouro e nos séculos XVIII e XIX se estruturou com a atividade dos tropeiros e a dos imigrantes europeus, respectivamente.

Esclarecemos ainda que durante a fase de delimitação do *corpus* outras formas de variação do gênero gramatical nas estruturas sob análise despertaram nosso interesse. Destacamos como mais recorrentes a oscilação na

¹ Para esta pesquisa entende-se por predicativos também os participios passivos, os quais foram devidamente analisados e controlados sob o nome *formas participais*, através da variável *núcleo do predicativo*.

marcação do gênero nas relações anafóricas, nas relações entre os elementos do sintagma nominal (SN) e a substituição de vocábulos classificados no português padrão como feminino por vocábulos flexionados no masculino e vice-versa. Contudo, a variação na concordância de gênero entre o sujeito e o predicativo foi o tipo de construção mais produtivo, isto é, o que apresentou o maior número de ocorrências, fator decisivo para que optássemos por tratar somente deste tipo de variação.

Uma avaliação preliminar dos dados nos levou a constatar que os núcleos de predicativos representados por nomes masculinos eram ligados, com frequência, a sujeitos representados por nomes femininos, o que foge à forma padrão. Por outro lado, núcleos de predicativo representados por itens lexicais classificados como masculino no léxico português eram sempre ligados a sujeitos também representados por um item na forma masculina, de acordo com a forma padrão.

Assim sendo, a fim de caracterizar a variação da concordância de gênero entre o sujeito e o predicativo na variedade lingüística em estudo nesta dissertação, foram considerados como dados para esta pesquisa apenas as construções em que um item lexical classificado como feminino dentro do léxico da língua portuguesa exerce a função de sujeito, o qual se liga, por sua vez, a um núcleo de predicativo biforme.

Este estudo parte do pressuposto de que a produção do fenômeno da variação da concordância de gênero entre o sujeito e o predicativo no falar do Paraná Tradicional rural pode ser explicada tanto pela atuação de variáveis lingüísticas, que representam os fatores estruturais da gramática da comunidade na

qual a variação na concordância de gênero está relacionada, quanto por variáveis sociais, que podem oferecer um quadro dos padrões de comportamento lingüístico na estrutura social da comunidade. Para tanto, optou-se pelos pressupostos teóricos da Sociolingüística Variacionista devido à propriedade desta em associar a atividade lingüística com a situação social do falante, afastando-se do tipo de análise que considera apenas categorias gramaticais para o entendimento das variedades lingüísticas ao estudar os fenômenos lingüísticos.

Com a análise dos dados pretendemos investigar algumas hipóteses em relação à produção do fenômeno em estudo:

a) A partir de uma observação inicial dos dados, complementada pela pesquisa a autores como Câmara Júnior (1995), Martins (1975) e Mira Mateus et al (1994) postulamos a hipótese de que a referência genérica do sujeito seria o principal fator condicionante da ausência de concordância de gênero entre o sujeito e o predicativo na amostra de fala analisada.

b) Conforme verificado por Scherre (1991), em trabalho sobre a variação na concordância de número em português, a interrupção do fluxo do processamento linear do pensamento do falante pode ser atribuída à ordem não canônica, a qual desfavoreceria a concordância. Esperamos que os dados desta pesquisa confirmem tal afirmação também para a concordância de gênero, de forma que a ordem que os elementos ocupam na construção possa ser um dos fatores que expliquem a ausência de concordância verificada na fala dos informantes deste estudo.

c) Tendo em vista nosso conhecimento de várias pesquisas variacionistas, entre elas as de Lucchesi (2000) e Dettoni (2003), apresentadas nos

capítulo 3, nas quais a fala de informantes mais jovens e escolarizados reflete um comportamento lingüístico marcado pela substituição de formas desprestigiadas pelas formas padrão da língua, será considerada também a hipótese da ocorrência de um processo de mudança nesta variedade, marcada pela perda gradual de estruturas sem concordância de gênero e o uso, cada vez maior, de construções padrão do português.

Serviram de fonte para o estudo que se desenvolve a seguir os dados obtidos pelos pesquisadores que participaram da elaboração do Atlas Lingüístico do Paraná (AGUILERA, 1994) – doravante ALPR². Foram inquiridos 130 informantes distribuídos em 65 pontos lingüísticos, de modo a representar todo o Estado. Em decorrência do grande número de informantes, delimitamos nossa investigação aos pontos lingüísticos correspondentes aos núcleos de povoamento dos períodos dos séculos XVII, XVIII e XIX e passamos a contar com dados de 40 informantes.

As entrevistas se encontram transcritas e registradas nas seguintes versões: *Atlas Lingüístico do Paraná* (AGUILERA, 1990a) - banco de dados sonoro, inédito; *Atlas Lingüístico do Paraná* (AGUILERA, 1990b) - versão digitada em 13 disquetes; *Atlas Lingüístico do Paraná* (AGUILERA, 1990c) - versão impressa, 13 volumes³.

Este trabalho é composto de cinco capítulos, além de uma parte introdutória e das considerações finais. O capítulo primeiro descreve sinteticamente

² O Atlas Lingüístico do Paraná, constituído por 92 cartas lexicais, 70 fonéticas e 29 sintéticas de isoglossas, é resultado de uma pesquisa que envolveu 65 municípios paranaenses, abrangendo mais de uma centena de comunidades rurais, em sua grande maioria.

³ Mesmo privilegiando a versão digitada em disquete para dar corpo a este trabalho, também sentimos a necessidade, muitas vezes, de fazer uso das versões impressa e sonora, como apoio durante a delimitação do *corpus*.

alguns aspectos históricos da formação do Estado do Paraná e enfoca a variedade do português falada no Paraná Tradicional rural, caracterizando-a a partir de seus traços mais típicos quanto à variação do gênero gramatical.

No segundo capítulo são abordados os pressupostos teóricos da Teoria da Variação que dão sustentação às análises desenvolvidas neste estudo.

O terceiro capítulo traz uma abordagem teórica sobre o gênero como categoria gramatical e seus mecanismos de flexão e de derivação. Em seguida, é apresentada uma revisão de alguns trabalhos sociolingüísticos e dialetológicos sobre variedades populares e regionais do Brasil, com destaque para os aspectos relativos ao gênero gramatical. Para finalizar o capítulo, são descritas algumas pesquisas variacionistas recentes sobre a concordância de gênero no português do Brasil.

O quarto capítulo explicita a metodologia empregada para o desenvolvimento desta dissertação. Nele são descritos o método utilizado na pesquisa, a formação do *corpus*, as variantes, as variáveis lingüísticas e sociais, bem como as hipóteses que nortearam a escolha de cada variável e, por fim, o processo de codificação e quantificação dos dados.

O quinto capítulo trata da análise variacionista da concordância de gênero entre os sujeitos e os predicativos, com base na apresentação e discussão dos resultados estatísticos das ocorrências do fenômeno lingüístico em estudo, expostos em forma de tabelas.

1 DESCRIÇÃO E DELIMITAÇÃO DO OBJETO DE ESTUDO

Este trabalho tem por objetivo analisar a variação na concordância de gênero em construções da variedade do português falada pelos informantes do ALPR (AGUILERA, 1990a, 1990b, 1990c, 1994). Antes de partirmos para as questões lingüísticas, exporemos sucintamente alguns aspectos da ocupação do Estado⁴, a fim de caracterizar melhor o objeto de estudo desta pesquisa. Em seguida, apresentamos uma breve descrição dos principais aspectos da variedade do português falada no Paraná Tradicional rural, caracterizando-a a partir de seus traços mais típicos quanto à variação do gênero gramatical.

1.1 O Paraná dos Séculos XVII, XVIII E XIX

A história do Paraná pode ser contada através da ocupação de seu território. Sua área abrange quase 200 mil quilômetros quadrados e possui um dos mais férteis solos do mundo, abrigando um rico entrelaçamento de raças e culturas. Não se pode deixar de reconhecer que a especificidade do Paraná está na sua formação diferenciada. Essa diversidade de raças foi traçada pelo cruzamento de muitos caminhos abertos pelo homem ao longo dos séculos. O Paraná é europeu, asiático, gaúcho, paulista, mineiro e ‘pé vermelho’⁵.

⁴ Esta seção baseia-se, em grande parte, nas informações contidas em Cardoso e Westphalen (1986) e em Martins (1995).

⁵ A expressão ‘pé vermelho’ faz alusão à cor da terra do Norte do Estado. É comumente empregada para

Geograficamente, o Estado do Paraná se localiza ao norte da região sul do Brasil, possui uma área de 199.709,1 km² e é formado por uma baixada no litoral, planaltos a oeste e a leste e uma depressão central. Seus principais rios são: Paraná, Iguaçu, Ivaí, Tibagi, Paranapanema, Itararé e Piquiri; Curitiba, Londrina, Ponta Grossa e Maringá são as maiores cidades do Estado. Segundo Cardoso e Westphalen (1986) a colonização do território paranaense se deu a partir do litoral e em três ondas povoadoras que viabilizaram sua ocupação, criando as seguintes comunidades regionais:

a) *Paraná Tradicional*: é o Paraná dos primeiros navegadores que, no século XVII, aportaram na baía de Paranaguá, em busca de ouro, criando as comunidades mais antigas; dos tropeiros que, no século XVIII, para fazer o comércio de gado, abriram caminhos e fizeram surgir cidades nos Campos Gerais e dos imigrantes europeus, os quais, por sua vez, no século XIX, atraídos pela promessa de fortuna, fortaleceram a agricultura, a extração e indústria da erva-mate e da madeira;

b) *Paraná Moderno do Norte*: tem sua história iniciada no século XX, quando o norte do Estado, ocupado por imigrantes europeus e pela migração de paulistas e mineiros, foi conquistado pelo café - considerado pelos agricultores mais antigos como o ouro verde extraído da terra roxa.

c) *Paraná Moderno do Oeste e Sudoeste*: surge por meio do povoamento da região por imigrantes gaúchos, a partir da abertura de novas fronteiras agrícolas no oeste e sudoeste e, mais recentemente, pela industrialização dos grandes centros urbanos.

Inicialmente, pretendíamos investigar a variação na concordância nominal de gênero em todas as entrevistas realizadas dentro do Estado pela equipe do ALPR (AGUILERA, 1990a, 1990b, 1990c, 1994), o que abrangeria os municípios pertencentes às três ondas povoadoras descritas acima. Entretanto, tendo em vista o tempo disponível para a elaboração deste trabalho, defrontamo-nos com dificuldades quanto à delimitação da amostra de fala a ser analisada devido ao número muito elevado de entrevistas, 130 no total, que representavam os 65 pontos lingüísticos estabelecidos para a elaboração do ALPR (AGUILERA, 1994).

Desse modo, o primeiro recorte que fizemos foi reduzir a investigação aos pontos lingüísticos relacionados ao início da colonização do Paraná. Seguindo essa distinção, ligada ao processo de ocupação e povoamento do Estado, os pontos lingüísticos selecionados se encaixam na comunidade denominada *Paraná Tradicional* (cf. figura 1, p. 74).

1.2 O Gênero Gramatical na Variedade Lingüística do Paraná Tradicional Rural

Em nossa caminhada, durante o processo de elaboração deste trabalho, tivemos contato com algumas dissertações que utilizaram dados da pesquisa de campo do ALPR (AGUILERA, 1990a, 1990b, 1990c) na produção de estudos de natureza discursiva, lexical e fonética. Entre eles, citamos as pesquisas de Almeida (1996), Rodrigues (2000) e Gusmão (2004).

A dissertação de mestrado de Almeida (1996), “Discurso rural: uma perspectiva ideológica”, apesar de utilizar dados de uma pesquisa dialetológica,

segue os pressupostos teóricos relacionados à Análise do Discurso e aos recursos lingüístico-textuais, apresentando as formações ideológicas contidas no discurso em análise, evidenciando sua realidade social. A ideologia que subjaz ao discurso de falantes rurais é, para isso, identificada, descrita e analisada tomando como *corpus* as respostas à questão de número 325 do questionário aplicado aos inquiridos, que consta de uma narrativa de experiência pessoal.

A hipérbole, a antítese, o eufemismo e a prosopopéia, bem como a repetição, o deslocamento e os operadores argumentativos foram identificados nas narrativas analisadas. A autora ressalta que por meio deles foram revelados alguns componentes que se fazem presentes na realidade social dos falantes rurais, tais como: o tabu lingüístico, a necessidade de consideração do heroísmo das pessoas de camadas pobres e a ênfase no sofrimento, oriundos da sua pobreza.

Almeida conclui que os efeitos de realidade confirmam os conceitos teóricos que os fundamentam, pois colaboram, em sua totalidade, para assegurar a veracidade dos fatos.

A pesquisa de Rodrigues (2000), “Para um glossário da fala popular rural paranaense”, envolveu o léxico no falar rural paranaense e resultou na construção de um glossário a partir da reunião, análise e seleção de obras de cunho léxico elaboradas para o referido Estado. Sua organização privilegiou os exemplos, extraídos, em grande parte, das entrevistas feitas para o ALPR (AGUILERA, 1990b, 1990c, 1994), com o intuito de ressaltar a importância desse componente para a compreensão da lexia e, ao mesmo tempo, para a divulgação da cultura popular rural.

Gusmão (2004) produziu uma dissertação intitulada “As realizações da vibrante na variedade lingüística rural do Paraná - uma abordagem geovariacionista” que toma como objeto de estudo o fonema /r/ na variedade lingüística rural falada no Paraná. Para isso, analisou 5.633 realizações desse fonema, obtidas nas entrevistas do ALPR (1990a). Foram selecionados os registros de setenta e dois informantes, de ambos os sexos, distribuídos em trinta e oito pontos lingüísticos, representando todo o território paranaense. Sob a perspectiva da Teoria da Variação, Gusmão buscou descrever contextos fônicos favoráveis ao uso de certa(s) variante(s) da vibrante, detectar influências sociais e estabelecer áreas geográficas de predomínio de determinadas articulações das variantes.

A análise quantitativa dos dados revelou que as diversas formas de realização de /r/ são determinadas por fatores lingüísticos e sociais. Em todos os contextos de /r/ estudados, com exceção de /r/ fraco, a primeira variável selecionada foi a *diatópica*, e entre as demais variáveis (lingüísticas e sociais) foi verificada uma maior presença das variáveis sociais, como a *faixa etária*, a *escolaridade* e o *gênero*.

No entanto, até o presente momento, não temos notícias de pesquisas realizadas a partir dos dados do ALPR (AGUILERA, 1990a, 1990b, 1990c) no desenvolvimento de estudos morfossintáticos, sobretudo relacionados à concordância nominal de gênero. Neste sentido, a opção por investigar a variedade falada na região rural do Paraná se apóia no fato de a variação do gênero gramatical se constituir num fenômeno morfossintático ainda pouco descrito e explorado pela pesquisa lingüística do Brasil e praticamente inédito nos estudos já realizados sobre o dialeto rural do Paraná.

Constatamos, ainda, que as informações advindas de estudos sobre a concordância de gênero são, de certo modo, bem distintas, sem que se possa definir com clareza a extensão de sua variação na fala popular. Segundo Lucchesi (2000) a variação na concordância nominal de gênero não se trata de um fenômeno comum entre as variedades do português, só sendo registrada em um nível significativo em certas comunidades rurais que passaram por um amplo e profundo contato lingüístico em sua história. Dentro do panorama lingüístico brasileiro, trata-se de um fenômeno, de certo modo, bem mais localizado, quase inexistente nas variedades cultas, mesmo nos seus registros mais informais.

Entretanto, com relação à concordância de número em estruturas nominais não se pode dizer o mesmo, visto que vários estudos e pesquisas a apontam como bastante geral e bem significativa, pois se encontra amplamente generalizada, chegando a ocorrer também entre as camadas mais cultas da sociedade.

Também achamos oportuno trabalhar com o falar rural paranaense por acreditarmos que essa variedade se encontra em processo de mudança em decorrência do fato de que cada vez mais caem as fronteiras que separam o homem do campo do homem dos centros urbanos. O desenvolvimento nas áreas de transportes e de comunicação, o fluxo de movimentos migratórios e o acesso das camadas populares à escola, após a democratização do ensino, possibilitou aos falantes da zona rural um maior contato com o universo urbano, ocasionando mudanças sócio-culturais que trouxeram, como um de seus reflexos, mudanças no sistema lingüístico de tais falantes.

Dentre as variedades de menor prestígio social, a rural é uma das que mais sofrem discriminação, devido ao seu caráter bastante característico em relação a vários fenômenos fonológicos, morfológicos e sintáticos. Seus usuários, muitas vezes, ainda são rotulados com o estigma da ignorância e da exclusão, apesar de vivermos num século marcado pela globalização, o que torna o mundo menor e que deveria, a princípio, facilitar a tolerância em relação aos vários tipos de diversidade, sejam elas culturais, comportamentais ou lingüísticas.

Uma análise preliminar das entrevistas selecionadas para esta pesquisa nos permitiu detectar vários tipos de variação de gênero na variedade pesquisada. Levantamos, assim, um número considerável de ocorrências que envolviam o fenômeno em estudo. Pudemos agrupá-las do seguinte modo:

I. Vocábulos classificados no português padrão como feminino, mas flexionados no masculino e vice-versa:

(1)... cheguei conhecer *uma guiza* de cascavel... (ME47) ⁶

(2)... tem muita *reumatismo*... (MI49)

(3)... o *lagarticho*. Ah! Esse dá na, na, lá na roça... (ME53)

II. Nas relações internas ao sintagma nominal:

(4)... quando é *mandioca brabo*... (MI49)

(5)... pois é *perna fino*, né... (FI52)

(6)... eu vi *mulher sentado*... (FI37)

III. Nas relações anafóricas nas quais nomes femininos são retomados por um pronome na forma masculina:

⁶ Os números e letras correspondem ao código utilizado para facilitar a identificação do informante. A primeira letra, **F** ou **M**, significa sexo do informante; as letras **I** e **E** correspondem à escolaridade – **I** para iletrado e **E** para escolarizado; o número corresponde à idade.

(7)... tem uma *cachoeira* que *ele* nasce lá em cima... (FE59)

(8)... *banana maçã* [...] *ele* não forma direito... (FE46)

IV. Na relação sujeito/predicativo:

(9)... a *espiga* ficou *ralo*... (MI35)

(10)... a nova⁷ que *ela* está bem *pequeninho*... (MI54)

Diante do quadro apresentado, percebe-se que a marcação do gênero gramatical na variedade lingüística falada no dialeto do Paraná Tradicional rural é bastante oscilante e variada. No entanto, mesmo sendo a variação da concordância de gênero nas relações internas ao sintagma nominal (II) bastante expressiva, percebemos que a variação na relação sujeito/predicativo (IV) foi mais produtiva.

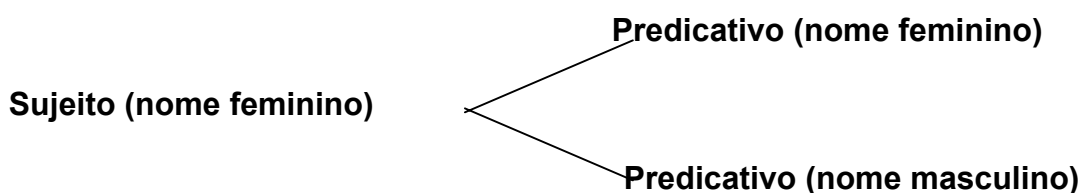
Isso nos motivou quanto à decisão de tomar como objeto de estudo para esta pesquisa, no conjunto das construções que apresentam variação de gênero no falar rural paranaense, estritamente as que envolvem uma relação sujeito/predicativo. Até então estávamos em dúvida se analisaríamos esta ou se tomaríamos como objeto de análise a *ausência de concordância de gênero* entre os elementos do SN. Na verdade, durante a fase inicial de delimitação dos dados que comporiam o *corpus*, chegamos a fazer um levantamento do número de ocorrências dos sintagmas nominais que não apresentavam *concordância de gênero* e obtivemos um percentual de 5% aproximadamente; contudo, na relação predicativa, apuramos um total de 40%.

Dentro desse total de 40% de *não concordância*, o aspecto que mais nos chamou atenção foi o uso, pelos falantes, de nomes masculinos representando núcleos de predicativos ligados, com freqüência, a sujeitos representados por nomes

⁷ Lua nova

femininos, o que foge à forma padrão. Ou seja, em nosso *corpus* de pesquisa não nos deparamos com construções cujo predicativo fosse representado por nomes de gênero masculino ligados a predicativos representados por nomes femininos.

Isto posto, apresentamos a seguinte configuração para o fenômeno variável da concordância de gênero na relação sujeito/predicativo no falar do Paraná Tradicional rural.



A seguir, seguem-se exemplos que ilustram os casos de relação de concordância nominal de gênero que constituem o *corpus* da presente pesquisa:

(11)... as *asas* por baixo é *branco*, as *costas* é *preto*... (FI52)

(12)... daí a *filha* dela era *pequena*... (FE42)

Em (11), os adjetivos *branco* e *preto*, em função predicativa, encontram-se flexionados no gênero masculino, mas se referem a nomes tradicionalmente designados pela norma gramatical como femininos - *asas* e *costas* -, o que aponta para um caso de ausência de concordância de gênero; em (12), o predicativo representado pelo adjetivo *pequena* se flexiona de acordo com o gênero do substantivo *filha*, sujeito da construção. Neste caso, a estrutura apresenta a concordância de gênero do português padrão.

De acordo com a gramática normativa, o predicativo pode ser representado por substantivo ou expressão substantivada, adjetivos, pronomes e numerais (CUNHA; CINTRA, 2001). No entanto, em nosso *corpus* de pesquisa não

registramos a ocorrência de numerais exercendo a função de predicativo. Em função disso, foram considerados contextos com núcleos de predicativo biformes, os quais se encontram representados por *substantivo* ou *expressão substantivada*, *pronomes* e *adjetivos*, inclusive as *formas participiais*, levando-se em conta que o participípio passado pode ocorrer em posição predicativa.

2 PRESSUPOSTOS TEÓRICOS

A Sociolinguística dispõe de diferentes modelos teórico-metodológicos para a análise da variação e da mudança. Na investigação do fenômeno linguístico em estudo nesta pesquisa nos fundamentamos na abordagem da Teoria da Variação, cujo precursor é William Labov. Esta linha é considerada como teoricamente coerente e metodologicamente eficaz para a descrição da língua em uso numa perspectiva sociolinguística.

A primeira seção do capítulo que se segue trata da apresentação de alguns fundamentos da Sociolinguística. A segunda seção abrange os pressupostos teóricos que orientam a pesquisa variacionista. Na última seção são descritos alguns aspectos que caracterizam a análise quantitativa, utilizada no exame dos dados.

2.1 Fundamentos da Sociolinguística

A área dos estudos sociolinguísticos envolve desde os problemas que surgem na comunicação entre os indivíduos, passando pela análise dos dialetos, idioletos, socioletos, questões de pronúncia e desenvolvimento linguístico, até barreiras linguísticas provenientes de concepções ideológicas, diferenças de raça e posição social, barreiras linguísticas estas que muitas vezes geram barreiras sociais. Afinal, a linguagem é condição e meio necessário para qualquer interação. Sem ela, não haveria formação de grupos e nenhuma atividade humana organizada. Porém, a linguagem não é só produto de integração social. Mas sim um sistema de sinais convencionais desenvolvidos, ensinados e apreendidos socialmente.

A ciência sociolingüística toma especialmente como objeto de estudo a variação e, segundo Mollica (2003, p. 10)

estuda a língua em uso nas comunidades de fala, voltando atenção para um tipo de investigação que correlaciona aspectos lingüísticos e sociais. Esta ciência se faz presente num espaço interdisciplinar, na fronteira entre língua e sociedade, focalizando precipuamente os empregos lingüísticos concretos, em especial os de caráter heterogêneo.

A variação, nesse caso, entendida como um fenômeno passível de ser descrito e analisado cientificamente, parte do princípio de que são os fatores sociais e estruturais que influenciam o uso de uma ou outra forma lingüística. Esses fatores também são denominados *variáveis independentes*, “no sentido que os usos de estruturas lingüísticas são motivados e as alternâncias se configuram, por isso, sistemáticas e estatisticamente previsíveis” (MOLLICA, 2003, p.10).

Ao traçar uma breve retrospectiva pelos estudos lingüísticos a partir do início do século XX, não se pode deixar de fazer referência a alguns estudiosos cujas obras já refletem a questão do social no campo da linguagem, tais como: Antoine Meillet, Mikhail Bakthin, Émile Benveniste e Roman Jakobson. Segundo Alkmim (2004), Meillet, Bakthin, Beneviste e Jakobson são consensuais na constatação de que a relação entre linguagem e sociedade é óbvia e complexa ao mesmo tempo. Entretanto, apresentam divergências quanto ao modo de tratar e explicar a questão da relação entre linguagem e sociedade.

As pesquisas realizadas por estes autores representam um marco divisor no campo da reflexão da Lingüística contemporânea. Após seu advento, aspectos ligados à linguagem, cultura e sociedade passam a ser considerados inseparáveis. Posteriormente, surgem estudiosos como John Gumperz, Eínar

Haugen, Dell Hymes, John Fischer, entre outros, que vieram a efetivar a caracterização e a definição da Sociolingüística como nova área de estudo. Todos se tornaram referências clássicas na tradição dos estudos focados na relação entre língua e sociedade.

Todavia, os propósitos de descrever a heterogeneidade lingüística e de encontrar um modelo capaz de dar conta da influência dos fatores sociais que atuam na língua somente passaram a ter êxito com os trabalhos de Labov, que, em 1963, publica seu conhecido trabalho sobre a comunidade da ilha de Martha's Vineyard, em Massachusetts. Nesse trabalho mostra que o papel dos fatores sociais é decisivo na variação lingüística observada na comunidade.

Em 1964, Labov conclui sua pesquisa sobre a estratificação social do inglês em New York, na qual apresenta um modelo de descrição e interpretação do fenômeno lingüístico no contexto social de comunidades urbanas. Este modelo, causador de grande impacto na Lingüística contemporânea, é conhecido como Sociolingüística Variacionista ou Teoria da Variação, também denominada por alguns de Sociolingüística Quantitativa, por operar com números e tratamento estatístico dos dados coletados (TARALLO, 1997, p. 8). A seção 2.2 tratará especificamente dessa vertente da Sociolingüística.

2.2 Teoria da Variação

Para Fasold (1984), a Sociolingüística trata da importância social da linguagem desde pequenos grupos sócio-culturais até comunidades maiores. Se

cada grupo apresentasse comportamento lingüístico idêntico, não haveria razão para se ter um olhar sociolingüístico sobre a sociedade.

A Sociolingüística, ao relacionar a atividade lingüística com a situação social do indivíduo e as condições ou meio em que este realiza suas experiências, reveste-se de cunho político-social, pois, segundo Alkmim (2004, p. 31), tem como “objeto principal o estudo da língua falada, observada, descrita e analisada no contexto social, em situações reais de uso”. Por conseguinte, sua investigação tem como ponto de partida a comunidade lingüística, “um conjunto de pessoas que interagem verbalmente e compartilham um conjunto de normas com respeito aos usos lingüísticos” (ALKMIM, 2004, p. 31).

É na comunidade que se percebe a diversidade ou a variação, pois toda comunidade se caracteriza pelo emprego de diferentes modos de falar. A esse respeito, argumenta Labov:

The procedures of descriptive linguistics are based upon the conception of language as a structured set of social norms. It has been useful in the past to consider these norms as invariants, shared by all members of the speech community. However, close studies of the social context in which language is used show that many elements of linguistic structure are involved in systematic variation which reflects both temporal change and extralinguistic social processes (apud MONTEIRO, 2000, p. 13)⁸.

Com isso, Labov remete a dois aspectos que devem ser considerados:

o primeiro se relaciona ao caráter eminentemente social dos fatos lingüísticos e o

⁸ “Os procedimentos da Lingüística se baseiam no entendimento de que a língua é um conjunto estruturado de normas sociais. No passado foi útil considerar que tais normas eram invariantes e compartilhadas por todos os membros da comunidade lingüística. Todavia, as análises do contexto social em que a língua é utilizada vieram demonstrar que muitos elementos da estrutura lingüística estão implicados na variação sistemática que reflete tanto a mudança no tempo quanto os processos sociais extralingüísticos” (Tradução de Monteiro, 2000, p. 13).

segundo se relaciona à percepção da variabilidade a que tais fatos estão submetidos. É necessário, então, que se descreva e analise essa variabilidade.

Desse modo, o pressuposto básico da Teoria da Variação Lingüística, (LABOV, 1972), é o de que a heterogeneidade, ou variação, é inerente a todo sistema lingüístico e não é aleatória, mas ordenada por restrições lingüísticas e extralingüísticas. E são essas restrições que levam o falante a usar certas formas e não outras quando faz uso da língua falada.

Sempre que duas ou mais maneiras de se dizer a mesma coisa, em um mesmo contexto, são utilizadas com certa freqüência e sistematicidade, em uma determinada comunidade de fala, acontece a variação (TARALLO, 1997). Para haver variação, portanto, é preciso avaliar se as diferentes possibilidades de expressão estão correlacionadas a determinados contextos estruturais específicos ou a dadas situações de uso da língua, de maneira sistemática e freqüente. Sobre tal abordagem, citamos as palavras de Naro (2003, p. 15):

O pressuposto básico do estudo da variação no uso da língua é o de que a heterogeneidade lingüística, tal como a homogeneidade lingüística, não é aleatória, mas regulada, governada por um conjunto de regras. Em outras palavras, tal como existem condições ou regras categóricas que obrigam o falante a usar categoricamente certas formas (*a casa*) e não outras (*casa a*), também existem condições ou regras variáveis que funcionam para favorecer ou desfavorecer variavelmente e com pesos específicos, o uso de uma ou outra das formas variáveis em cada contexto.

Conforme Mollica (2003, p. 11), *variantes* são formas lingüísticas inerentes à variação lingüística e caracterizam-se como formas alternativas que compõem um fenômeno variável, chamado *variável dependente*. A concordância nominal de gênero entre o sujeito e o predicativo, por exemplo, é uma variável lingüística (ou um fenômeno variável), pois se realiza por meio de duas variantes,

duas alternativas possíveis e semanticamente equivalentes: a *concordância* quanto ao gênero entre o sujeito e o predicativo ou a *ausência da concordância*. Uma variável é definida como dependente no sentido de que o emprego das variantes não é aleatório, mas influenciado por grupos de fatores (ou *variáveis independentes*) de natureza social ou estrutural.

O modelo de análise lingüística proposto por Labov e pesquisadores afins prevê, entre outros, um tratamento estatístico dos dados obtidos com o objetivo de desenvolver também um suporte quantitativo para subsidiar a análise. Em síntese, segundo Scherre (1998, p. 43):

Ao pesquisador variacionista cabe identificar os fenômenos lingüísticos variáveis de uma dada língua, inventariar suas variantes, definindo as variáveis dependentes, levantar hipóteses que dêem conta das tendências sistemáticas da variação lingüística, operacionalizar as hipóteses através de variáveis independentes ou grupos de fatores de natureza lingüística e não lingüística, identificar, levantar e codificar os dados relevantes, submetê-los a tratamento estatístico adequado e interpretar os resultados obtidos à luz das hipóteses levantadas.

Para a descrição e análise dos fenômenos lingüísticos de uma língua é necessário, pois, que se proceda ao levantamento de um número significativo de dados da língua. Neste sentido, faz-se necessário garantir uma coleta de dados de boa qualidade, que representem, o mais fielmente possível, o vernáculo de uma determinada comunidade de fala.

Tendo em vista que os fatores que atuam sobre o comportamento lingüístico dos falantes podem ser internos ou externos ao sistema lingüístico, uma análise, seguindo os moldes da Sociolingüística Variacionista, deve incorporar tanto os *fatores lingüísticos* quanto os *fatores sociais*. Os fatores lingüísticos são aqueles relacionados aos contextos da estrutura que condicionam o uso de uma ou outra

variante na comunidade onde a variação está relacionada, como os de natureza fono-morfo-sintáticos, os semânticos, os discursivos e os lexicais. Os *fatores sociais* ou *extralingüísticos* são determinados pelo sexo, escolaridade, idade, nível social e o lugar onde o informante mora. O importante, então, é o que possibilita a variação.

De acordo com Camacho

[...] o que a Sociolingüística faz é correlacionar variações existentes na expressão verbal a diversidades sociais, entendendo cada domínio, tanto o lingüístico, como o social, como fenômenos estruturados e regulares. Se um falante enuncia o verbo “vamos” como [vãmus] e outro falante o enuncia como [vãmu], podemos afirmar, com base nos postulados da Sociolingüística, que essa variação não é o resultado aleatório de um uso arbitrário e inconseqüente dos falantes, mas um uso sistemático e regular de uma propriedade inerente aos sistemas lingüísticos, que é a possibilidade de variação (CAMACHO, 2004, p. 50).

As variantes podem permanecer estáveis nos sistemas, alternando-se durante pouco tempo ou mesmo por séculos; também podem sofrer mudanças, quando uma das formas desaparece, e a outra prevalece.

Porém, nem todos os fatos da língua estão sujeitos a variações. Existem regras gramaticais conhecidas como categóricas, já que um falante não pode violá-las, como lembra Monteiro (2000, p. 58). Essas regras definem o que é e o que não é possível na língua. Ainda segundo Monteiro, em romeno, por exemplo, o artigo se pospõe ao nome, já em português e em outras línguas o artigo sempre antecede o nome e se houver qualquer alteração desse emprego, a construção se torna agramatical. Essa ordem não pode ser alterada pelo falante, sob pena de dificultar ou inviabilizar a compreensão dos enunciados. Esse conjunto de regras internas costuma ser denominado de *invariantes*.

2.3 A Análise Quantitativa

A análise quantitativa realizada pelo VARBRUL consiste na análise de cada ocorrência do fenômeno lingüístico em estudo, ou seja, examina cada escolha diferenciada e recorrente de uma estrutura ou forma lingüística feita pelo falante em seu desempenho, que não seja categoricamente previsível por informações contextuais. Daí se poder falar em *regras variáveis*. Estas seriam a expressão da escolha pelo falante entre dois ou mais sons, palavras ou estruturas alternativas distintas que expressam a relação entre elementos do ambiente lingüístico e do contexto social.

A análise quantitativa visa, assim, examinar a relação entre essas diferentes escolhas e os fatores a elas associadas.

Para a configuração de uma *regra variável* o pesquisador precisa vencer as seguintes etapas⁹:

- a) Definição da *variável dependente*;
- b) Definição das *variáveis independentes*;
- c) Delimitação da amostra e obtenção os dados;
- d) Transcrição e codificação os dados;
- e) Quantificação dos dados: a análise quantitativa permite obter os seguintes resultados:
 - (i) a seleção, em valores estatísticos, das *variáveis independentes* mais relevantes na produção do fenômeno analisado;

⁹ Para maiores detalhes acerca das etapas para configuração de regras variáveis consultar Brescancini

(ii) as freqüências de uso e o peso relativo correlacionados a cada um dos valores das *variáveis independentes*.

(iii) o nível de significância dos resultados obtidos.

f) Interpretação dos resultados: compreender e explicar os resultados numéricos é o estágio mais importante da análise. Muitas vezes implica reavaliar o sistema de codificação ou as hipóteses lingüísticas utilizadas.

Não se pode esquecer, durante a análise, que os valores numéricos relacionados aos fatores funcionam apenas como um direcionamento para chegar até as perguntas que motivaram a pesquisa, porém, não as responde. Para isso, o pesquisador deve ter uma boa fundamentação baseada na teoria lingüística e o conhecimento da estrutura social da comunidade em exame, os quais são elementos decisivos e indispensáveis na interpretação dos resultados numéricos.

Dentro desse modelo matemático, para fenômenos binários, pesos relativos acima do valor de *0,50* favorecem o uso da variante tomada como base, enquanto pesos relativos com valor abaixo de *0,50* desfavorecem-na; no caso desta pesquisa, tomamos como base a variante *ausência de concordância*. Os pesos relativos em torno de *0,50* são considerados neutros.

3 O GÊNERO NO PORTUGUÊS DO BRASIL

O capítulo três está organizado em cinco seções. Com base em alguns autores, é traçada uma caracterização do gênero como categoria gramatical; são abordados também alguns pontos da controversa questão de sua inserção entre os mecanismos de flexão ou derivação, compondo a primeira e segunda seção, respectivamente. Na terceira seção são abordados alguns aspectos conceituais a respeito da concordância de gênero entre o sujeito e o predicativo. A quarta seção é reservada à revisão de alguns trabalhos importantes da pesquisa sociolingüística e da pesquisa dialetológica sobre variedades populares e regionais do Brasil, com destaque para os aspectos relativos ao gênero gramatical.

Dedicamos a última seção à descrição de alguns trabalhos recentes que comportam estudos variacionistas sobre a concordância de gênero no português do Brasil.

3.1 O Gênero como Categoria Gramatical

Vários estudiosos das línguas concebem o gênero como uma organização dos nomes em classes. Para Lyons (1979, p. 299) o termo gênero deriva de uma palavra de sentido extremamente geral, derivada do latim – “*genus*” -, que indica *classe* ou *tipo*. Hockett (apud DETTONI 2003, p. 23) considera que “gênero são classes de nomes que se refletem no comportamento das palavras associadas”. Isto significa que, na língua portuguesa, as evidências para a existência

do gênero devem ser buscadas nos artigos e adjetivos que estiverem em relação de concordância com o nome. Segundo Câmara Júnior (1966, p. 5), é o artigo que melhor cumpre esse papel, pois “o gênero de um substantivo está na flexão do artigo que o determina ou pode determinar”.

Os estudiosos têm vinculado a concepção de gênero como classes em que os nomes se organizam aos reflexos dessa categoria gramatical na morfossintaxe das línguas que a possuem. Geralmente, esses reflexos são atestados por meio das relações de concordância que se estabelecem entre os nomes e os artigos e adjetivos que com eles se relacionam. Lyons (1979, p. 299) defende que, do ponto de vista gramatical, na maioria das línguas, os nomes são classificados em gêneros para dar conta da referência pronominal e da concordância do adjetivo. Nas palavras de Lucchesi (2000, p. 152),

a concordância é o mecanismo sintático que explicita a presença de gêneros em uma dada língua, permitindo, inclusive, que se determine o seu número. Diante disso, pode-se dizer que a análise da categoria gramatical do gênero não compreende apenas a sua estruturação no léxico de uma língua, como também sua expansão na estrutura da sentença, através do mecanismo de concordância. Este compreende tanto as relações de concordância entre os constituintes do Sintagma Nominal como as relações entre um Sintagma predicativo com o nome a que se refere.

No que diz respeito à estrutura lexical dos nomes em português, os gramáticos, em princípio, adotavam definições semânticas para o gênero. Atualmente, porém, a tradição gramatical define o masculino e o feminino como duas classes gramaticais nas quais os nomes da língua estão distribuídos (ROCHA LIMA, 1976; CUNHA; CINTRA, 2001). Câmara Júnior (1995), em seus estudos sobre a descrição tradicional de nossas gramáticas, informa que o gênero é a distribuição dos nomes em duas classes mórficas, o masculino e o feminino, não sendo coerente

uma definição baseada na associação entre gênero e distinção dos sexos, já que “o gênero abrange todos os nomes substantivos portugueses, quer se refiram a seres animais, providos de sexo, quer designem apenas ‘coisas’” (p. 88).

3.2 Flexão ou Derivação?

A inclusão do gênero entre os mecanismos de derivação ou de flexão não é um ponto pacífico entre os estudiosos da língua portuguesa. A Lingüística moderna vem propondo uma revisão dessa questão, contrariando a tradição gramatical que, geralmente, tem considerado o gênero nos nomes como um processo de flexão.

Não é objetivo desta dissertação um aprofundamento maior nessa questão; porém, julgamos importante focar o que alguns lingüistas têm a dizer sobre o assunto e tecer alguns comentários a respeito de suas conclusões.

A fim de se delimitar o que é flexão e o que é derivação, a distinção básica que tem sido apontada de modo recorrente é a de que a derivação está relacionada à criação de novas palavras que passam a integrar o léxico de uma língua, enquanto a flexão diz respeito à possibilidade de uma única e mesma palavra assumir várias formas, segundo o contexto sintático, sem alterar sua classe gramatical.

Tais considerações levam à questão do que seja a flexão, a qual, em função de sempre ter sido caracterizada como um processo próprio da palavra, não

pode ser desvinculada do mecanismo sintático da concordância. Para haver concordância, faz-se necessário que as palavras se flexionem para tomar formas adequadas. Em um sintagma como *as belas meninas* os traços semânticos [*masculino, singular*] se referem ao nome *menino*, porém sua expressão mórfica se estende a todos os constituintes flexionáveis do SN.

Rocha (1998, p.207) destaca que “a flexão deixa de ser uma característica da palavra, para ser um mecanismo morfossintático, que se serve ao mesmo tempo de regras morfológicas e sintáticas”.

Para Câmara Júnior (1966), em português, todos os substantivos têm um gênero determinado. O referido lingüista também acrescenta que

quer apareça, quer não apareça a flexão, todo nome, em cada contexto, será imperativamente *masculino* ou *feminino*. Assim, *livro*, *poeta*, *dente*, são exclusivamente masculinos, como são exclusivamente femininos *tribo*, *rosa*, *ponte*, com as mesmas vogais finais, respectivamente. Da mesma sorte, conforme o contexto, são masculinos ou femininos, sem mudança de forma, *artista*, *intérprete* ou *mártir*. (CÂMARA JÚNIOR, 1966, p. 4).

Câmara Júnior ainda admite haver flexão de gênero nos nomes portugueses em que ocorre a adição do sufixo flexional - *a* ao nome, com a supressão da vogal temática. Isso pode ser observado em casos como ‘lobo - lob-*a*’ e ‘menino - menin-*a*’ em que o autor atenta para o fato de que, nesses casos, a flexão ocorre, mas se trata de um traço redundante do gênero, pois sempre será acompanhada de concordância. Ele ainda prevê que sufixos que podem também fazer uma distinção de gênero, em português, são derivacionais, a exemplo de nomes como ‘imperador - impera-*triz*’ e ‘galo - gal-*inha*’.

Apesar dos critérios propostos por Câmara Júnior (1995) para distinguir flexão de derivação, convém considerar que muitos nomes da língua portuguesa nem chegam a possuir flexão de gênero, ao passo que outros a possuem como um traço redundante. Com isso, não se pode estabelecer padrões constantes de classificação, visto que a categoria de gênero nos nomes, do ponto de vista da flexão, não apresenta uma completa regularidade. Quanto a isso, Rocha (1998, p.196) tem a dizer:

apenas uma parte insignificante dos substantivos (aqueles que se referem a seres sexuados) pode receber uma marca morfológica distintiva de gênero (*gato/gata, abade/abadessa*, etc.). Segundo pesquisas realizadas (Rocha, 1981), 95% dos substantivos se referem a seres não-sexuados e 4,5% a seres sexuados. Mesmo assim, desses 4,5%, nem todos recebem uma marca morfológica de gênero, como *criança, cônjuge, homem, jacaré, selvagem*, etc. A quase totalidade dos substantivos em português não apresenta uma marca morfológica de gênero, ou seja, a quase totalidade pertence a um gênero único [...].

Ainda de acordo com Rocha (1998, p. 207), o que a gramática tradicional chama de flexão, nos nomes, é um processo de derivação, ou seja, aplica-se a uma regra de derivação que permite formar uma nova palavra.

A fim de definir nesta pesquisa nossa posição em relação ao que seja flexão e derivação, baseamo-nos na distinção, mais freqüentemente enfocada, de que a derivação se refere ao processo de formação de novos itens lexicais que passam a integrar o léxico, enquanto a flexão trata das diferentes formas que um mesmo item lexical pode assumir de acordo com as exigências do contexto sintático. Assim sendo, esclarecemos que neste trabalho nos concentraremos nas relações de concordância de gênero entre o sujeito e o predicativo tomando como base exclusivamente os contextos com núcleos de predicativos bifformes, ou seja, aqueles que, conforme descrito pela gramática normativa, são flexionados em gênero.

3.3 A Concordância Nominal de Gênero na Relação Sujeito/Predicativo

Para a tradição gramatical brasileira o português do Brasil apresenta dois tipos de concordância: a concordância verbal, que ocorre entre o verbo e o sujeito, e a concordância nominal, que se verifica em gênero e número entre o adjetivo e o pronome (adjetivo), o artigo, o numeral ou o particípio e o substantivo ou pronome a que se referem (BECHARA, 1986). Este tipo de concordância se dá entre os elementos constituintes do SN (aØ garotaØ bonitaØ/as garotas bonitas) e também entre o sujeito e o predicativo (elaØ estáØ cansadaØ/elas estão cansadas).

Para definir as variantes que iriam direcionar todo o processo de delimitação das construções componentes do nosso *corpus* de análise, seguimos a definição de concordância nominal de gênero atribuída pela gramática normativa. Com isso, a natureza do predicado na gramática tradicional que nos interessa é a do predicado nominal, pois é ele que contém o predicativo do sujeito.

Nem sempre os verbos de ligação funcionam apenas como elo entre o sujeito e o seu predicativo. Há verbos que se empregam ora como de ligação (ou copulativos), ora como significativos. Os exemplos de Cunha e Cintra (2001, p. 133) elucidam melhor a questão:

Estavas triste	Estavas em casa
Andei muito preocupado	Andei muito hoje
Fiquei pesaroso	Fiquei no meu posto
Continuaram silenciosos	Continuamos a marcha

Na coluna à esquerda, os verbos *estar*, *andar*, *ficar* e *continuar* funcionam como verbos de ligação; na coluna à direita, como verbos significativos. A

partir disso, deve-se sempre atentar para o valor que apresentam em determinado contexto lingüístico para classificá-los com acerto e, conseqüentemente, não incorrer em equívocos quanto ao devido valor do núcleo do predicativo.

3.4 A Variação do Gênero nas Variedades Populares do Português do Brasil

Já em 1920, surgem pesquisadores preocupados com a questão do gênero nas variedades populares do português do Brasil. É o caso de Amadeu Amaral, autor do livro “O dialeto caipira” e que foi o primeiro a estudar, no Brasil, um dialeto local à luz da Lingüística - no caso, o linguajar dos caipiras do interior de São Paulo. Sobre a concordância de gênero nesse dialeto, Amaral (1976, p. 70) registra em seu livro a variação na concordância dos adjetivos que se situam à direita do nome, e nos processos de predicação: “o adjectivo e o particípio passado deixam, freqüentemente, de sofrer a flexão genérica, sobretudo se não aparecem contíguos aos substantivos: *as crianças távum **quétu**, as criações ficarum **pestiado**”.* Amaral (1976, p. 73) ainda faz a seguinte observação no item referente aos pronomes:

Um facto que merece atenção, apesar de pertencer mais ao linguajar dos pretos boçais do que propriamente ao dialecto caipira: a invariabilidade genérica do pronome *êle*, junta à invariabilidade numeral. Quando se trata de indicar pluralidade, o pronome *ele* se pospõe ao artigo def. *os*, e tanto pode referir-se ao gênero masculino, como ao feminino: *osêle, zêle fôro zimbora - êles (ou elas) foram-se embora.*

Rodrigues (1974) faz uma descrição do dialeto caipira falado nas comunidades rurais e populares do Município de Piracicaba no Estado de São

Paulo. Em seu livro “O dialeto Caipira na Região de Piracicaba” encontramos construções que mostram a variação na concordância de gênero:

“A roupa meu nenhum servia”

“A mãe meu vinha tratar do menino”

“Sem dar dor nenhum”

“Aquele coisa estufado”

“Tenho natureza de gente novo”

“Avó meu lidava ali comigo” (RODRIGUES, 1974, p. 55).

Pelos exemplos acima pode-se perceber que, além dos adjetivos, a variação na concordância de gênero se estende para os possessivos e pronomes que se situam à direita do nome, podendo também atingir os determinantes que estão à esquerda do nome, como em “Aquele coisa estufado”. Com relação aos pronomes pessoais de terceira pessoa, que seguem a regra de flexão de gênero dos nomes, não há registros de variação.

Marroquim (1934, p. 100), em pesquisa denominada “A língua do Nordeste (Alagoas e Pernambuco)”, afirma que “há perfeita concordância de gênero entre o adjetivo e o substantivo”; e diz ainda não ter encontrado nenhuma ocorrência semelhante às que Amaral (1976) apresenta como típicas do dialeto caipira. Marroquim elenca apenas alguns exemplos de adjetivos que, apesar de geralmente uniformes no português, apresentam flexão mórfica de gênero no dialeto que ele pesquisou, tais como os pares opositivos “*monstro-monstra*” e “*severgonha-severgonho*” (MARROQUIM, 1934, p. 102).

Também Nascentes (1953), no seu “O linguajar carioca”, ao estudar a fala popular do Rio de Janeiro não indica a variação na concordância de gênero. É

verificada, apenas, uma diferença na fixação do gênero gramatical de alguns itens lexicais com relação à língua padrão. De acordo com o autor, “*cometa, sistema, fantasma, telefonema* são femininos; *alface, chaminé, cal* aparecem no masculino” (p. 83).

Veado (1982), em seu livro “Comportamento lingüístico no dialeto rural – MG”, analisa a língua falada na região de Januária, no vale do São Francisco, no norte de Minas Gerais. Trata-se de um trabalho baseado numa amostra de 45 informantes adolescentes e adultos, a maioria não escolarizada. Sobre a concordância de gênero, Veado (1982, p. 57) relata que

a concordância de gênero se processa no dialeto rural tal qual prevista pelos nossos gramáticos, e tal qual se processa no uso efetivo da língua em situação informal de comunicação. Isto é, o substantivo serve de referência: o artigo, o pronome, o numeral e o adjetivo assimilam o mesmo gênero do substantivo.

A autora conclui que, nesse dialeto rural, a regra de concordância de gênero ocorre “tal qual na língua ‘padrão’ escrita e oral” (p. 59).

Recentemente, os estudos lingüísticos dos dialetos rurais brasileiros foram contemplados com pesquisas voltadas para os dialetos falados por comunidades de descendentes de negros africanos, possíveis remanescentes de quilombos. Citamos, entre estes, os estudos de Careno (1991) e Petter (1999).

Careno (1991), em tese de doutorado denominada “A linguagem rural do Vale do Ribeira: a voz e a vez das comunidades negras”, pela UNESP, investigou uma dessas comunidades. A autora desenvolveu sua pesquisa junto a comunidades rurais localizadas no Vale do rio Ribeira, na região sul do interior de São Paulo, cuja população se caracteriza por uma concentração demográfica afro-brasileira bastante

elevada. De acordo com Careno, essas comunidades constituem remanescentes de quilombos e seus habitantes seriam descendentes diretos de escravos fugidos, falantes de um tipo de dialeto crioulo. Tal fato se refletiria na linguagem atual dos integrantes dessas comunidades, que apresentam traços desse antigo dialeto falado por seus ancestrais, os escravos.

Analisando a concordância de gênero na estrutura morfossintática desses falares, Careno (1991, p. 207) apresenta exemplos que revelam que a variação não se limita aos adjetivos à direita do nome, tal como previsto por Amadeu Amaral, atingindo também os determinantes que antecedem o nome. É o que demonstram os exemplos a seguir:

“Aqui o *nosso cumida* aqui é...”.

“Ali era *muito subida*”.

Petter (1999), no artigo “A linguagem do Cafundó: crioulo ou anticrioulo?” relata em sua pesquisa na comunidade do Cafundó, um bairro rural da cidade de Salto de Pirapora, a 150 km de São Paulo, a análise que fez da fala pertencente a um núcleo de descendentes africanos, a qual se caracteriza pela presença de um léxico de origem banto mesclado à estrutura morfossintática do português. O gênero gramatical é um dos aspectos morfossintáticos focalizados por Petter (1999, p. 113), que considera o adjetivo a categoria gramatical mais suscetível à variação em gênero, pois, quando anteposto ao nome, favorece a concordância; quando posposto, desfavorece-a. São destacadas algumas estruturas sem concordância como “*lenha moiado*” e “*coisa sério*”.

Por fim, citamos o artigo de Callou (1998) no qual a autora relata um trabalho muito interessante que realizou em uma comunidade denominada Mato

Grosso, no município de Rio de Contas, localizada no sertão do Estado da Bahia, cuja caracterização sócio-histórica não apresenta nenhuma relação com a presença de escravos de origem africana e seus descendentes. Nesse trabalho Callou confronta dados de dois *corpora* a fim de verificar possíveis mudanças em tempo real de curta duração no português brasileiro. Os dados do *corpus* 1 foram levantados por ocasião de sua dissertação de mestrado, em 1964, – “O falar de Mato Grosso (Bahia) – fonêmica, aspectos da morfossintaxe e do léxico” - pela UnB; o *corpus* 2 foi composto por dados retirados de registros magnetofônicos, realizados quase trinta anos depois, na década de 90, por Alan Baxter.

Quanto aos aspectos morfossintáticos que caracterizam esse dialeto, a autora descreve algumas flutuações na categoria de gênero que não se encaixam dentro dos padrões gramaticais da língua. São registradas algumas formas que, apesar de enquadradas no gênero feminino, passam ao masculino - gênero marcado pelo uso do artigo - como em “*o bronquite*”, “*o lebre*”, “*um coisa*” (CALLOU, 1998, p. 264-65). Também são apresentadas algumas estruturas com variação de gênero semelhantes às que têm sido registradas em outros dialetos, como “*as coisa muito barato*” e “*esse daqui é a mulher dele*”.

Pelo que foi exposto nos trabalhos abordados ao longo desta seção observamos que a variação do gênero gramatical em alguns dialetos populares do Brasil não se caracteriza por um padrão de variação regular e estável, pois as informações e conclusões que fornecem são, de certo modo, contrastantes.

Rodrigues (1974), Careno (1991), Petter (1999), e Callou (1998) registram casos onde há uma tendência para o apagamento das marcas do gênero feminino, sobretudo nos elementos à direita do núcleo, que permanecem

preferencialmente na forma não marcada, o masculino. Nascentes (1953) e Callou (1998) destacam uma alteração no gênero de algumas formas lingüísticas que, mesmo sendo normalmente femininas, passam a masculinas; porém, segundo Marroquim (1934) e Veado (1982), não há variação de gênero nos dialetos por eles pesquisados;

Já os estudos e pesquisas realizados, nas últimas décadas, sobre a variação na concordância de número no português falado do Brasil, são menos divergentes. As palavras de Scherre (1994) definem com clareza a extensão e a regularidade da variação na concordância de número na fala popular do Brasil:

O fenômeno da variação na concordância de número no português falado no Brasil, longe de ser restrito a uma região ou classe social específica, é característico de toda comunidade de fala brasileira, apresentando diferenças mais de grau do que de princípio, [...]. Dos trabalhos realizados, conclui-se, portanto, que o fenômeno da variação de número no português do Brasil pode ser caracterizado como um caso de variação inerente, tendo em vista que ocorre em contextos lingüísticos e sociais semelhantes e apresenta tendências sistemáticas de variação altamente previsíveis (SCHERRE, 1994, p. 38).

3.5 Pesquisas Variacionistas Sobre a Concordância de Gênero

3.5.1 Lucchesi (2000): A variação na concordância de gênero em uma comunidade de fala afro-brasileira: novos elementos sobre a formação do português popular do Brasil

Em tese de doutorado pela UFRJ, Lucchesi (2000) desenvolveu uma pesquisa envolvida com o debate sobre a relevância do contato entre línguas na

formação da realidade lingüística brasileira tomando como ponto de partida a questão crucial que envolve a polêmica sobre as origens crioulas do português popular do Brasil. Sua pesquisa estuda a variação na concordância de gênero registrada na variedade do português falada em Helvécia, comunidade formada, basicamente, por descendentes de escravos africanos, localizada no sul da Bahia, no município de Nova Viçosa.

A fim de demonstrar que a evolução interna do sistema do português não dá conta de explicar as alterações que ocorrem na morfologia flexional do português popular do Brasil, mas que estas requerem um exame das situações de contato entre línguas que se deram em determinados contextos históricos e sociais específicos, o autor realiza sua pesquisa em duas dimensões: um estudo vertical que se concentra na descrição da natureza e do estágio dos processos de variação e mudança lingüística na comunidade estudada, e um estudo horizontal, no qual são comparados os padrões de fala observados entre os falantes mais idosos da comunidade de Helvécia com os de outra comunidade rural, denominada Mato Grosso e situada no município de Rio das Contas, também pertencente ao Estado da Bahia.

Lucchesi levanta a hipótese de que “o contato entre línguas, sobretudo com as línguas africanas, teve um importante papel no desenvolvimento do português do Brasil, decisivamente em suas variedades populares” (Lucchesi 2000, p. 34).

O autor esclarece que as diferenças entre as duas comunidades estudadas se devem à maior ou menor frequência de variação na concordância de gênero que elas apresentam em decorrência das interferências mais ou menos

diretas do contato entre línguas. Isso faz com que Helvécia seja caracterizada por uma história de contato intenso e massivo, ao contrário da comunidade de Mato Grosso, cujo contato foi superficial.

A investigação teve como base um *corpus* constituído a partir da fala de 18 informantes da comunidade. Com o objetivo de permitir uma análise contrastiva entre os padrões de variação de Helvécia e os de Mato Grosso, organizou-se um *corpus* de controle, formado por entrevistas de oito informantes com mais de sessenta anos em cada uma das comunidades.

A análise variacionista foi feita a partir de um recorte que limita o exame da variação na concordância de gênero à estrutura do SN. Lucchesi observa que mesmo tendo estimado um nível de variação na concordância de predicativo maior do que o verificado na concordância no interior do SN optou por focar esta última em sua pesquisa. Tal escolha se justifica pelo fato de que a variação da concordância no interior do SN se constitui num fenômeno muito mais localizado do que a variação de gênero com o predicativo, só sendo observada em comunidades de fala brasileiras que passaram em sua formação por um massivo processo de transmissão lingüística irregular, o que atende com maior eficiência seu objetivo no sentido de estabelecer uma relação empiricamente motivada entre determinados fatos lingüísticos atuais do português popular do Brasil e o processo anterior de transmissão lingüística irregular.

No *corpus* analisado foram encontradas 4.023 ocorrências de SNs femininos, das quais apenas 179 representavam estruturas em que a regra de concordância de gênero não foi plenamente aplicada. Em termos de frequência relativa essas ocorrências correspondem a apenas 5% do total, o que “aponta para

um estágio final do processo de mudança de sentido na fixação do uso da regra de concordância de gênero na comunidade de fala de Helvécia-Ba” (LUCCHESI, 2000, p. 213).

Foi analisada a variação na concordância de gênero a partir de duas abordagens distintas: uma sintagmática ou não-atômica, na qual são consideradas as relações que unem todos os constituintes do SN, e a outra mórfica ou atômica, na qual a realização da categoria de gênero é tratada separadamente em cada constituinte do SN.

As variáveis lingüísticas, em número de oito, estabelecidas para a abordagem sintagmática foram: função sintática do SN; posição do SN em relação ao verbo; constituição sintagmática do SN; constituição morfológica do nome núcleo do SN; vogal temática do nome núcleo do SN; caracterização semântica do nome núcleo do SN; aplicação da regra de concordância de número no SN como um todo; forma fonética do SN.

O grupo de fatores *constituição sintagmática do SN* foi o que se mostrou estatisticamente mais significativo, superado somente pela variável social *faixa etária*. Com os resultados obtidos por meio da seleção desse grupo de fatores foi verificado que a estrutura mais simples, formada por um determinante e um nome núcleo do sintagma, é a que mais favorece a realização da concordância de gênero.

Para a abordagem mórfica, foram testadas as variáveis lingüísticas: *classe gramatical do constituinte; posição do constituinte em relação ao nome núcleo; saliência fônica; marcação do número no constituinte*. Nessa abordagem, o primeiro grupo de fatores selecionado pelo programa como estatisticamente

significativo foi a variável classe *gramatical do constituinte*; a variável social *faixa etária* foi selecionada em segundo lugar.

Para definir o quadro de variação presente na comunidade estudada, em termos de variação estável ou mudança em progresso, o autor elegeu as variáveis sociais: *faixa etária*; *sexo*; *nível de escolaridade*; *estada fora da comunidade*. A variável *faixa etária* foi dividida em três faixas distintas de forma a caracterizar três gerações dentro da mesma comunidade, o que permite observar os prováveis processos de mudança em curso no tempo aparente.

Dessas variáveis, a selecionada como estaticamente mais relevante foi a *faixa etária*, cujos resultados, segundo o autor, indicam um processo de mudança em progresso, uma vez que a variante inovadora – a aplicação da regra de concordância de gênero – é a mais expressiva entre os falantes mais jovens. Foram também selecionadas a variável *sexo*, com os homens liderando a mudança, e a variável *estada fora da comunidade*. Esses resultados levaram o autor a concluir que “o percentual de variação na ordem de cinco por cento, observado na comunidade como um todo, demonstra que o processo de mudança se encontra em um dos seus estágios finais” (LUCCHESI, 2000, p. 294).

Em seguida, Lucchesi estabelece uma comparação entre o quadro de variação observado em Helvécia e o da comunidade de Mato Grosso, cujo objetivo é demonstrar que a variação na concordância de gênero encontrada em Helvécia é consequência de um processo anterior de transmissão lingüística irregular produzido por um contato mais intenso entre línguas do que o que se deu em Mato Grosso.

Foi analisado um *corpus* constituído pela fala de oito informantes com mais de sessenta anos, de ambas as comunidades. Na comunidade de Helvécia, a

análise demonstrou que a variação na concordância de gênero na estrutura do SN é condicionada pelos mesmos fatores que se aplicaram à análise da amostra como um todo, incluindo as três faixas etárias, embora alcance agora um nível mais elevado (10%).

Foram identificados 1.456 sintagmas nominais femininos em Mato Grosso, dos quais somente 28 não apresentaram a realização de uma concordância plena, correspondendo a um índice de apenas 2%. A análise quantitativa dos dados da comunidade de Mato Grosso contou com as mesmas variáveis lingüísticas estabelecidas para a análise dos dados de Helvécia.

Entretanto, os resultados obtidos com a análise contrastiva são bem diferentes, levando o autor a observar que o quadro de variação sistêmica e estruturada que se verifica ainda hoje em Helvécia, contrariamente à variação aleatória encontrada em Mato Grosso, é o resultado do contato direto e intenso do português com as línguas africanas que se deu naquela primeira comunidade no período colonial.

Neste trabalho, a variação na concordância de gênero é tomada como um forte argumento a favor da hipótese de uma crioulização prévia do português no período colonial. Apesar desta abordagem não ser contemplada nesta dissertação, cujos objetivos fogem ao debate sobre as origens crioulas do português popular do Brasil, não podemos deixar de mencioná-la e descrevê-la por se tratar de uma pesquisa variacionista sobre a concordância de gênero que se constitui, a nosso ver, um estudo de fundamental importância e referência para todos os pesquisadores variacionistas que se dedicam a estudar o fenômeno da concordância.

3.5.2 Dettoni (2003): A concordância de gênero na anáfora pronominal: variação e mudança lingüística no dialeto da baixada cuiabana – MT

Dettoni analisou a variação na concordância de gênero na variedade do português falada na região da baixada cuiabana em sua tese de doutorado pela UFMG.

A baixada cuiabana é a região do Estado do Mato Grosso formada pela capital, Cuiabá, e municípios vizinhos que se desenvolveram, desde o período colonial, às margens dos rios Cuiabá e Paraguai e de seus afluentes e defluentes. É uma área que se distingue das outras regiões do Estado por suas peculiaridades nos costumes, na cultura popular local e, especialmente, na linguagem.

O foco central da análise é a marcação do gênero gramatical na relação anafórica, nas construções em que um nome feminino é, com frequência, retomado pela forma pronominal masculina *ele*, como em:

“*Bananinha*, dia que Rafael vim ele panha, porque *ele*, com esse frio, *ele* custa a madurar”.

“Uai, *mandioca*, a senhora casca *ela*, lava bem *lavadinho*. A senhora vai com *ele* no ralo, se não tem catitu, rela *ele* no ralo, daí a senhora imprensa a massa no tipiti...”

Para a análise da marcação do gênero na relação anafórica, tomada como variável sociolingüística, só foram consideradas as construções em que o antecedente do pronome em função anafórica era constituído por um nome classificado como feminino no léxico português, por serem estas as construções que apresentam variação na marcação do gênero. De acordo com Dettoni, em seu

corpus os nomes classificados como masculinos no léxico português são sempre retomados por um anafórico também na forma masculina, ou seja, não apresentam variação.

A amostra que compõe o *corpus* da pesquisa contém 882 dados e foi organizada a partir de entrevistas informais realizadas com dezoito falantes nativos, distribuídos em três faixas etárias (15 a 25 anos, 30 a 45 anos e acima de 60 anos) e três níveis de escolaridade (iletrados, quarta e oitava série).

Dettoni partiu da hipótese de que há marcação variável de gênero na variedade lingüística da baixada cuiabana, e que esta marcação variável seria condicionada pelo *traço de animacidade* que caracteriza o nome que funciona como núcleo em uma relação de concordância. Considerou-se também a hipótese da ocorrência de um processo de mudança lingüística em curso nesta variedade, decorrente do fato de a fala da geração mais jovem refletir um comportamento lingüístico diferenciado, com a perda gradual das estruturas com variação de gênero e o uso, cada vez mais freqüente, de construções típicas do português padrão.

As variáveis lingüísticas que o programa VARBRUL selecionou, como estatisticamente relevantes, foram: *tipo de referência do antecedente; distância entre antecedente e anafórico; grau de escolaridade; faixa etária; natureza morfológica do antecedente, presença x ausência de determinante junto ao antecedente.*

A variável *tipo de referência do antecedente*, selecionada como a variável estatisticamente mais relevante, levou a autora a refutar sua hipótese inicial de que nomes com o *traço [inanimado]* constituiriam o ambiente mais favorável ao uso de *ele* e a perceber que a referência genérica tende a favorecer a retomada pelo anafórico masculino.

Os valores obtidos com o programa estatístico, tanto em frequência quanto em pesos relativos, referentes ao efeito desta variável, sempre se mostraram estatisticamente significativos. Mesmo nos cruzamentos que foram feitos entre as outras variáveis e a variável *tipo de referência do antecedente*, os resultados obtidos sempre mantiveram a *referência genérica* como o fator mais favorecedor da retomada por *ele*.

Os dados levaram Dettoni a observar ainda que, além da *referência genérica*, as características morfológicas do antecedente, que pode ser constituído por um nome sem flexão de gênero, associadas à ausência de um determinante junto ao antecedente, também se revelaram fatores favorecedores da retomada pronominal por *ele*.

Um outro aspecto da pesquisa que merece ser destacado se refere à distribuição do fenômeno variável nas diversas faixas etárias. Foi constatado que, na fala da geração mais idosa, o fenômeno é mais generalizado, podendo abarcar antecedentes de características variadas. À medida que se caminha em direção a gerações menos adultas, restringem-se os contextos de ocorrência da variação: na fala dos informantes jovens, a ocorrência do fenômeno se apresenta mais restrita.

Com relação à variável *escolaridade*, a pesquisadora considera que o acesso, cada vez maior, dos falantes a níveis de escolaridade mais altos tem contribuído para a assimilação de formas de expressão mais prestigiadas da língua, com a conseqüente neutralização das marcas típicas da variedade regional.

A autora evidencia, entretanto, que a mudança na concordância de gênero no dialeto da baixada cuiabana não se traduz apenas por uma mudança

motivada por fatores sociais, mas também por um forte condicionamento lingüístico que vem ocorrendo no nível morfossintático e fonológico do dialeto.

3.5.3 Souza (2000): A concordância de gênero entre o sujeito e o predicativo na fala da comunidade quilombola da Caçandoca - SP

Dissertação de mestrado que aborda a descrição e a análise da concordância de gênero entre o sujeito e o predicativo, tendo como amostra de fala a língua portuguesa da comunidade negra remanescente de escravos da Caçandoca - município de Ubatuba, Estado de São Paulo. Seu objetivo principal foi construir um *corpus* considerável para futuras análises da fala dos habitantes da comunidade da Caçandoca, a fim de contribuir para um maior conhecimento da realidade cultural das minorias de origem africana e de reunir elementos que permitam definir a real contribuição desses segmentos na formação da língua falada no Brasil.

Para a análise, foram realizadas entrevistas com 24 informantes (12 homens e 12 mulheres) selecionados pelo critério de sexo e faixa etária.

O trabalho também seguiu o modelo variacionista da Sociolingüística Quantitativa, proposta por Labov (1972). Delimitados 2.413 predicativos, a quantificação dos dados foi processada pelo programa VARBRUL, que acusou um percentual geral em torno de 7% para a ausência de concordância de gênero.

O autor efetuou a descrição e a análise de cinco variáveis lingüísticas e três sociais. Foram selecionadas pelo programa como estatisticamente significativas

as variáveis: *característica formal do sujeito, tipo de estrutura do predicativo e núcleo do predicativo*. Entre as variáveis sociais, a *faixa etária* foi a mais relevante, ocupando a segunda posição na seqüência das mais significativas.

A variável *núcleo do predicativo* foi dividida em quatro fatores: *substantivo, adjetivo, pronome e numeral*. Pelos resultados atingidos com a aplicação do VARBRUL o autor constatou que o fator *substantivo* torna a aplicação da regra de concordância de gênero quase categórica, pois de 649 ocorrências, 634 são de concordância de gênero, resultando num percentual muito alto (98%). Os dados levaram Souza a confirmar sua hipótese inicial de que o adjetivo seria a categoria gramatical mais sensível à variação de gênero. Por meio dos resultados estatísticos, ele constatou que a variável apresenta grande efeito sobre todos os outros grupos de fatores da pesquisa, visto que os percentuais dos fatores *substantivo* (98%) e *pronome* (96%) são quase categoricamente de concordância.

Os números fornecidos pelo VARBRUL para a variável *faixa etária* demonstraram que a ausência de concordância de gênero é mais recorrente na fala dos informantes mais idosos.

O autor chega à conclusão de que a correlação das variáveis sociais e lingüísticas revela em que estágio se encontra o processo de aquisição da regra de concordância de gênero na comunidade de Caçandoca, e que isso talvez indique, retrospectivamente, um longo processo diacrônico, em que um sistema prévio de flexão de gênero é substituído gradualmente sob a influência da língua culta.

3.5.4 Karim (2004): A variação na concordância de gênero no falar da comunidade de Cáceres – MT

Dissertação de mestrado pela UNESP que focalizou as variantes lingüísticas *presença* e *ausência de concordância de gênero*. A autora avaliou que o fenômeno da variação na concordância de gênero na fala dos habitantes nativos da comunidade é um fato comum e correlacionado, de maneira significativa, a fatores *extralingüísticos*, em especial a idade e a escolaridade.

O *corpus* de análise foi constituído com base em amostras de fala obtidas por meio de entrevistas gravadas na cidade de Cáceres-MT, durante o ano de 2001. A amostra de fala é composta por 36 entrevistas realizadas com informantes da comunidade, feitas de acordo com as técnicas da pesquisa sociolingüística. Os fatores mais decisivos na aplicação da regra foram os *sociais*.

Para definir a variável dependente a autora considerou a atuação ou não do mecanismo de concordância de gênero no sintagma nominal e verbal somente em palavras femininas. Ela concluiu que o emprego da concordância, segundo a norma padrão está, principalmente, circunscrito ao nível do SN, onde foi encontrado um alto índice de *presença de concordância*. Foi observado, por outro lado, que o cacerense utiliza a regra de concordância no sintagma verbal (SV) com uma freqüência bem menor que no SN; dessa forma, a regra não-padrão –*ausência de concordância* - na relação entre o predicativo e o sujeito da oração é muito mais utilizada, atingindo 30% de aplicação da regra.

Ressalte-se que a variante *ausência parcial de concordância* não apresentou uma diferença significativa em relação aos casos de ausência total de

concordância, razão pela qual as ocorrências dessas duas variantes foram amalgamadas na discussão dos demais grupos de fatores.

Sobre o fator *sexo*, Karim (2004) verificou que não há diferenças significativas entre homens e mulheres quanto à variação de concordância de gênero no nível do SN. A autora acredita que os índices apresentados na variante *presença de concordância* possam refletir uma tendência à homogeneidade dos padrões de comportamento lingüístico da comunidade cacerense, nesse contexto. Todavia, no nível do SV a variante *ausência de concordância* apresentou um índice mais alto na fala dos homens do que na fala das mulheres, demonstrando, neste caso, que as mulheres estariam fazendo mais uso da forma padrão do que os homens.

O resultado obtido pelo fator condicionador *grau de escolaridade* vai no sentido esperado, ou seja, o que se verifica é que os falantes com grau de escolaridade mais alto aplicam mais a concordância do que aqueles que possuem apenas o 1º grau.

Quanto à variável *idade*, sua influência foi relevante e decisiva para que se pudessem detectar as variações. A análise dos resultados obtidos ofereceu meios de verificar a forte tendência de os informantes da faixa etária mais velha utilizarem a forma não-padrão. Com isso, observa-se um contraste bem marcado entre as duas faixas mais jovens e a terceira faixa, dos mais velhos. Enquanto os primeiros se mantêm relativamente próximos aos índices gerais, os falantes mais velhos se distanciam significativamente daqueles resultados, revelando uma fala marcada fortemente pela variante *ausência de concordância*.

Karim conclui que os resultados obtidos na pesquisa mostram como os fatores sócio-econômicos e culturais criam as condições para o início de um nivelamento lingüístico, que tende a eliminar a marca característica da região, em função do padrão de realização da regra de concordância de gênero.

Ao finalizar este capítulo não podemos deixar de salientar a importância das pesquisas aqui citadas para o estudo das diversas variedades dialetais do português brasileiro. Conforme já mencionado na introdução, ainda são poucos os estudos que têm a variação da concordância de gênero como objeto de suas análises, entretanto, acreditamos que tais estudos possam contribuir de forma relevante no sentido de se obter um perfil mais completo da diversidade da língua portuguesa no Brasil.

Destacamos ainda que as informações e conclusões de Souza (2000), Lucchesi (2000), Dettoni (2003) e Karim (2004) em suas pesquisas foram fundamentais para a o desenvolvimento desta dissertação, pois serviram como uma fonte muito rica de pesquisa que nos norteou em muitos momentos, desde a postulação das hipóteses de trabalho até a descrição e análise dos dados.

Apesar de abordarem a variação na concordância de gênero em situações distintas e em diferentes variedades do português brasileiro, todos os referidos pesquisadores assinalam em seus trabalhos que a correlação de fatores sociais e lingüísticos revela um processo de mudança que aponta para a fixação da regra de concordância de gênero em que o fenômeno da variação vai sendo substituído gradualmente pela aquisição da regra de concordância sob a influência da língua padrão.

4 PROCEDIMENTOS METODOLÓGICOS

De acordo com Scherre e Naro (2003, p. 147) muitos pesquisadores variacionistas utilizam dados lingüísticos gravados especialmente para a pesquisa em que estão trabalhando no momento, ou seja, já vão para o campo gravar os dados sabendo qual assunto irão estudar e selecionam seus informantes de forma a compor uma amostra estratificada, a qual é mais indicada para pesquisas variacionistas. Entretanto, outras vezes, devido a razões de ordem prática, decidem-se por um *corpus* previamente levantado e transcrito, sem atenção especial para o fenômeno escolhido como objeto de estudo da pesquisa, o que é o caso da presente pesquisa¹⁰.

Conforme mencionado anteriormente, para descrever e analisar um fenômeno de variação lingüística, detectado no falar do homem do campo proveniente das regiões pertencentes ao Paraná Tradicional, adotamos uma metodologia que se sustenta na perspectiva da Sociolingüística Variacionista Laboviana, desenvolvida pelo pesquisador norte-americano William Labov. Como suporte para a análise quantitativa dos dados utilizamos o Pacote Computacional VARBRUL, desenvolvido por Pintzuk (1988).

Na amostra de fala dos informantes do ALPR (AGUILERA, 1990b) que analisamos, abordamos a concordância de gênero na relação entre o sujeito e os predicativos biformes.

¹⁰ Os dados que serviram de fonte para esta dissertação foram retirados das entrevistas coletadas e organizadas por Aguilera (1990a, 1990b, 1990c) para a elaboração do ALPR, publicado em 1994, conforme mencionado na introdução, p. 18-19.

Dada à especificidade de cada etapa deste trabalho, dividimos o capítulo que se segue em cinco seções: pontos lingüísticos de origem do informante; informantes – contendo dados característicos às suas idades, ao sexo e à escolaridade; a definição da variável dependente; as variáveis independentes e o levantamento das hipóteses de trabalho e a codificação e quantificação dos dados.

4.1 Os Pontos Lingüísticos

Para a elaboração do ALPR, Aguilera (1990a, 1990b, 1990c, 1994) estabeleceu dentro do Estado 65 pontos lingüísticos. Em cada ponto foi aplicado, durante os anos de 1985 a 1989, um questionário contendo 325 questões que abrangiam dois grandes campos semânticos: Terra e Homem. Ao todo são 130 entrevistas (um homem e uma mulher) em cada ponto, com duração entre 1h30 minutos às 3h, dependendo do informante.

Nossa intenção inicial era trabalhar com todas as entrevistas dos 65 pontos lingüísticos, porém, antecipadamente, já sabíamos que a quantidade de dados seria muito grande. Assim, o primeiro recorte que fizemos foi reduzir a investigação aos pontos que correspondem ao Paraná Tradicional, segundo divisão proposta por Cardoso e Westphalen (1986), já explicitada no capítulo 1.

Entretanto, apesar de optarmos apenas pelos pontos pertencentes ao Paraná Tradicional, contávamos, ainda com 32 pontos, ou seja, 64 entrevistas – 32 de informante feminino e 32 de informante masculino. Fizemos um segundo recorte e passamos a contar com 40 entrevistas a serem analisadas. Para isso,

reduzimos a investigação a 20 pontos ao escolher aqueles correspondentes aos municípios pólo de cada microrregião pertencente ao Paraná Tradicional, conforme figura 1.



Figura 1 – Os 20 pontos Lingüísticos selecionados para a pesquisa

Como se vê no mapa, os pontos lingüísticos de origem dos informantes são representados pelos municípios: Jacarezinho (ponto 6); Ibaiti (ponto 18); Jaguariaíva (ponto 26); Cerro Azul (ponto 37); Pitanga (ponto 41); Prudentópolis (ponto 42); Ponta Grossa (ponto 43); Antonina (ponto 45); Guaraqueçaba (ponto 46); Guarapuava (ponto 51); Irati (ponto 52); Curitiba (ponto 54); Paranaguá (ponto 55); Pato Branco (ponto 58); Palmas (ponto 60); União da Vitória (ponto 61); São Mateus do Sul (ponto 62); Lapa (ponto 63); Rio Negro (ponto 64); Guaratuba¹¹ (ponto 65).

Quando iniciamos a pesquisa tínhamos como objetivo incluir os pontos lingüísticos de origem do informante como uma das *variáveis independentes* a fim de constatar se o *fator geográfico* teria alguma influência sobre a variação. Porém, ao iniciar a delimitação das ocorrências percebemos que em muitos pontos lingüísticos não havia estruturas que representassem a variante *ausência de concordância*. Tal constatação nos levou a não incluir os pontos lingüísticos na quantificação dos dados, pois a falta destes para um dos fatores da variável representada pelos pontos lingüísticos acarretaria vários “*Knockouts*”¹², o que comprometeria a pesquisa. Isso não descarta o valor do *fator geográfico* no fenômeno de variação que abordamos neste trabalho.

¹¹ Os municípios de Antonina, Guaraqueçaba e Guaratuba não são municípios pólo de sua microrregião. Eles fazem parte da microrregião que tem Paranaguá como pólo, porém foram selecionados para fazerem parte do *corpus* devido ao fato de terem sido criados ainda no século XVII, no início da ocupação do Paraná, que se inicia por seu litoral.

¹² “*Knockouts*” são fatores que constituem o ambiente de 100% ou de 0% de aplicação da regra, por isso são desconsiderados no arquivo de condições. Para maiores detalhes verificar Brescancini (2002, p. 33).

4.2 Os Informantes

O quadro a seguir representa dados referentes aos informantes e seus pontos lingüísticos de origem:

Ponto	Sexo	Escolaridade	Idade
06	feminino	Analfabeto	35
06	masculino	Analfabeto	43
18	feminino	Analfabeto	38
18	masculino	Analfabeto	41
26	feminino	Escolarizado (1 a 4 anos)	46
26	masculino	Escolarizado (1 a 4 anos)	44
37	feminino	Analfabeto	38
37	masculino	Analfabeto	39
41	feminino	Analfabeto	41
41	masculino	Analfabeto	40
42	feminino	Analfabeto	40
42	masculino	Escolarizado (1 a 4 anos)	53
43	feminino	Escolarizado (1 a 4 anos)	38
43	masculino	Analfabeto	46
45	feminino	Analfabeto	52
45	masculino	Analfabeto	54
46	feminino	Escolarizado (1 a 4 anos)	59
46	masculino	Analfabeto	35
51	feminino	Analfabeto	46
51	masculino	Escolarizado (1 a 4 anos)	40
52	feminino	Analfabeto	46

52	masculino	Analfabeto	44
54	feminino	Escolarizado (1 a 4 anos)	42
54	masculino	Escolarizado (1 a 4 anos)	52
55	feminino	Analfabeto	42
55	masculino	Escolarizado (1 a 4 anos)	48
58	feminino	Escolarizado (1 a 4 anos)	50
58	masculino	Escolarizado (1 a 4 anos)	47
60	feminino	Analfabeto	59
60	masculino	Analfabeto	56
61	feminino	Escolarizado (1 a 4 anos)	53
61	masculino	Analfabeto	51
62	feminino	Escolarizado (1 a 4 anos)	50
62	masculino	Analfabeto	49
63	feminino	Escolarizado (1 a 4 anos)	48
63	masculino	Analfabeto	46
64	feminino	Analfabeto	54
64	masculino	Analfabeto	49
65	feminino	Analfabeto	49
65	masculino	Analfabeto	49

Quadro 1 - Ponto lingüístico de origem do informante e as Variáveis Sociais Sexo, Idade e Escolaridade

O quadro 2 apresenta o número de informantes divididos por sexo de acordo com a faixa etária à qual pertencem:

Faixa Etária/Sexo	Feminino	Masculino	Total
Faixa I: 30-46 anos	11	10	21
Faixa II: 47-65 anos	9	10	19

Quadro 2: Distribuição dos informantes por faixa etária e sexo

Embora o número de informantes com idades de 30 a 46 anos (21) - *faixa etária I* - seja maior que o número de informantes de 47 a 65 anos (19) - *faixa etária II*, a diferença entre os dois grupos é pouco representativa (2).

Escolaridade/Sexo	Feminino	Masculino	Total
Analfabeto	12	14	26
Escolarizado	8	6	14

Quadro 3: Distribuição dos informantes por escolaridade e sexo

Nota-se, pelo quadro acima, que a maioria dos informantes é analfabeta (26); os outros 14 informantes, classificados como escolarizados, são aqueles que concluíram ou tiveram algum contato com o ensino básico, ou seja, até a quarta série do ensino fundamental¹³. O quadro ainda nos mostra que há mais informantes analfabetos do sexo masculino.

Sobre o critério que adotamos na distribuição das variáveis *grau de escolaridade* e *faixa etária* (cf. quadros 2 e 3) temos a ressaltar que tomamos como base a pesquisa de Gusmão (2004, p. 87). A autora esclarece em seu trabalho que, de início, pretendia trabalhar com quatro níveis de grau de escolaridade - analfabeto, MOBREAL, primário incompleto e primário completo - e quatro níveis de faixa etária - até 30 anos, 31-40 anos, 41-50 anos, mais de 50 anos - mas, em virtude de uma assimetria na distribuição dos dados entre os fatores destes grupos, precisou efetuar amalgamações nas duas variáveis a fim de conferir uma maior ortogonalidade aos dados.

¹³ Só trabalhamos com dois níveis de escolaridade devido ao fato de que todos os informantes inquiridos para a elaboração do ALPR terem cursado, no máximo, até a quarta

Gusmão destaca que obteve a confirmação de que esse seria o melhor tratamento a ser dado às variáveis *escolaridade* e *faixa etária* por meio do cruzamento entre os fatores de ambas, do qual resultaram várias células vazias, o que não seria ideal para o cálculo dos pesos relativos.

Levando em consideração que mais da metade dos nossos informantes – vinte e um deles - eram os mesmos, já que, em sua pesquisa, conforme citado na introdução deste trabalho, Gusmão também utilizou dados obtidos através das entrevistas do ALPR (AGUILERA, 1990a), optamos por adotar o mesmo critério da referida autora. Porém, decidimos distribuir a variável *faixa etária* em três níveis (diferentemente dela, que usou quatro níveis, pois contava com informantes com menos de 30 anos, o que não era nosso caso – todos os informantes tinham mais de 30 anos).

Com a *faixa etária* composta por três níveis e o *grau de escolaridade* por quatro, cruzamos as duas variáveis. A exemplo da pesquisa de Gusmão, deparamo-nos com o mesmo problema, o que reforçou nossa opção. No cruzamento das variáveis não verificamos a existência de células vazias, todavia a faixa etária de 31-40 anos contava com vinte e dois informantes, a faixa de 41-50 com seis e a faixa com mais de 50 com doze. O *grau de escolaridade* totalizava vinte e seis informantes analfabetos, dez com primário incompleto, dois com primário completo e dois com MOBREAL.

Essa irregularidade na distribuição dos dados surgiu em razão de não trabalharmos com uma amostra aleatória estratificada que é a mais indicada para pesquisas que seguem a linha variacionista.

4.3 A Variável Dependente

Definidas as 40 entrevistas a serem analisadas, partimos para o levantamento de todas as ocorrências que envolviam o fenômeno em estudo. Dentre a extensão e riqueza de dados que podem ser investigados, abordamos somente a concordância nominal de gênero, a qual não se trata de um fenômeno comum à grande maioria das variedades regionais do país e é objeto de poucos estudos dentro do campo da pesquisa lingüística, (cf. p. 26).

No capítulo 1 (seção 1.2) são descritos os traços mais típicos quanto à variação do gênero gramatical na variedade do português falada no Paraná Tradicional rural, organizados em três grupos de construções sintáticas, configuradas da seguinte maneira:

- (i) *na relação sujeito/predicativo;*
- (ii) *nas relações internas ao sintagma nominal;*
- (iii) *nas relações anafóricas.*

Porém, para a descrição e análise da variação na concordância de gênero nesta dissertação, enfocamos apenas o grupo (i), selecionando todos os casos de concordância e ausência de concordância envolvendo o sujeito e o

predicativo da construção por ter sido este grupo que, entre os três, apresentou um índice maior de casos de ausência de concordância - 40%.

A *variável dependente* a ser analisada foi constituída pela atuação ou não do mecanismo de concordância de gênero na relação entre o sujeito e o predicativo, atendo-se apenas aos contextos com núcleos de predicativos biformes, ou seja, aqueles que, segundo a gramática normativa, flexionam-se em gênero.

A partir da observação da ocorrência das variantes no material pesquisado, selecionamos as variáveis lingüísticas: *ausência de concordância de gênero/presença de concordância de gênero*. Os casos em que o sujeito concorda em gênero com o núcleo do predicativo foram classificados como presença de concordância de gênero, tal como ilustram os exemplos a seguir:

sujeito (nome feminino) → predicativo (nome feminino):

(13)... a *mulher* era muito *malvada*, né... (FE48)

(14)... a *grama* está *orvalhada*... (MI48)

Nos exemplos acima, o falante aplicou o mecanismo da concordância de gênero entre o sujeito e o predicativo.

Os casos de ausência de concordância podem ser ilustrados com os seguintes exemplos:

sujeito (nome feminino) → predicativo (nome masculino):

(15)... chuva deu pedra, tá muito *frio*, essa *chuva*... (FI46)

(16)... essa *doença* é *conhecido*... (MI44)

4.4 As Variáveis Independentes

Conforme a perspectiva teórico-metodológica da Teoria da Variação (MOLLICA, 2003) definimos como *variáveis independentes* as variáveis sociais *sexo, grau de escolaridade e faixa etária* (quadro 1, p. 76) e as variáveis lingüísticas *tipo de referência do sujeito, característica formal do sujeito, verbo utilizado, núcleo do predicativo, ordem dos elementos na construção, elementos intervenientes entre o verbo e o predicativo*. A seguir, descrevemos os fatores que compõem cada variável testada e as motivações as estabeleceram.

As variáveis lingüísticas propostas para a análise dos dados foram levantadas em função das características morfossintáticas da categoria de gênero em português. O estabelecimento dessas variáveis foi motivado também, em parte, pela sugestão de trabalhos que analisam a variação na concordância nominal em outras variedades do português do Brasil. Citamos, entre eles, os trabalhos de Souza (2000) e Dettoni (2003) - descritos na seção 3.5 - sobre a variação na concordância de gênero e o trabalho de Scherre (1991), "A concordância de número nos predicativos e participios passivos". Apesar de não abordar a variação na concordância de gênero, o trabalho de Scherre norteou o levantamento de alguns grupos de fatores de nossa pesquisa, os quais foram reproduzidos e testados com o intuito de verificar o alcance que teriam com relação à ausência da concordância de gênero.

A variável *tipo de referência do sujeito* se baseia na expectativa de que a *referência genérica do sujeito* favorece a ausência de concordância. A partir da constatação de que núcleos de predicativos representados por nomes masculinos eram ligados, com frequência, a sujeitos representados por nomes femininos e da pesquisa a autores como Martin (1975) e Câmara Júnior (1995), que chamam a atenção para a tendência de o masculino ser, de modo geral, a forma usada para os

casos em que se faz uma referência genérica, levantamos a hipótese de que tal tendência poderia exercer influência na escolha do falante em fazer uso de estruturas sem concordância de gênero ao se referir de uma forma genérica ao nome que representa o sujeito.

O grupo de fatores *característica formal de sujeito* foi estabelecido com a finalidade de avaliar se a presença do sujeito exerceria alguma influência sobre a ausência da concordância. A partir de observações iniciais sobre os dados levantados, acreditamos, inicialmente, que o *sujeito não preenchido* tende a favorecer a ausência da concordância de gênero. Para a composição desta variável nos baseamos em Souza (2000) e definimos dois grupos de fatores: *sujeito preenchido* e *sujeito não-preenchido*. Este último foi aquele que encontramos no contexto, porém sem estar explícito.

Por meio da variável *verbo utilizado* objetivamos verificar se o tipo de verbo de ligação interferia na escolha que o informante fazia entre a variante ausência x *presença de concordância*. Para esse fim, subdividimos a variável em três fatores: *verbo ser*, *verbo estar*, *verbo ficar*.

A variável *núcleo do predicativo* foi subdividida em *substantivos*, *pronomes (demonstrativo, possessivo, pessoal)*, *adjetivos* e *formas participiais*. Ao propor esta variável, pretendíamos confrontar nossos resultados com as conclusões de Souza (2000), que demonstrou em sua pesquisa que as *formas participiais* em construções predicativas favorecem a ausência de concordância.

A variável *ordem dos elementos na construção* seguiu a mesma divisão proposta por Scherre (1991), ficando composta por dois fatores: *ordem canônica* (sujeito+verbo+predicativo) e *ordem não canônica* (inversão de posição entre o

sujeito, o verbo e o predicativo na construção). Com a definição desta variável pretendíamos confirmar se nossas impressões iniciais, ao fazer o recorte para compor os dados da presente pesquisa, se confirmariam no sentido de que a *ordem não canônica* favoreceria ausência de concordância de gênero.

Ao levantar variável *elementos intervenientes entre o verbo e o predicativo* procuramos testar se esses elementos exerciam alguma interferência na relação sujeito/predicativo. Formulamos a hipótese de que as estruturas presentes entre o verbo e o predicativo, em decorrência do fato de se postarem logo após o verbo, distanciariam conseqüentemente o sujeito (ou a idéia deste) do predicativo, o que poderia influenciar a escolha do falante por uma ou outra variante. Os fatores determinados foram: *presença de elemento interveniente* e *ausência de elemento interveniente*.

As variáveis sociais *sexo, grau de escolaridade e faixa etária do informante* (cf. capítulo 4, seção 4.2) foram estabelecidas no sentido de avaliar se a produção do fenômeno em estudo apresenta uma configuração comumente interpretada como um padrão de variação estável em que, de acordo com Labov (1981) as mulheres realizam mais as formas de prestígio do que os homens; quanto maior o grau de escolaridade, maior o favorecimento das formas de prestígio; a idade mais avançada desfavorece-as. Sendo assim, a expectativa para estas variáveis é a de que falantes masculinos, com menor grau de escolaridade e mais idosos tendem a apresentar uma maior freqüência para a variável *ausência de concordância*.

4.5 Codificação e Quantificação dos Dados

Conforme citado anteriormente, nesta pesquisa empregamos o programa estatístico VARBRUL como suporte para a análise quantitativa dos dados, desenvolvido por Pintzuk (1988). No entanto, optamos pela versão mais recente desse programa, o *Goldvarb* (ROBINSON; LAWRENCE; TAGLIAMONTE, 2001) que é uma aplicação para análise multivariacional baseada no programa previamente circulado - *Goldvarb 2.0* (RAND; SANKOFF, 1990).

A preparação dos dados a serem analisados, a tabulação cruzada de duas *variáveis independentes* e a análise da *regra variável* são passos indispensáveis para a análise quantitativa e podem ser obtidos usando somente duas janelas, “*Token*” e “*Results*”, e seus respectivos menus.

Como trabalhamos com o programa VARBRUL, atribuímos códigos identificáveis pelo programa a todas as ocorrências que delimitamos na amostra de fala pesquisada, de acordo com cada variável previamente estabelecida, digitando-os em um arquivo apropriado.

De acordo com Brescancini (2002, p. 25), o programa toma um conjunto de dados lingüísticos, organizando-os de acordo com a *variável dependente*, em “ambientes possíveis” do ponto de vista lingüístico e extralingüístico. Em seguida, por meio de um algoritmo, oferece informações estatísticas para cada fator condicionador de uma *regra variável*. Os símbolos utilizados na codificação dos dados foram os seguintes:

Variável dependente

P - Presença de concordância de gênero

A - Ausência de concordância de gênero

1 *Tipo de referência do sujeito*

% - Referência genérica

\$ - Referência específica

2 *Característica formal do sujeito*

T - Sujeito preenchido

N - Sujeito não-preenchido

3 *Verbo utilizado*

q - Verbo ser

r - Verbo estar

s - Verbo ficar

4 *Núcleo do predicativo*

* - Substantivo

- Adjetivo

£ - Pronome (demonstrativo, possessivo, pessoal)

@ - Formas participiais (Particípio passado)

5 *Ordem dos elementos na construção*

b - Ordem canônica

a - Ordem não canônica

6 *Elementos intervinientes entre o verbo e o predicativo*

u - Presença de elemento interviniente

v - Ausência de elemento interviniente

7 *Sexo do informante*

M - feminino

H - masculino

8 *Grau de escolaridade do informante*

1 - analfabeto

2 - escolarizado

9 *Faixa etária do informante*

3 – Faixa etária I: de 30 a 46 anos

4 – Faixa etária II: de 47 a 65 anos

5 ANÁLISE DOS DADOS

O *corpus* estabelecido para esta pesquisa é formado por 462 construções que representam a relação sujeito/predicativo. Nessas construções será analisada a variação da concordância nominal de gênero que se estabelece entre o sujeito e o seu predicativo. Para tanto foram analisadas 40 entrevistas de informantes inquiridos na pesquisa de campo do Atlas Lingüístico do Paraná - ALPR - (AGUILERA, 1990a, 1990b, 1990c). Os informantes, de ambos os sexos e distribuídos em duas faixas etárias, representam regiões pertencentes ao Paraná Tradicional, conforme descrito na metodologia.

Este capítulo apresenta os resultados estatísticos das ocorrências do fenômeno lingüístico em estudo, rodados por meio do programa VARBRUL, a análise dos dados com base na teoria da Sociolingüística Variacionista, bem como exemplos referentes ao fenômeno tomado como objeto de estudo.

5.1 Resultados da Análise Quantitativa

Antes de iniciarmos a análise dos dados, achamos oportuno esclarecer que o produto final do processamento do programa VARBRUL também é

comumente chamado de rodada por aqueles que o utilizam. Uma rodada conta com dois tipos de análise: a primeira é conhecida como análise unidimensional e apresenta as porcentagens de aplicação da regra para cada uma das variáveis codificadas; a segunda, denominada análise multidimensional oferece como produto final os resultados em peso relativo para cada fator de uma variável.

No final da análise, o programa seleciona as variáveis independentes estatisticamente relevantes.

Uma vez estabelecidos os grupos de fatores, procedeu-se a uma primeira rodada no programa VARBRUL. Convém destacar que os resultados estatísticos que se seguem foram lidos a partir da segunda rodada do programa, já que a primeira mostrou a necessidade de se eliminar o fator *pronomes* da variável *núcleo do predicativo*, o qual foi desconsiderado no processamento do programa por aparecer apenas nas construções em que havia concordância de gênero, ou seja, o fator constituiu um ambiente de 100% de aplicação da regra de concordância (28 ocorrências) contra 0% de não aplicação, caracterizando um “*Knockout*”.

Com a segunda rodada, o VARBRUL selecionou, por ordem de relevância, as variáveis:

- 1) *Núcleo do predicativo*;
- 2) *Ordem dos elementos na construção*;
- 3) *Característica formal do sujeito*;
- 4) *Sexo*;
- 5) *Tipo de referência do sujeito*.

A tabela 1 apresenta os resultados das freqüências que constituem um quadro geral da variação na aplicação da regra de concordância entre o sujeito e o

predicativo como um todo, sem ainda estarem especificadas as freqüências relativas dos fatores, as quais serão analisadas na próxima seção.

TABELA 1: Número de Ocorrências e Percentuais das Variantes da CG¹⁴

Variável Dependente	Total de Ocorrências	Freqüência
Ausência de CG	186	40%
Presença de CG	276	60%
Total	462	100%

Como se vê na tabela, na linguagem do Paraná Tradicional rural a freqüência da presença de concordância de gênero é maior do que a freqüência da ausência de concordância.

A seguir, apresentamos e analisamos os dados fornecidos para cada grupo de fatores a fim de atender nosso interesse maior que é estabelecer os fatores lingüísticos e sociais que determinam a escolha de uma ou outra variante pelo falante, visto que os dados expostos na tabela acima possibilitam apenas uma visão geral do uso da concordância de gênero no falar rural paranaense.

5.1.1 Variáveis lingüísticas estatisticamente significativas

Para a leitura das tabelas das variáveis estatisticamente significativas observa-se que a primeira coluna, denominada fatores, apresenta a descrição de cada fator testado pelo VARBRUL; na segunda coluna apresentamos o número de casos da variante focalizada como base para a configuração da *regra variável*-

ausência de concordância - e o total de casos das duas variantes (soma dos casos com ausência e presença de concordância); a terceira coluna mostra a frequência relativa ou percentual da variante focalizada e a quarta coluna revela o peso relativo de cada fator na aplicação da *regra variável*.

Os pesos relativos acima de *0,50* favorecem a variante tomada como base e os pesos abaixo de *0,50* desfavorecem-na; pesos em torno de *0,50* são considerados neutros. Quanto maior o peso relativo, maior o índice de favorecimento da *ausência de concordância*.

(i) Núcleo do predicativo

Esta variável foi dividida em três fatores: *substantivo*, *adjetivo* e *formas participiais*. Estas últimas foram consideradas na análise devido à propriedade de o particípio passado também ocorrer em posição predicativa exercendo funções semelhantes às do adjetivo: participios passados e adjetivos têm em comum a característica de poderem ser flexionados em gênero e número em contextos predicativos, em concordância com o nome a que se referem, atribuindo-lhe qualificações. Neste caso, teríamos formações como: *as meninas estão animadas/o menino está animado*, onde o particípio passado *animado/animada* se flexiona conforme o nome a que se refere – *menino/menina*.

Já mencionamos anteriormente que, de acordo com a gramática normativa, os predicativos são representados por substantivo ou expressão substantivada, adjetivo, pronome e numeral. Entretanto, não pudemos incluir em nossa análise as categorias *numeral* e *pronome*. A primeira não foi registrada em nenhuma das ocorrências do *corpus* e a segunda foi desconsiderada no

¹⁴ Concordância de gênero.

processamento do programa por aparecer somente em estruturas com concordância de gênero, caracterizando um “*Knockout*”.

Fatores e exemplos da variável ‘Núcleo do predicativo’

a) Substantivo

- *concordância*

... minha *mulher* era *professora*... (MI49)

- *não concordância*

... a *madeira* arribá é o *rei* das madeiras... (MI54)

b) Adjetivo

- *concordância*

... a *madeira* é *escura*... (ME44)

- *não concordância*

... *criança* é *bobo*, né... (FI59)

c) Formas participiais

- *concordância*

... a *criança* tava *assustada*... (MI46)

- *não concordância*

... *mãe-de-ouro* é *encantado*... (FI35)

O *núcleo do predicativo* foi selecionado como primeira variável significativa na produção do fenômeno em estudo. Contamos com os seguintes números para a variável:

TABELA 2¹⁵: Efeito da Variável Núcleo do Predicativo

¹⁵ Fatores da variável, número de casos da variante focalizada - *ausência de concordância* - e o total de casos das duas variantes (soma dos casos com ausência e presença de concordância), frequência da ausência de concordância e peso relativo de cada fator na aplicação da regra variável.

Fatores	Apl/Total	Freqüência Relativa	Peso relativo
Substantivo	4/19	21%	0,27
Adjetivo	75/243	30%	0,42
Formas participiais	107/200	54%	0,61
Total	186/462	40%	

Os resultados indicam que a variável *núcleo do predicativo* exerce influência sobre a ausência de concordância de gênero. As formas participiais se constituiriam, no caso, na categoria gramatical mais propensa à ausência de concordância, alcançando os índices mais altos entre os três fatores testados (54%).

O substantivo, bem menos usado em construções predicativas do que o adjetivo e as formas participiais, por sua vez, seria a categoria menos favorecedora da *ausência de concordância*, ou seja, o falante utilizou muito pouco o substantivo na relação sujeito/predicativo, mas quando o fez, na maioria das vezes, realizou a concordância.

Quando analisou a concordância de gênero nos predicativos Souza (2000) constatou que a estrutura passiva favorecia a ausência da concordância de gênero. Na presente análise também chegamos à mesma constatação, pois as formas participiais, ao atingirem um peso de 0,61, apontam para um favorecimento da ausência de concordância.

(ii) Ordem dos elementos na construção

Ao definir esta variável, pretendíamos confirmar se nossas impressões iniciais, ao fazer o recorte para compor os dados da presente pesquisa, se confirmariam no sentido de que a *ordem não canônica* favoreceria a ausência de concordância de gênero.

Encontramos respaldo para a postulação desta hipótese em Scherre (1991), que testou a mesma variável na concordância de número entre o sujeito e o predicativo. A variável ficou composta por dois fatores: *ordem canônica* (sujeito+verbo+predicativo) e *ordem não canônica* (deslocamento do sujeito ou do predicativo).

Fatores e exemplos da variável 'Ordem dos Elementos na Construção'

a) Ordem canônica

- *concordância*

... a *lagartixa* é mais *preta*... (MI43)

- *não concordância*

... a *semente* é *gêmeo*... (MI44)

b) Ordem não canônica

- *concordância*

... tava *inflamada* a *garganta* né... (MI43)

- *não concordância*

... é *bonito* as *estrela*... (FE53)

Esta variável foi apontada pelo programa como a segunda mais significativa ao favorecimento da *ausência da concordância*. A tabela 3 ilustra a variável em números:

TABELA 3: Efeito da Variável Ordem dos Elementos na Construção

Fatores	Apl/Total	Frequência relativa	Peso relativo
Ordem canônica	152/411	37%	0,47
Ordem não canônica	34/51	67%	0,73
Total	186/462	40%	

A distribuição das percentagens e a leitura dos pesos relativos comprovam o que esperávamos: a *ordem não canônica* é bastante favorecedora da ausência de concordância, com peso relativo em torno de 0,73 e percentual de 67%.

Isso pode ser atribuído à interrupção do fluxo do processamento linear do pensamento do falante. Encontramos respaldo para esta constatação em Scherre (1991), que testou a mesma variável na concordância de número entre o sujeito e o predicativo. No entanto, seus pesos relativos não foram significativos: 0,48 para a *ordem canônica* favorecendo a aplicação da regra de concordância e 0,52 para a *ordem não canônica*.

Um exame na amostra de fala analisada nos levou a verificar que 24 das 34 construções em que prevalece a *ordem não canônica* são compostas pela estrutura *Verbo+Predicativo+Sujeito* (... *é bonito as estrela...*). Quando o falante quer se referir a algo no discurso mais à frente, pospõe claramente o sujeito em relação ao verbo e ao predicativo, prevalecendo, então esta estrutura.

Presumimos que isso ocorre devido ao fato de que o falante, por ter em vista as pressões de uso que atuam sobre seu desempenho, acabe por fazer a sua escolha de modo que seja preservada a competência comunicativa na transmissão da mensagem.

(iii) Característica formal do sujeito

A partir de observações iniciais sobre os dados levantados, acreditamos, inicialmente, que o *sujeito não-preenchido* tende a favorecer a ausência da concordância de gênero.

Para a composição desta variável nos baseamos em Souza (2000)¹⁶ e definimos dois grupos de fatores: *sujeito preenchido* e *sujeito não-preenchido*. Este último foi aquele que encontramos no contexto, porém sem estar explícito.

Fatores e exemplos da variável 'Característica Formal do Sujeito'

a) Sujeito preenchido

- *concordância*

... a *madeira* é mais *curta* assim, é mais *grossa*, né... (ME52)

- *não concordância*

... essa *doença* é *conhecido*... (ME44)

b) Sujeito não-preenchido

- *concordância*

... pra cá nós plantava em terra vermelha e pra cá *era branca*... (FI38)

- *não concordância*

... algum outro faz a *pamonha* [...] aí já não é doce, é *salgado*... (MI46)

A variável *característica formal do sujeito* foi selecionada como a terceira variável significativa na produção do fenômeno em estudo e apontou o *sujeito não-preenchido* como o fator com maior índice de ausência de concordância de gênero (52%). Veja-se a tabela 4:

TABELA 4: Efeito da Variável Característica Formal do Sujeito

Fatores	Apl/Total	Frequência relativa	Peso relativo
Sujeito preenchido	118/332	36%	0,46
Sujeito não-preenchido	68/130	52%	0,61
Total	186/462	40%	

¹⁶ Souza (2000) constatou o *sujeito não preenchido* favorece a aplicação da concordância de Gênero na fala de seus informantes

Os dados comprovam nossa expectativa inicial ao selecionarmos este grupo de fatores: o *sujeito não-preenchido* favoreceria a ausência de concordância.

Isso nos leva a inferir que quando os informantes utilizam o recurso do não preenchimento do sujeito se referem a ele com menor clareza de pensamento na seqüência da estrutura predicativa, o que talvez possa estar condicionando a ausência de concordância, isto é, por já ter se referido anteriormente ao sujeito o falante pode não se preocupar em realizar a concordância ou não oferecer muita atenção ao fato.

(iv) Tipo de referência do sujeito

Durante a fase de composição da amostra de fala a ser analisada nesta dissertação verificamos um aspecto que se destacou: a frequência com que núcleos de predicativos representados por nomes masculinos eram ligados a sujeitos representados por nomes femininos. Entretanto, não encontramos nomes masculinos na função de sujeito ligados a nomes femininos na função de predicativo.

A partir daí, e fundamentados em autores como Martin (1975) e Câmara Júnior (1995) que apontam uma tendência de o masculino ser, de modo geral, a forma usada para os casos em que se faz uma referência genérica, formulamos a hipótese de que esta tendência estaria influenciando o falante a fazer uso de estruturas sem concordância de gênero entre o sujeito e o predicativo ao se referir de uma forma genérica ao nome que representa o sujeito.

Também encontramos na pesquisa de Dettoni (2003) alguns resultados que nos motivaram a levantar a hipótese em questão. Dettoni testou a mesma variável em sua pesquisa sobre a variação na concordância de gênero na anáfora pronominal e concluiu ser a *referência genérica* o fator que mais favorece a ausência de concordância (cf. seção 3.5).

De acordo com Mira Mateus et al (1994), quando os *nominais* (substantivos na tradição gramatical), em seu uso, referem-se a um único objeto identificado, individualizando-o em relação ao conjunto de todos os objetos que se definem pela propriedade de 'ser x' (p.ex. 'ser casa', 'ser livro', 'ser criança' etc), apresentam descrições definidas e seu uso característico é denominado *uso referencial* - ou *referência específica*. Os possessivos, os demonstrativos, os pronomes pessoais, os nomes próprios e os artigos definidos (conforme o contexto), encaixam-se nesta denominação.

Contudo, os nomes comuns nem sempre são usados com uma *referência específica*. É o caso de nomes que, mesmo acompanhados por expressões de individuação como os artigos definidos, não se referem a um único objeto identificado, mas ao conjunto de um determinado objeto encarado na sua totalidade.

Em casos assim, o artigo que acompanha o nome funciona como uma expressão de *quantificação universal* e caracteriza o *uso genérico* dos nominais. A seguir, encontram-se exemplos que ilustram os dois tipos de referência:

(a) *Uso referencial*

O *teu* carro está mal estacionado.

(b) *Uso genérico*

O *carro* é um meio de transporte.

Para testar a hipótese levantada, estabelecemos a variável *tipo de referência do sujeito* com dois fatores: *referência genérica* e *referência específica*. Com esta variável objetivamos avaliar se o tipo de referência que estava sendo feita pelo falante para se referir ao nome que exerce a função de sujeito na construção favoreceria, de algum modo, a ausência da concordância de gênero.

Fatores e exemplos da variável 'Tipo de Referência do Sujeito':

a) Referência genérica:

- *concordância*

... o lagarto é um lagarto e é grande né, é maior, e a *lagartixa* já é *pequeninha*... (MI54)

- *não concordância*

... apareceu a serpente pra Eva, *mulher* é *fraco* mesmo... (FA59)

b) Referência específica:

- *concordância*

... *minha mãe* ainda era *viva*... (F159)

- *não concordância*

... *as criança* era *pequeno*, eu tinha torrado um café... (FE50)

Os resultados da análise são apresentados na tabela 5:

TABELA 5: Efeito da Variável Tipo de Referência do Sujeito

Fatores	Apl/Total	Frequência relativa	Peso relativo
Referência genérica	117/262	45%	0,57
Referência específica	69/200	34%	
Total	186/462	40%	

A tabela indica um favorecimento da ausência de concordância pela referência genérica do sujeito. Os pesos relativos (0,57) para o fator *referência genérica* e para o fator *referência específica* (0,41), apresentam uma diferença

significativa. Mesmo não tendo sido selecionada como a mais influente para a produção do fenômeno em estudo nesta pesquisa, o *tipo de referência* correspondeu às nossas expectativas iniciais ao postularmos a hipótese de que a *referência genérica* atuaria no sentido de favorecer a ausência da concordância de gênero nas relações entre sujeito e predicativo.

5.1.2 Variáveis lingüísticas não significativas estatisticamente

(i) Verbo utilizado

A fim de verificar se o tipo de verbo de ligação utilizado pelo falante afetaria a sua escolha entre as variantes *ausência/presença de concordância*, ao mesmo tempo em que teria influência sobre os outros grupos de fatores da pesquisa, subdividimos a variável em três fatores: *verbo ser*, *verbo estar* e *verbo ficar*.

Mesmo a variável não tendo sido selecionada pelo VARBRUL não podemos deixar de acrescentar que na análise multidimensional do programa de regras variáveis o grupo de fatores *verbo utilizado* foi o último eliminado pelo “*step down*”¹⁷.

¹⁷ O VARBRUL seleciona as variáveis significativas por meio de comparações progressivas entre as variáveis independentes e da projeção, em pesos relativos, para os seus respectivos fatores utilizando diversos níveis de análise em um método denominado “*step up*”. Este método realiza a comparação de cada uma das variáveis com a média global corrigida ou *input* e trabalha sucessivamente até que as variáveis independentes adicionadas não sejam mais selecionadas. Já o “*step down*” é o método que atua de forma inversa ao atribuir pesos relativos a todos os fatores de todas as variáveis independentes num só nível de análise. Em seguida, com retiradas de cada uma delas, verifica-se se todas as variáveis selecionadas não são eliminadas, se as selecionadas também são eliminadas e se existem variáveis que não são nem selecionadas ou eliminadas. Para maiores esclarecimentos, consultar Sankoff (1988) e

Fatores e exemplos da variável 'Verbo Utilizado'

a) Verbo ser

- *concordância*

... a flor é amarela... (FI41)

- *não concordância*

... é preto, né, semente pra comer é preto... é meio misturado... (MI39)

b) Verbo estar

- *concordância*

... a palha dele tá amarela... (FI49)

- *não concordância*

... temos a nova que ela tá bem pequenininho... (MI54)

c) Verbo ficar

- *concordância*

... sei que minha mãe ficou revoltada... (FI42)

- *não concordância*

... a espiga ficou ralo... (MI35)

Os resultados alcançados pela variável se encontram na tabela a seguir:

TABELA 6: Efeito da Variável Verbo Utilizado

Fatores	Apl/Total	Frequência relativa	Peso relativo
Verbo ser	120/314	38%	0,51
Verbo estar	49/111	44%	0,42
Verbo ficar	17/37	46%	0,61

Scherre e Naro (2003).

Total	186/462	40%
--------------	----------------	------------

De acordo com os números, destacados na tabela 6, predominam as estruturas cujos sujeitos e predicativos são ligados pelo verbo *ser*, o que é comum quando se trata de estruturas predicativas. Mas isso não significa que este verbo favoreça a ausência de concordância. Ao contrário, o verbo *ser* atingiu o maior número de realizações da variante padrão e um peso relativo muito próximo à neutralidade, 0,51.

O verbo *estar*, mesmo com um percentual de ausência de concordância maior (44%) do que o verbo *ser*, atingiu um peso relativo inferior, 0,42, também próximo à neutralidade.

O verbo *ficar*, no entanto, alcançou um peso relativo maior, 0,61, o que aponta para um favorecimento da ausência de concordância. Porém, o programa não selecionou a variável como estatisticamente significativa, o que pode ter decorrido de uma distribuição desequilibrada dos dados, já que o verbo *ficar* foi bem menos recorrente nas estruturas analisadas (8% dos casos) do que os verbos *ser* (68%) e *estar* (24%).

(ii) Elementos intervenientes entre o verbo e o predicativo

O estabelecimento deste grupo de fatores, num primeiro momento, foi ao encontro da proposta de Scherre (1991, p. 66-67) que, em seu trabalho sobre a concordância de número nos predicativos e participios passivos, dividiu o *material interveniente* entre o verbo e o predicativo em três categorias para a variável. Entretanto, ao codificar os dados, decidimo-nos por definir apenas duas categorias para esta variável devido ao fato de que a outra categoria definida por Scherre –

presença de intensificadores – apresentou-se em número muito reduzido (quatro) nas estruturas sem concordância de gênero. A partir disso, resolvemos incluí-la no fator *presença de elemento interveniente*. Dentro deste fator agrupamos, então, os artigos, advérbios e pronomes indefinidos.

Fatores e exemplos da variável ‘Elementos Intervenientes entre o Verbo e o Predicativo’

a) Presença de elemento interveniente

- *concordância*

... a serra *é muito alta* e a lomba *é mais baixa*... (MI51)

- *não concordância*

... tem a banana conga [...] ... *é tudo nativo*... (MI54)

b) Ausência de elemento interveniente

- *concordância*

... tá bichada né, aquela *goiaba* no pé *tá bichada* (FI42)

- *não concordância*

... mulher *é fraco*... (MI59)

Ao levantar a hipótese sobre a influência desta variável levamos em consideração o fato de que, por se postar imediatamente logo após o verbo, um elemento interveniente afastaria o sujeito do predicativo, podendo exercer algum peso sobre a ausência da concordância de gênero por parte do falante. Porém este grupo não correspondeu às nossas expectativas.

De acordo com a tabela 7, os números atingidos com a quantificação dos dados foram:

TABELA 7: Efeito da Variável Elementos Intervenientes entre Verbo e Predicativo

Fatores	Apl/Total	Frequência relativa	Peso relativo
Presença de elemento interveniente	72/166	43%	0,51
Ausência de elemento interveniente	114/296	39%	0,49
Total	186/462	40%	

Os dados quantificados sugerem que os *elementos intervenientes* entre o verbo e o predicativo não interferem nas relações de concordância entre os elementos analisados, a exemplo do que verificou Scherre (1991) para a concordância de número. Além de não ser selecionada pelo “*step up*”, a variável *elementos intervenientes entre o verbo e o predicativo* foi a primeira eliminada pelo “*step down*”.

5.1.3 Variáveis sociais

(i) Sexo do informante

A variável *sexo do informante* foi selecionada como a única variável social estatisticamente significativa. Das cinco variáveis selecionadas, esta foi a terceira na ordem de seqüência das mais significativas. Com relação ao fator sexo, é tradicionalmente aceito que nos casos de variação estável os falantes do sexo feminino tendem a se aproximar mais da norma padrão do que os falantes do sexo masculino.

É de Fischer (apud PAIVA, 2003, p. 34) a primeira referência à correlação entre variação lingüística e o fator sexo em um trabalho intitulado *Influências sociais na escolha de variantes lingüísticas*. O autor verificou que o sexo do falante pode exercer uma grande influência em processos variáveis de diferentes níveis nos quais é maior a probabilidade de as mulheres demonstrarem preferência pelas variantes lingüísticas mais prestigiadas socialmente, especialmente em fenômenos que indicam variação estável.

Posteriormente, a constatação de Fischer foi corroborada por outros estudos sócio-variacionistas, entre os quais citamos Vazzata-Dias (2000) e Scherre (1991), em pesquisa sobre a variação na concordância de número nos predicativos e participios na fala de informantes nascidos no sul do Brasil e de informantes do Programa de Estudos sobre o Uso da Língua (PEUL)¹⁸, respectivamente, e Karim (2004), em pesquisa sobre a variação na concordância de gênero (cf. seção 3.5).

Contudo, mesmo a literatura apontando para a predominância de casos em que falantes do sexo feminino se aproximam mais da norma padrão, existem estudos que mostram o contrário, como, por exemplo, Seara (2000) em pesquisa sobre a variação do sujeito *nós* e *a gente* na fala florianopolitana, e Lucchesi (2000) e Souza (2000) que, em suas pesquisas sobre a variação na concordância de gênero (cf. seção 3.5), verificaram que os homens tendem a usar mais a variante formal do que as mulheres.

Em nosso estudo, os resultados da variável *sexo do informante* não apresentaram diferenças muito significativas entre homens e mulheres. Ainda assim, os percentuais e os pesos relativos, embora próximos do ponto neutro (0,57 e 0,41),

¹⁸ Grupo de pesquisa sediado no Departamento de Lingüística e Filologia da Faculdade de Letras da Universidade Federal do Rio de Janeiro. Para maiores informações consultar Naro et al (1986).

apontam, relativamente, para um maior favorecimento da ausência de concordância na fala das mulheres. A frequência da variável lingüística *ausência de concordância de gênero* entre o sujeito e o predicativo na fala das mulheres atingiu um percentual de 47%, enquanto que na fala dos homens atingiu um percentual de 31%.

Os dados fornecidos pelo programa demonstram que os homens realizam mais a concordância de gênero no falar do Paraná Tradicional rural. Os resultados para a variável sexo são apresentados na tabela abaixo:

TABELA 8: Efeito da Variável Sexo

Fatores	Apl/Total	Frequência relativa	Peso relativo
Feminino	125/265	47%	0,57
Masculino	61/197	31%	0,41
Total	186/462	40%	

Como se pode observar, as mulheres utilizam mais estruturas sem concordância. Na verdade, os homens estariam mais afetados pela pressão normativa que as mulheres. A análise da influência dessa variável vai ao encontro aos resultados obtidos por Lucchesi (2000) que, em sua pesquisa sobre a concordância de gênero no sintagma nominal em Helvécia, verificou que os homens superam as mulheres no uso da forma padrão – peso relativo de 0,56 para a realização da regra entre os homens contra 0,45 para as mulheres.

Diante disso, talvez possamos inferir que, no dialeto do Paraná Tradicional rural, os homens se aproximam mais das formas de maior prestígio social pelo fato de terem oportunidade de interagir com outros grupos sociais, pois no contexto cultural das comunidades rurais eles tendem a ter mais contato com o mundo exterior. Silva e Paiva (1996, p. 368) destacam que eles possuem “uma vida social

mais intensa e um maior contato com seus iguais”, o que os levaria ao “compromisso de partilhar idéias e atitudes do grupo que participam”.

(ii) Grau de escolaridade do informante

A hipótese para esta variável é que quanto menor fosse o grau de escolaridade do informante, mais propensão ele teria de fazer uso de construções caracterizadas pela ausência de concordância de gênero entre o sujeito e o predicativo, como também o inverso seria verdadeiro, isto é, quanto maior fosse seu *grau de escolaridade* mais chance ele teria de realizar a concordância.

Esta hipótese é respaldada por pesquisas como a de Lucchesi (2000), Souza (2000), Dettoni (2003) e Karim (2004). Todas essas pesquisas (cf. seção 3.5) atestam que o acesso dos falantes à escolaridade tem contribuído para a assimilação das formas de prestígio da língua, com a conseqüente neutralização das formas desprestigiadas.

Os resultados que obtivemos para a variável *grau de escolaridade* corresponderam à expectativa, com os analfabetos superando os escolarizados em relação ao uso da variante não-padrão. Este grupo de fatores, no entanto, não foi selecionado pelo programa de regras variáveis.

Vejamos, então, na tabela 9, os resultados referentes à variável:

TABELA 9: Efeito da Variável Grau de Escolaridade

Fatores	Apl/Total	Frequência relativa	Peso relativo
Analfabeto	126/303	42%	0,52
Escolarizado	60/159	38%	0,46
Total	186/462	40%	

A partir da tabela apresentada, percebe-se, no cômputo geral, que o *grau de escolaridade* exerce pouca influência sobre a escolha do falante em relação à aplicação da *regra variável*: tanto escolarizados quanto analfabetos apresentam percentuais muito próximos em relação a esta variável. Os dados percentuais confirmam que as diferenças entre analfabetos e escolarizados quanto à variação de concordância são muito pequenas, na ordem de 4% apenas e os pesos relativos se encontram muito próximos do ponto neutro (0,52 para analfabetos e 0,46 para escolarizados).

Os índices apresentados para a variante *ausência de concordância* parecem refletir uma tendência à uniformidade dos padrões de comportamento lingüístico das duas faixas. É preciso assinalar, porém, que o resultado obtido pelo fator condicionador *grau de escolaridade* aponta em direção à hipótese inicial, ou seja, o que se verifica é que os falantes escolarizados aplicam mais a concordância que os falantes analfabetos.

A princípio, o problema poderia ser relacionado a uma distribuição desequilibrada em relação à escolaridade dos informantes na amostra analisada: na *faixa I* (de 30 a 46 anos) somente cinco dos vinte e um informantes eram escolarizados; na *faixa II* (47 a 65 anos), nove informantes o eram. Mas, a partir da observação de que os falantes escolarizados revelam percentuais elevados quanto ao uso da forma não-padrão, nos afastamos desse raciocínio. No geral, quem freqüentou a escola não deveria apresentar um índice tão alto - 38% - de *ausência de concordância*.

Talvez isso se deva ao fato de termos trabalhado apenas com informantes que, no máximo, freqüentaram até a quarta série do ensino

fundamental. Desses, somente quatro concluíram os estudos, dez não chegaram à quarta série. Por trabalharmos com dados obtidos junto a informantes detentores de níveis tão baixos de escolaridade, não se pode ignorar que aqueles que tiveram esse contato mínimo com a escola, muitas vezes, desenvolveram somente a habilidade de escrever o próprio nome, o que tem um efeito mais simbólico no plano da cidadania do que funcional, pois o indivíduo que sabe assinar seu nome tem o direito de receber seus documentos e utilizar os serviços públicos sem se expor a uma situação de inferioridade proveniente da demonstração do seu analfabetismo num mundo regido pelo registro escrito.

(iii) Faixa etária do informante

A idade do falante tem sido abordada pela sociolinguística como um dos fatores que caracterizam o fenômeno lingüístico estável como envolvendo ou não o processo de mudança lingüística.

No entanto, em nossa pesquisa, o grupo de fatores *faixa etária* não se mostrou estatisticamente significativo. Os resultados são apresentados na tabela a seguir:

TABELA 10: Efeito da Variável Faixa Etária

Fatores	Apl/Total	Freqüência relativa	Peso relativo
Faixa etária I - 30 a 46 anos	64/179	36%	0,47
Faixa etária II - 47 a 65 anos	122/283	43%	0,52
Total	186/462	40%	

Nota-se, pela tabela, que, assim como o *grau de escolaridade*, a *faixa etária* não mostra uma grande diferença no uso das variantes entre informantes

da *faixa etária I*, os mais jovens, e da *faixa etária II*, os mais velhos, sugerindo apenas um pequeno favorecimento da ausência de concordância para a faixa etária dos mais velhos: os pesos relativos, ambos próximos à neutralidade, indicam não ter a variável *faixa etária* uma influência significativa sobre a variação da concordância.

Tais índices poderiam ser atribuídos à diferença entre o número de informantes das duas faixas etárias, resultante da forma como foram divididas, conforme esclarecido na seção 4.2. Entretanto, após a quantificação dos dados pelo VARBRUL, percebemos que mesmo tendo dois informantes a mais, a *faixa etária I*, dos mais jovens, apresentou os índices mais baixos de ausência de concordância.

Acreditamos ser este dado relevante, pois, mesmo com a exclusão da *faixa etária do informante* pelo programa e os pesos relativos apontando para uma aparente indiferença quanto à sua influência sobre a ausência ou presença de concordância, a expectativa de que a faixa etária I tenderia a apresentar números mais baixos de produtividade no fenômeno se cumpriu de acordo com o esperado, indicando que o uso da variante não-padrão é mais acentuado na fala dos mais velhos.

Como vimos, o emprego da variante não-padrão – *ausência de concordância* - pelos informantes pode ser explicado tanto por fatores lingüísticos quanto por fatores sociais, incluindo também aqueles que não foram selecionados como estatisticamente relevantes pelos programa de regras variáveis.

De modo geral, no que se refere às variáveis lingüísticas e às sociais que podem condicionar a concordância de gênero, podemos estabelecer uma comparação entre esta pesquisa e as pesquisas variacionistas abordadas na seção 3.5, capítulo 3, e citadas também neste capítulo.

Dentre estas, a que melhor permite estabelecer comparações é a pesquisa de Souza (2000) na comunidade de Caçandoca, pois, assim como ele, também nos concentramos em analisar as relações entre o sujeito e o predicativo e testamos muitas variáveis em comum, devido à própria natureza morfossintática do fenômeno em análise. Os resultados também se assemelharam quanto ao desempenho de alguns fatores lingüísticos no sentido de favorecer a ausência de concordância, tais como o *sujeito não-preenchido* e o *tipo de núcleo do predicativo*.

O trabalho de Dettoni (2003) nos orientou quanto à expectativa de testar a variável *tipo de referência do sujeito* que, conforme análise demonstrada acima, apresentou-se relevante para a aplicação da *regra variável*. Assim como na variação da concordância de gênero na anáfora pronominal no dialeto da baixada cuiabana, estudada por Dettoni, o fator *referência genérica* tem influência sobre a ausência de concordância de gênero na relação sujeito/predicativo no dialeto do Paraná Tradicional rural.

As pesquisas de Lucchesi (2000) sobre a concordância de gênero no sintagma nominal em Helvécia, na Bahia e de Karim (2004) sobre a concordância de gênero na comunidade de Cáceres, Mato Grosso, juntamente com Souza (2000) e Dettoni (2003) nos auxiliaram na análise das variáveis sociais. Assim como esses pesquisadores verificamos em nosso estudo que o sexo, a idade e a escola influenciam na realização da concordância.

CONSIDERAÇÕES FINAIS

Nesta pesquisa, os instrumentos teóricos e metodológicos da Sociolingüística foram fundamentais para que pudéssemos mostrar a freqüência da variação e identificar os condicionadores mais decisivos da variação da concordância de gênero que envolve as relações entre o sujeito e o predicativo.

Iniciamos o trabalho com uma breve descrição da variedade do português falada no Paraná Tradicional rural, enfatizando alguns de seus aspectos mais incidentes no que se refere à variação do gênero gramatical. Para melhor caracterizar o fenômeno analisado foram apresentados alguns estudos que descreveram e analisaram a variação na concordância de gênero em diversas variedades do português faladas no país, como nos Estados da Bahia, Mato Grosso e São Paulo.

Dentre estes estudos, mereceram destaque especial aqueles cujos autores utilizaram os pressupostos teórico-metodológicos da Sociolingüística Variacionista, os quais também utilizamos no decorrer desta pesquisa a fim de associar os condicionamentos lingüísticos aos sociais para o estudo da freqüência da variação na concordância nominal de gênero. Mesmo sem terem sido colhidas especialmente para a investigação do fenômeno em estudo, as entrevistas utilizadas no trabalho se revelaram fontes muito produtivas para a análise aqui desenvolvida.

Dentre as seis variáveis lingüísticas propostas como condicionadoras da produção do fenômeno pesquisado, quatro foram ao encontro de nossa

expectativa inicial ao selecioná-las: *núcleo do predicativo, ordem dos elementos na construção, tipo de referência do sujeito, característica formal do sujeito*.

Os dados fornecidos pelo programa de regras variáveis indicaram o fator *formas participiais* do grupo de fatores *núcleo do predicativo* como o maior favorecedor da ausência de concordância. Com isso, não se confirmou nossa hipótese principal de que a *referência genérica* do nome que representa o sujeito da construção constituiria o ambiente mais favorável à ausência de concordância de gênero entre este e o seu predicativo.

Ainda assim, pode-se afirmar, com base no percentual e no peso relativo fornecidos pela análise quantitativa para a variante *ausência de concordância*, que a referência genérica pode influenciar a ausência de concordância entre as categorias analisadas.

Os resultados indicaram que a *ordem não canônica* favorece a ausência de concordância, correspondendo à hipótese formulada de que o fenômeno também pode ser atribuído à interrupção do fluxo do processamento linear do pensamento dos falantes, o que não ocorreu para a variação da concordância de número pesquisada por Scherre (1991).

Quanto ao fator condicionador *sexo*, mesmo selecionado pelo VARBRUL, apresentou números pouco expressivos em relação às diferenças na fala de homens e mulheres. Ainda assim, os números em percentuais e em pesos relativos indicaram ser menor a ausência de concordância na fala dos homens. Na verdade, então, os homens estariam mais orientados em direção às formas de prestígio do que as mulheres. Tal fato difere da previsão geralmente aceita de que as mulheres lideram as mudanças em direção às formas padrão.

As variáveis *grau de escolaridade* e *faixa etária* não foram apontadas como estatisticamente relevantes na aplicação da *regra variável*. Com relação ao *grau de escolaridade* isso tal talvez possa ser atribuído ao fato termos utilizado apenas informantes que cursaram até a quarta série do ensino fundamental, o que, atualmente, significa muito pouco em termos de aprendizagem.

Ainda assim, não desprezamos os resultados que, em termos percentuais, mesmo com poucas diferenças, nos levam a concluir que os falantes mais velhos e analfabetos utilizam com maior freqüência a variante não-padrão. Acreditamos, com isso, que a maioria das comunidades rurais do país tem sido atingidas, nas últimas décadas, por processos de mudança lingüística.

A mudança em questão pode estar sendo condicionada por vários fatores, dentre os quais são citados: o êxodo rural, a expansão dos meios de comunicação de massa, a melhora das estradas rurais e a abertura de rodovias, o maior acesso à tecnologia, ao sistema educacional, aos meios de transportes e aos serviços públicos, em geral. Como consequência, o estilo de vida dos falantes rurais não foge às pressões da sociedade capitalista, em que a variedade padrão é a mais valorizada.

Podemos afirmar que o estudo lingüístico das comunidades rurais, não só no Paraná, como no resto do país, é de grande relevância não apenas no âmbito da ciência sociolingüística, mas também por representar uma importante contribuição para o conhecimento e caracterização da identidade cultural dos brasileiros de origem rural, que vêm sendo estigmatizados ao longo dos anos e assim contribuir para a descrição da Língua Portuguesa falada no Brasil.

Com esses resultados, concluímos, por enquanto, nossa reflexão sobre a variação na concordância de gênero no falar rural paranaense. Porém, estamos conscientes de que, quanto mais estudamos um fenômeno, mais aspectos surgem para serem investigados. E, ao final deste trabalho, um questionamento se faz presente: será que se, hoje, quase 20 anos após a coleta dos dados que analisamos, colhêssemos novos dados nos mesmos pontos lingüísticos, os resultados seriam os mesmos? Pensamos que não, principalmente se levarmos em conta que a língua está em constante mutação e considerarmos que, nos últimos anos, o acesso à escolarização e às facilidades da vida moderna em muito têm atuado no sentido de levar os falantes a fazerem um uso cada vez maior das formas de prestígio.

Pretendemos, com as informações presentes nesta pesquisa, contribuir para uma melhor compreensão do fenômeno em estudo em dialetos rurais do Brasil, como também para que se compreenda melhor a natureza do gênero gramatical no português brasileiro.

REFERÊNCIAS

AGUILERA, V. de A. *Atlas Lingüístico do Paraná*. Curitiba: Imprensa Oficial do Paraná, 1994.

_____. *Atlas Lingüístico do Paraná: banco de dados sonoro*. 1990a, inédito.

_____. *Atlas Lingüístico do Paraná: entrevistas*. 1990b. Versão digitada, 13 disquetes.

_____. *Atlas Lingüístico do Paraná: entrevistas*. 1990c. Versão impressa.

ALKMIM, T. M. Sociolingüística. In: MUSSALIN, F.; BENTES, A. C. *Introdução à Lingüística: domínios e fronteiras*. 5. ed. São Paulo: Cortez, 2004.

ALMEIDA, J. E. de. *Discurso rural: uma perspectiva ideológica*. 1996. Dissertação (Mestrado em Letras) - Universidade Estadual de Londrina, Londrina.

AMARAL, A. *O Dialeto Caipira*. 3. ed. São Paulo: HUCITEC, 1976.

BAXTER, A. N.; LUCCHESI, D. A Relevância dos processos de pidginização e crioulização na formação da língua portuguesa no Brasil. *Revista de Estudos Lingüísticos e Literários, Salvador*, n. 19, p.65-83, mar., 1997.

BECHARA, E. *Moderna gramática portuguesa*. 30. ed. São Paulo: Editora Nacional, 1986.

BRESCANCINI, C. R. A análise de regra variável e o programa VARBRUL 2S. In: BISOL, L.; BRESCANCINI, C. R. (Orgs.). *Fonologia e Variação: recortes do português brasileiro*. Porto Alegre: EDIPUCRS, 2002. p. 13-75.

CALLOU, D. M. I. Um estudo em tempo real em dialeto rural brasileiro: questões morfossintáticas. In: GROßE, S.; ZIMMERMANN, K. (Eds.). *“Substandard” e mudança no português do Brasil*. Frankfurt am main: TFM: 1998, p. 255-272.

CAMACHO, R. G. Sociolingüística. In: MUSSALIN, F.; BENTES, A. C. *Introdução à Lingüística: domínios e fronteiras*. 5. ed. São Paulo: Cortez, 2004.

CÂMARA JÚNIOR, J. M. Considerações sobre o gênero em português. *Estudos Lingüísticos – Revista Brasileira de Lingüística Teórica e Aplicada*, Rio de Janeiro, v. 1, n. 2, p. 1-9, dez. 1966.

_____. *Estrutura da Língua Portuguesa*. 23. ed. Petrópolis: Vozes, 1995.

CARDOSO, J. A.; WESTPHALEN, C. M. *Atlas Histórico do Paraná*. Curitiba: Livraria do Chain, 1986.

CARENO, M. F. do. *A linguagem rural do vale do Ribeira: a voz e a vez das comunidades negras*. 1991. Tese (Doutorado em Lingüística) - Universidade Estadual Paulista, Assis.

CUNHA, C.; CINTRA, L. *Nova Gramática do Português Contemporâneo*. 3. ed. Rio de Janeiro: Nova Fronteira, 2001.

DETTONI, R. do V. *A concordância de gênero na anáfora pronominal: variação e mudança lingüística no dialeto da Baixada Cuiabana – Mato Grosso*. 2003. 256f. Tese (Doutorado em Estudos Lingüísticos) - Universidade Federal de Minas Gerais, Belo Horizonte.

FASOLD, R. W. *The Sociolinguistics of Society*. Basil Blackwell, 1984.

GUSMÃO, A. B. *As realizações da vibrante na variedade lingüística rural do Paraná: uma abordagem geovariacionista*. 2004. 156 f. Dissertação (Mestrado em Estudos da Linguagem) - Universidade Estadual de Londrina, Londrina.

KARIM, J. M. *A variação na concordância de gênero no falar da comunidade de Cáceres-MT*. 2004. 102f. Dissertação (Mestrado em Lingüística e Língua Portuguesa) - Universidade Estadual Paulista, Araraquara.

LABOV, W. *Sociolinguistic Patterns*. Oxford: Basil Blackwell, 1972.

_____. *Sociolingüística*. Paris: Les Édition de Minuit, 1976.

_____. What can be learned about change in progress from synchronic descriptions? In: SANKOFF, David & CEDERGREN, Henrietta (Eds.). *Variation Omnibus*. Canadá: Lingüístico Research, Inc., 1981, p. 177-199.

LUCCHESI, D. *A variação na concordância de gênero em uma comunidade de fala afro-brasileira* – Novos elementos sobre a formação do português popular do Brasil. 2000. 364f. Tese (Doutorado em Lingüística) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

LYONS, J. *Introdução à Lingüística Teórica*. Tradução de Rosa Virgínia Mattos e Silva e Hélio Pimentel. São Paulo: Companhia Editora Nacional/EDUSP, 1979.

MARROQUIM, M. *A língua do Nordeste (Alagoas e Pernambuco)*. São Paulo: Ed. Nacional, 1934.

MARTINS, R. *História do Paraná*. 3. ed. Curitiba: Guaíra, 1995.

MARTIN, J. G. Gênero? *Revista Brasileira de Lingüística*, Petrópolis, n. 2, p.3-8, 1975.

MIRA MATEUS et al. *Gramática da Língua Portuguesa*. 4. ed. Lisboa: Editorial Caminho, 1994.

MOLLICA, M. C. Fundamentação teórica: conceituação e delimitação. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.) *Introdução à Sociolingüística: o tratamento da variação*. São Paulo: Contexto, 2003.

MONTEIRO, J. L. *Para compreender Labov*. Petrópolis: Vozes, 2000.

NARO, A. J. Modelos quantitativos e tratamento estatístico. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.) *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003.

_____. et al. *Relatório final de pesquisa: Projeto Subsídios do Projeto Censo à Educação*. Rio de Janeiro: UFRJ, 1986. Mimeo.

NASCENTES, A. *O linguajar carioca*. Rio de Janeiro: Simões, 1953.

PAIVA, M. da C. A variável gênero/sexo. In: MOLLICA, Maria Cecília; BRAGA, Maria Luíza (Orgs.) *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003.

PETTER, M. M. T. *Relatório Científico apresentado à FAPESP Projeto Vestígios de Dialeto Crioulos de Base Lexical Portuguesa em Comunidades Afro-Brasileiras Isoladas*. São Paulo: 1995 mimeo.

_____. A linguagem do Cafundó: crioulo ou anticioulo? In: ZIMMERMANN, K. (ed.) *Lenguas criollas de base lexical española y portuguesa*. Frankfurt am Main: Vervuet, 1999.

PINTZUK, S. *VARBRUL programs*. Heslington: University of York, 1988.

RAND, D.; SANKOFF, D. *Goldvarb: a variable rule application for Macintosh*. Montreal: Université de Montreal, 1990.

ROBINSON, J. S.; LAWRENCE, H. R.; TAGLIAMONTE, S. A. *GOLDVARB 2001: a multivariate analysis application for Windows*. Heslington: University of York, 2001.

ROCHA LIMA, C. H. da. *Gramática normativa da língua portuguesa*. 18. ed. Rio de Janeiro: José Olympio, 1976.

ROCHA, L. C. de A. *Estruturas morfológicas do português*. Belo Horizonte: UFMG, 1998.

RODRIGUES, A.N. *O dialeto caipira na região de Piracicaba*. São Paulo: Ática, 1974.

RODRIGUES, R. E. *Para um glossário da fala popular rural paranaense*. 2000. Dissertação (Mestrado em Estudos da Linguagem) - Universidade Estadual de Londrina, Londrina.

SANKOFF, D. Variable rules. In: AMMON, U.; DITTMAR, N; MATTHEIER, K. J. (Eds.) *Sociolinguistics – An international handbook of the science of language and society*. Berlin/New York: Walter de Gruyter, 1988. p. 984-998.

SCHERRE, M. M. P. A concordância de número nos predicativos e nos participios passivos. *Organon*, Porto Alegre, v. 5, n. 18, p. 52-70, 1991.

_____. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa*, Lisboa, n. 12, p. 37-49, 1994.

_____. Sobre a influência de variáveis sociais na concordância nominal. In: OLIVEIRA E SILVA, G. M.; SCHERRE, M. M. P. (Orgs.). *Padrões sociolingüísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro, 1996.

_____. Pressupostos teóricos e suporte quantitativo. In: OLIVEIRA E SILVA, G. M.; SCHERRE, M. M. P. (Orgs.). *Padrões Sociolingüísticos*. 2. ed. Rio de Janeiro: Tempo Brasileiro, 1998.

_____; NARO, A. J. Análise quantitativa e tópicos de interpretação do Varbrul. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). *Introdução à Sociolingüística: o tratamento da variação*. São Paulo: Contexto, 2003.

SEARA, I. C. A variação do sujeito *nós* e *a gente* na fala florianopolitana. *Organon*, Porto Alegre, v. 14, n. 28/29, p. 179-194, 2000.

SILVA, G. M.; PAIVA, M. A. Visão de conjunto das variáveis sociais. In: OLIVEIRA E SILVA, G.; SCHERRE, M. M. (Orgs.). *Padrões sociolingüísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro 1996.

SOUZA, A. C. S. de. A Variação da Concordância de Gênero entre o Sujeito e o Predicativo na Linguagem do Cafundó. ESTUDOS LINGÜÍSTICOS, XLVI, **Anais** ...Bauru, v. 28, p.208-214, 1999.

_____. A Concordância de gênero entre o sujeito e o predicativo na fala da comunidade Quilombola da Caçandoca. 2000. 100f. Dissertação (Mestrado em Lingüística e Semiótica Geral) - Universidade de São Paulo, São Paulo.

TARALLO, F. *A pesquisa Sociolingüística*. 5. ed. São Paulo: Ática, 1997 (Série Princípios).

VAZZATA-DIAS, J. F. A Concordância de Número nos Predicativos/Participios na Fala do Sul do Brasil – motivações extralingüísticas. *Letras de Hoje*, Porto Alegre, v. 35, n. 1, p. 209-228, mar., 2000.

VEADO, R. M. A. *Comportamento Lingüístico do Dialeto Rural*. Belo Horizonte: UFMG/PROED, 1982.

VOTRE. S. Escolaridade. In: MOLLICA, M. C. (Org.) *Introdução à Sociolingüística variacionista*. Cadernos didáticos da UFRJ. 2. ed. Rio de Janeiro, p. 75-79, 1994.