

UNIVERSIDADE ESTADUAL PAULISTA
INSTITUTO DE BIOCÊNCIAS, LETRAS E CIÊNCIAS EXATAS
CAMPUS DE SÃO JOSÉ DO RIO PRETO

MIRCIA HERMENEGILDO SALOMÃO

A VARIAÇÃO DE PLURALIDADE NAS ESTRUTURAS PREDICATIVAS
DA VARIEDADE FALADA NA REGIÃO DE SÃO JOSÉ DO RIO PRETO

São José do Rio Preto - SP
2010

MIRCIA HERMENEGILDO SALOMÃO

A VARIAÇÃO DE PLURALIDADE NAS ESTRUTURAS PREDICATIVAS
DA VARIEDADE FALADA NA REGIÃO DE SÃO JOSÉ DO RIO PRETO

Dissertação apresentada ao Instituto de
Biociências, Letras e Ciências Exatas da
Universidade Estadual Paulista, Campus de São
José do Rio Preto, para obtenção do título de
Mestre em Estudos Linguísticos (Área de
Concentração: Análise Linguística)

Orientador: Prof. Dr. Roberto Gomes Camacho

São José do Rio Preto - SP
2010

Salomão, Mircia Hermenegildo.

A variação de pluralidade nas estruturas predicativas da variedade falada na região de São José do Rio Preto / Mircia Hermenegildo Salomão. - São José do Rio Preto : [s.n.], 2010.

162f. : il. ; 30 cm.

Orientador: Roberto Gomes Camacho

Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Linguística. 2. Gramática comparada e geral - Sintaxe. 3. Língua portuguesa - Predicativos. 4. Língua portuguesa – Português falado. I. Camacho, Roberto Gomes. II. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU – 81'367

Ficha catalográfica elaborada pela Biblioteca do IBILCE
Campus de São José do Rio Preto - UNESP

BANCA EXAMINADORA

MEMBROS TITULARES

Prof. Dr. Roberto Gomes Camacho - orientador
UNESP – São José do Rio Preto

Profa. Dra. Luciane Ester Tenani
UNESP – São José do Rio Preto

Prof. Dr. Ronald Beline Mendes
USP - São Paulo

MEMBROS SUPLENTE

Profa. Dra. Gisele Cássia de Sousa
UNESP – São José do Rio Preto

Profa. Dra. Angela Cecília de Souza Rodrigues
USP - São Paulo

São José do Rio Preto, 10 de setembro de 2010

DEDICATÓRIA

Dedico este trabalho primeiramente a Deus, pois sem ele, nada disso seria possível e não estaríamos reunidos para desfrutarmos juntos destes momentos que nos são tão importantes.

Aos meus pais Orestes e Fátima, pelo amor incondicional, pela dedicação e compreensão, em todos os momentos dessa caminhada.

Ao meu irmão Miredson, pela paciência, cooperação e apoio durante esses dois anos de luta.

Dedico este trabalho também ao Wagner, meu namorado, não só por suportar minhas queixas nos momentos mais difíceis, mas também pelo carinho e amizade compartilhados comigo durante a graduação e o mestrado.

A todos os familiares que participaram dessa empreitada: meus padrinhos Alcides e Yvone, Nelson, Teresinha e toda a sua família. Aos Amigos Beto e Divina, Lucas e Gustavo.

A todos meus amigos de caminhada, principalmente a Ana Helena, Ana Lídia, Geisa, Juliana, Marcelo, Jesuelem, Viviane, Raquel e Bárbara, por compartilharem comigo momentos de alegrias e tristezas nessa etapa, que graças a Deus, está sendo vencida.

Agradeço, de maneira especial, meu amigo Fernando, que sempre me ajudou a organizar melhor as minhas idéias durante a graduação. Agradeço também seu empenho na tradução do meu resumo para língua inglesa.

AGRADECIMENTOS

Agradeço especialmente ao Prof. Dr. Roberto Gomes Camacho, meu orientador, pela oportunidade de aprender a cada dia mais sobre Linguística. Agradeço, pela paciência, tranquilidade e compreensão nestes três anos de orientação, pelos “puxões de orelha” dados nos momentos certos e por me ensinar a importância do trabalho e do conhecimento construído dia após dia.

Ao Prof. Dr. Sebastião Carlos Leite Gonçalves, pela amizade e contribuição dada para o desenvolvimento dessa dissertação.

À Profa. Dra. Gisele Cássia de Sousa, à Profa. Dra. Luciane Ester Tenani e ao Prof. Dr. Ronald Beline Mendes pela leitura e pelos valiosos comentários no Exame de Qualificação e na Banca de Defesa.

À Profa. Dra. Roxana Guadalupe Herrera Alvarez, que me mostrou a importância da pesquisa e da busca de conhecimento em todos os aspectos.

Às Professoras Clélia Jubran, Erotilde G. Pezatti, Maria Helena Viera-Abraão, que, de forma direta, me ajudaram a expandir os conhecimentos nas diversas áreas da linguística.

Agradeço à Profa. Dra. Cláudia Brescancini (PUC-RS), pela compreensão e pela ajuda na utilização do pacote estatístico GOLDVARB. Sem essa ajuda provavelmente não conseguiria terminar as análises deste trabalho.

A todos os professores do Departamento de Estudos Linguísticos e Literários, do Departamento de Educação e do Departamento de Letras Modernas, que contribuíram na minha formação durante esses seis anos de Instituto.

Agradeço à Fundação de Amparo à Pesquisa do Estado de São Paulo por acreditar no meu trabalho e por me auxiliar financeiramente durante essa empreitada com a bolsa de mestrado (processo nº: 2007/57203-0)

SALOMÃO, Mircia Hermenegildo. **A variação de pluralidade nas estruturas predicativas da variedade falada na região de São José do Rio Preto**. 2010. 162f. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas. Universidade Estadual Paulista, São José do Rio Preto

RESUMO

A atuação de processos de redução e de supressão de segmentos consonantais em fenômenos redundantes, como a concordância nominal, tende a suprimir marcas de pluralidade, provocando a possibilidade de ambiguidade referencial. Um fenômeno fonológico que reflete o processo gramatical de concordância é o da supressão de /S/, que pode afetar a marcação de pluralidade em contexto nominal. Com base nisso, o objetivo deste trabalho é o de submeter a um tratamento variacionista, de base quantitativa, dados de marcação variável de plural no SN e no SA em contexto de predicativo, obtidos em corpus coletado na região de São José do Rio Preto. O trabalho procura examinar se a marcação de pluralidade nos predicativos pode ser explicada com base em motivações exclusivamente formais, ou exclusivamente funcionais, ou ainda, com base na interação entre ambas, que consistiriam, assim, em motivações em competição (DU BOIS, 1985). A análise dos resultados mostra, por um lado, que um princípio funcional, como o as Condições de Distintividade de Kiparsky (1972), atua positivamente na marcação variável de pluralidade, mas não tem força suficiente para governar isoladamente o processo. Mostra, por outro, que as motivações formais, gerenciadas pelo Princípio do Paralelismo Formal, também exercem influência significativa na marcação variável de plural no predicativo, mas não como força predominante, na medida em que se submete a uma influência significativa das variáveis extralinguísticas *escolaridade* e *idade*. Esse comportamento dos dados não permite dar a esse princípio o poder explanatório que Labov (1996) lhe atribui, isto é, como a atuação mecânica de forças inerentemente internas ao sistema linguístico. Por essa razão, a explicação mais plausível para a marcação variável de pluralidade nos predicativos, na variedade estudada, é a de que esse processo sofre a influência de motivações em competição, nos termos propostos por Du Bois (1985).

Palavras-chave: sintaxe de concordância, predicativo, paralelismo formal, variação de plural motivações em competição, condições de distintividade

SALOMÃO, Mircia Hermenegildo. **The plural variation in predicative structures in the variety spoken in the São José do Rio Preto region.** . 2010. 162f. Thesis (Master degree in Linguistics studies) - Instituto de Biociências, Letras e Ciências Exatas. Universidade Estadual Paulista, São José do Rio Preto, São Paulo, Brazil.

ABSTRACT

Both the reduction and the suppression of consonantal segments when those occur in contexts of nominal agreement tends to drop plural markers, thus creating referential ambiguities. Amongst those phenomena reflecting a process of grammatical agreement is the suppression of /S/ in spoken Portuguese, which characterizes a phonological phenomenon that may affect plurality marking in the NP. Considering that, the present work employs a variationist and quantitative method to analyse data containing instances of predicative variable plural marking both in the NP and AP taken from a corpus collected in the area of São José do Rio Preto. This research aims to attest to whether plurality marking in the predicative context can be explained in terms of exclusively formal or exclusively functional motivations, or else, in terms of the interaction between these two kinds of motivations, thus consisting in competing motivations (DU BOIS, 1985). On the one hand, the analysis of the results evinces that a functional principle such as Kiparsky's Distinctiveness Conditions (1972) acts in a positive way in the variable plural marking; yet it alone does not suffice to govern the entire process. On the other hand, this work demonstrates that formal motivations governed by the Principle of Formal Parallelism have a significant influence on predicative plural marking, but not in a predominant fashion, inasmuch as the latter is subject to significant influence of extralinguistic variables such as 'education' and 'age'. The situation of the data do not authorize that such a principle be bestowed the explanatory power that Labov (1996) ascribes to it, i.e., that idea of a mechanical acting of inherent forces within the inner linguistic system. For that reason, the most plausible explanation is that the predicative plural marking in the variety under scrutiny constitutes a process influenced by competing motivations, in Du Bois terms (1985).

Keywords: syntax of agreement, predicative, formal parallelism, plural variation, competing motivations, distinctiveness conditions

LISTA DE TABELAS

Tabela 1: Número de ocorrências analisadas e percentual de aplicação e não-aplicação da regra de concordância nominal nas estruturas predicativas e nos participios passivos.	98
Tabela 2: Relação entre a variável número de marcas precedentes no sujeito e no verbo e a marcação de pluralidade.....	99
Tabela 3: Relação entre a variável paralelismo formal no discurso e a marcação de pluralidade	104
Tabela 4: Relação entre a variável características formais do verbo e a marcação de pluralidade	106
Tabela 5: Relação entre o material interveniente e a marcação de pluralidade.....	109
Tabela 6: Relação entre a variável características formais do sujeito e a marcação de pluralidade	110
Tabela 7: Relação entre tipo estrutural do sujeito e a marcação de pluralidade.....	112
Tabela 8: Relação entre saliência fônica e marcação de pluralidade.....	114
Tabela 9: Relação entre configuração estrutural do predicativo/particípio e marcação de pluralidade	116
Tabela 10: Relação entre escolaridade e a marcação de pluralidade.....	118
Tabela 11: Relação entre a idade do informante e a marcação de pluralidade.....	120
Tabela 12: Relação entre a variável gênero do informante e a marcação de pluralidade	121
Tabela 13: Cruzamento das variáveis Idade x Escolaridade e a marcação de pluralidade	123
Tabela 14: Cruzamento das variáveis Características formais do verbo e Paralelismo Formal	125
Tabela 15: Distribuição dos predicativos com plural explícito nos quatro níveis de escolaridade	127
Tabela 16: Distribuição dos predicativos com plural explícito em relação ao gênero	128
Tabela 17: Distribuição dos predicativos com plural zero nos quatro níveis de escolaridade	128
Tabela 18: Distribuição dos predicativos com plural zero em quatro faixas etárias do Iboruna	129
Tabela 19: Cruzamento das variáveis Características formais do verbo e Escolaridade	131
Tabela 20: Distribuição dos verbos sem marca de pluralidade do Ensino Superior em quatro faixas etárias	133
Tabela 21: Cruzamento entre as variáveis Paralelismo Formal no discurso e Escolaridade .	135
Tabela 22: Distribuição dos predicativos precedidos de predicativos com plural explícito do Ensino Médio em quatro faixas etárias do Iboruna	137
Tabela 23: Distribuição dos predicativos precedidos de predicativos com plural explícito do Ensino Médio em relação à variável gênero	138
Tabela 24: Distribuição dos predicativos precedidos de predicativos com plural explícito do Ensino Superior em quatro faixas etárias do Iboruna	138

Tabela 25: Distribuição dos predicativos precedidos de predicativos com plural zero dos informantes do Ensino Médio em relação à variável idade	139
Tabela 26: Distribuição dos predicativos precedidos de predicativos com plural zero dos informantes do Ensino Médio em relação à variável gênero	140
Tabela 27: Distribuição dos predicativos precedidos de predicativos com plural zero dos informantes do Ensino Superior nas quatro faixas etárias do Banco de Dados Iboruna.....	140
Tabela 28: Distribuição dos predicativos precedidos de predicativos com plural zero dos informantes do Ensino Superior em relação à variável gênero	141

LISTA DE GRÁFICOS

Gráfico 1: Percentual geral de presença/ausência de CN nos predicativos e participios passivos	98
Gráfico 2: Peso relativo da relação entre número de marcas precedentes no sujeito e no verbo e a marcação de pluralidade	99
Gráfico 3: Peso Relativo da relação entre a variável paralelismo formal no discurso e marcação de plural	104
Gráfico 4: Peso Relativo da relação entre às características formais do verbo e marcação de pluralidade	106
Gráfico 5: Peso relativo da relação entre escolaridade e a marcação de pluralidade	118
Gráfico 6: Peso relativo da relação entre idade e a marcação de pluralidade	120
Gráfico 7: Peso Relativo do cruzamento das variáveis Idade e Escolaridade e a marcação de pluralidade	123
Gráfico 8: Peso relativo do cruzamento entre as variáveis Características formais do verbo e Paralelismo Formal no Discurso.....	126
Gráfico 9: Peso relativo do cruzamento entre as variáveis Características formais do verbo e Escolaridade	132
Gráfico 10: Peso relativo do cruzamento entre as variáveis Paralelismo Formal no Discurso e Escolaridade	135

LISTA DE QUADROS, FIGURAS E ABREVIATURAS

QUADROS

Quadro 1: Variáveis linguísticas e extralinguísticas utilizadas nos trabalhos de Scherre (1991) e Dias (1996)	48
Quadro 2: Variáveis controladas na constituição da Amostra Censo	73
Quadro 3: Distribuição e identificação dos informantes da Amostra Censo por variáveis sociais (RUBIO, 2008).....	74
Quadro 4: Distribuição dos informantes proporcionalmente à densidade populacional das cidades da região.....	75
Quadro 6: Variáveis sociais controladas na constituição da subamostra.....	79

FIGURAS

Figura 1: Mapa da região administrativa de São José do Rio Preto	67
Figura 2: Localização de São José do Rio Preto na região Noroeste de São Paulo	67

ABREVIATURAS

CN - Concordância Nominal
CV - Com Concordância Verbal
PB - Português Brasileiro
SN - Sintagma Nominal
PR - Peso Relativo
AC - Amostra Censo ou Comunidade
AI - Amostra Interação
NURC - Norma Linguística Urbana Culta
VARSUL - Variação Linguística Urbana do Sul do país
PEUL - Programa de Estudos sobre o Uso da Língua

SUMÁRIO

INTRODUÇÃO	14
1. PRESSUPOSTOS TEÓRICOS	20
1.1. A Teoria da Variação e Mudança.....	20
1.1.1. O compromisso da análise sociolinguística com explicações funcionalistas.....	27
1.1.2. O compromisso da análise sociolinguística com explicações formalistas	30
1.2. Motivações em competição	34
1.3. A literatura sobre marcação de pluralidade	38
1.3.1. Literatura sobre a marcação de pluralidade no predicativo	45
1.3.2. A concordância nominal nas estruturas predicativas da fala urbana.....	46
1.3.3. A concordância nominal das estruturas predicativas e passivas em comunidades afro-brasileiras	57
1.4. O problema da norma e o preconceito linguístico	59
2. UNIVERSO DE PESQUISA E PROCEDIMENTOS METODOLÓGICOS	65
2.1. Caracterização histórico-social da região de São José do Rio Preto.....	65
2.2. O Projeto ALIP – Breve histórico	70
2.2.1. A constituição do Banco de Dados Iboruna	72
2.2.2. A Amostra Censo.....	73
2.3. Composição da amostra	76
2.3.1. Da seleção das ocorrências.....	80
2.3.2. Da definição da variável dependente	81
2.4. As variáveis linguísticas.....	83
2.5. As variáveis extralinguísticas	93
2.6. Da quantificação e da análise dos dados.....	95
3. ANÁLISE DOS RESULTADOS	97
3.1. Análise geral dos resultados.....	97
3.2. Análise dos fatores linguísticos estatisticamente relevantes	98
3.3. Análise dos fatores linguísticos não selecionados	108
3.4. Análise das variáveis extralinguísticas	117
3.5. Cruzamento entre fatores extralinguísticos	122
3.6. Cruzamentos entre fatores linguísticos	125
3.7. Cruzamento entre fatores linguísticos e fatores extralinguísticos	131
4. ALGUMAS GENERALIZAÇÕES E IMPLICAÇÕES TEÓRICAS.....	142
4.1. O efeito das motivações externas	142
4.2. O efeito das motivações internas	145

4.2.1. O Princípio do Paralelismo Formal	147
4.2.2. O princípio da Saliência Fônica	149
4.3. Motivações em Competição: Palavras Finais	150
REFERÊNCIAS BIBLIOGRÁFICAS	153

INTRODUÇÃO

Na década de 80 houve um desenvolvimento acentuado do estudo das relações de concordância no português do Brasil. A partir dos estudos da variação na marcação de pluralidade no português, sob o enfoque sociolinguístico, os trabalhos realizados, na esteira da pesquisa pioneira de Poplack, (1980) para o espanhol americano, tornaram possível formular princípios restritivos de natureza linguística e extralinguística na aplicação de regras variáveis como a de marcação de plural nos nomes. Essa regra é, muitas vezes, motivada por mecanismos de erosão fonológica, como processos de redução e/ou supressão consonantal, entre os quais se enquadra apagamento da marca de plural realizada por uma das pronúncias do arquifonema fricativo / S /.

A pesquisa sociolinguística e a dialetológica têm demonstrado que há outros fatores, não apenas de ordem semântica e gramatical, capazes de acelerar o processo de redução, como os fatores de natureza externa ao sistema linguístico. Dessa maneira, os fatores externos (grau de escolaridade, faixa etária e gênero), juntamente com os fatores internos, permitem acionar e acelerar o processo de redução de segmentos fonológicos, sendo extremamente relevantes para os estudos da concordância nominal e da variedade não-padrão.

É possível que os processos fonológicos de redução de segmentos fonológicos, como o de marcação de número no Sintagma Nominal (SN), tenham sido acelerados por conta da redundância desse mecanismo flexional, com a presença de marca formal de plural em todas as palavras constituintes de um sintagma; haja vista a intuição clara de que o /S/ de palavras como *lápiz* é frequentemente muito menos apagado do que núcleo de SNs como *os meninos*. Apesar dessa característica sintática, a variante padrão concorre, na prática quotidiana, com sentenças alternativas em que a indicação de pluralidade pode ocorrer num único constituinte, geralmente o que se localiza mais à esquerda do SN sujeito e, portanto, mais à esquerda da estrutura sentencial como um todo.

Em predicativos com a mesma estrutura de SNs não-predicativos, é possível afirmar que as restrições são as mesmas, mas outros fatores, como a marcação de plural no SN sujeito e no SV, interferem igualmente com esse processo variável, tratado nas gramáticas normativas como uma regra de concordância nominal e, como tal, de aplicação compulsória em todos os contextos de uso.

Esse conflito entre uma regra variável na fala e obrigatória na norma gramatical torna plenamente justificável o estudo da variação de pluralidade no predicativo, considerando, principalmente, que a ausência de certas marcas não provoca qualquer perturbação no valor funcional da sentença. Outra motivação considerável é a necessidade de se conhecer todas as variedades do português para que a dialetologia social possa elaborar um mapa completo dos mesmos processos em diferentes comunidades.

Sob a vertente variacionista, dentre os estudos já realizados, no português brasileiro, sobre a marcação variável de plural no SN é possível citar, em posição de destaque, os trabalhos de Scherre (1978; 1988; 1991; 1994; 1998a; 1998b; 2001; 2005), os de Scherre e Naro (1991; 1993; 1997; 2000; 2003b; 2006) e o de Braga (1976), sobre o dialeto carioca. Exclusivamente sobre predicativos, é possível mencionar os trabalhos de Dias (1996) e Dias e Fernandes (2000), sobre a variedade da região sul.

Essas referências mostram que, nessa área, a sociolinguística vem atuando de modo consistente no Brasil com o mapeamento de processos variáveis em vários dialetos do português brasileiro. Pode-se afirmar, entretanto, que os estudos ainda se encontram em um patamar distante de atingir um panorama completo da manifestação variável de plural em todas as variedades do português, devido à grande diversidade dialetal no território nacional, principalmente se consideradas as variedades faladas no interior dos estados, que ainda merecem um cuidado especial.

Além de ser reduzido o número de trabalhos, na dialetologia brasileira, sobre a marcação de plural no predicativo, é necessário investigar essa regra variável numa variedade dialetal ainda pouco conhecida, como a que se fala em uma parte da região noroeste do estado de São Paulo, nucleado em torno de São José do Rio Preto. Por isso, este trabalho se debruça sobre uma amostra de casos extraída do Córpus Iboruna, que justamente representa a fala dessa comunidade.

Além disso, como é recente o córpus em análise, não há estudos da concordância nominal em estruturas predicativas aplicados a ele, caracterizando, dessa maneira, uma justificativa relevante para este estudo, principalmente porque todo dialeto, linguisticamente heterogêneo, é único em relação a outros. O dialeto da região de Rio Preto apresenta características próprias como qualquer outro. Para exemplificar uma de suas marcas mais significativas, pode-se mencionar o “r” retroflexo, que constitui uma característica distintiva de outras regiões do estado, já investigadas por Guioti (2002).

Além dessa curiosidade dialetal, este trabalho pretende discutir a natureza explanatória de motivações funcionais e de motivações formais para o enfoque de processos variáveis. Explicações funcionais têm sido rechaçadas na pesquisa variacionista recente, em particular, por Labov (1996) e por Scherre (1988), em proveito de motivações de natureza estritamente formal, derivadas de mecanismos estruturais, internos, do sistema linguístico.

Por conseguinte, além de debruçar-se sobre a variação de plural no predicativo e dar notícia sobre a variedade falada em São José do Rio Preto, este trabalho tem por objetivo específico examinar que tipo de motivação é o mais relevante.

Esta pesquisa adota o modelo da Sociolinguística Quantitativa, aplicando-o a alguns princípios funcionalistas como as *condições de distintividade* (KIPARSKY, 1971) e o *princípio de economia* (HAIMAN, 1983), a fim de examinar as motivações (linguísticas e extralinguísticas) em competição (DU BOIS, 1985) na indicação de pluralidade nas estruturas

predicativas da variedade selecionada. Este trabalho procura, então, estabelecer a devida interação entre os dois tipos de motivações (formais e funcionais), assim comparáveis, para ver se elas consistem realmente em motivações em competição para o mesmo conjunto de dados. Além disso, como um interesse teórico, pretendemos contribuir, no final, para a discussão do alcance explanatório das motivações. São mais decisivas as motivações internas, de natureza intrinsecamente formal, ou devem ser elas matizadas por motivações internas de natureza funcional e por motivações externas de natureza extralinguística?

A teoria funcionalista tem grande relevância teórica para o desenvolvimento deste trabalho, na medida em que estudos como os de Du Bois (1985) argumentam a favor de motivações externas e internas em competição. É possível sustentar, desse modo, uma necessidade teórica de incorporar à pesquisa linguística questões como essas, que auxiliam na compreensão da natureza da interação complexa entre língua e fala, entre discurso e gramática.

Como o objeto de estudo deste trabalho, vale repetir, incide sobre a marcação de plural nas estruturas predicativas, é possível deduzir que o núcleo dessas construções pode ser ocupado por adjetivos, participios passivos e substantivos isolados, como nos exemplos (1), (2) e (3)

- (1) (...) e ela abandonô(u) ela foi embora não aguentô(u)... veja bem o quanto que... que esses filhos e esse marido foi **ruim** pra ela... (AC 110/NR/L 162-163)
- (2) (...) aí ela disse que o marido dela tinha sido afastado mas passô alguns dias... realmente... todos eles foram **mandado** embora... (AC 114/NR/L 270-271)
- (3) (...) ele foi brincá(r) assim com ela... – “ai vamo(s) ficá **noivo** vamo ficá **noivo**” – brincano e ela levô a sério e falô assim – “ai vamo” – ficô toda empolgada – “ai vamo vamo ficá **noivo**” – aí ela encheu o saco então teve que colocá aliança né? (AC 044/NR/L 139-142)

O objeto de estudo exclui, portanto, os predicativos com mais de um constituinte, exemplificados em (4) e (5) abaixo.

- (4) *Elas são crianças bastante agitadas*
- (5) *Os estrangeiros são pessoas muito bem educadas.*

As razões para essa decisão metodológica são as seguintes: primeiramente, há uma baixa frequência desse tipo de SN no corpus; em segundo lugar, haveria sérias dificuldades de medir o grau de saliência fônica de estruturas com diversos constituintes, já que seria necessário controlar esse fator com todos os elementos do SN.

Diante dos objetivos da pesquisa aqui relatados, este trabalho se encontra organizado basicamente em quatro seções. A seção I apresenta os pressupostos teóricos considerados para a elaboração dessa pesquisa, como a Teoria da Variação e Mudança Linguística, os estudos sobre a marcação de pluralidade no SN, os estudos sobre a CN nos predicativos, o compromisso dos estudos sociolinguísticos com as explicações formalistas e as funcionalistas e as Motivações em Competição, proposta por Du Bois (1985).

Já a seção II expõe os procedimentos metodológicos adotados. Primeiramente, caracteriza-se histórica e socialmente a região de São José do Rio Preto; em segundo lugar, faz-se um breve histórico sobre a criação do Projeto ALIP; em terceiro lugar, trata-se da constituição do Banco de Dados Iboruna para mostrar a constituição da subamostra utilizada nesta pesquisa. Finalmente, são descritos os fatores de análise, as hipóteses de trabalho e os procedimentos de quantificação dos dados.

A seção III apresenta a análise geral dos resultados e os cruzamentos mais relevantes entre os grupos de fatores estatisticamente significativos. Analisa-se os fatores estatisticamente relevantes, selecionados pelo pacote estatístico *Goldvarb*, e os fatores menos relevantes, a fim de se poder discutir algumas questões relacionadas, principalmente, às variáveis relevantes que não foram selecionadas pelo programa estatístico. Além disso, serão discutidos os cruzamentos dos fatores extralinguísticos com os fatores linguísticos, na

tentativa de detectar a influência das variáveis sociais sobre o princípio do Paralelismo Formal.

A seção IV, por fim, traz uma análise mais refinada dos dados, com a discussão dos efeitos das motivações internas e das motivações externas. Discutem-se, nesse caso, os efeitos dos Princípios do Paralelismo Formal e do Princípio de Saliência Fônica com a finalidade de mostrar uma atenuação da força a ele atribuído por Labov (1996), em função da influência direta de fatores sociais. É importante observar que a seção IV reforça a importância das motivações formais e funcionais em competição, nos termos propostos por Du Bois (1985). Após a ampla discussão sobre as motivações em competição, a seção IV se encerra com as palavras finais e os encaminhamentos futuros deste trabalho.

Após essa apresentação geral da organização deste trabalho, as páginas seguintes, contidas na seção I, tratarão do arcabouço teórico que rege os procedimentos metodológicos da pesquisa.

1. PRESSUPOSTOS TEÓRICOS

A seção, a seguir, vem dividida em duas partes. A primeira trata da sociolinguística variacionista e as implicações funcionalistas e formalistas em relação ao princípio de motivações em competição, que podem ser de grande relevância para este trabalho. Já a segunda, traz uma revisão dos estudos sobre concordância nominal e sobre os estudos de marcação de pluralidade no predicativo, foco desta pesquisa. Ainda nessa seção, discute-se a noção de preconceito linguístico, questão de natureza ideológica profundamente arraigada na regra normativa de concordância, principalmente na variedade aqui abordada.

1.1 A Teoria da Variação e Mudança

Cada língua é um produto social, e cada sociedade constitui-se por meio da linguagem. A sociolinguística está preocupada com a “vida real” das questões da língua no contexto social. Embora a linguística formal construa uma linguagem simplificada cujo comportamento pode ser previsto, a sociolinguística tenta lidar com a língua como um fenômeno social. Foi Saussure, o pioneiro da linguística estrutural, como uma ciência autônoma, que identificou a linguagem como um fato social. A sociolinguística pode ser entendida como uma tentativa de levar a sério a caracterização da língua feita por Saussure, pois o estruturalismo adotou uma versão forte do “axioma da categoricidade” (CHAMBERS, 1996), ou seja, os dados de base de natureza linguística deveriam eliminar a variabilidade existente na linguagem em proveito de uma visão homogênea.

Dessa forma, a fala foi excluída dos trabalhos, pois, segundo Saussure, ela é “acessória e acidental”. Weinreich, Labov e Herzog (2006) argumentam que a teoria saussureana não consegue acomodar a heterogeneidade inerente da linguagem justamente por tratar a linguagem como um sistema homogêneo. Saussure entende a heterogeneidade dentro do uso

linguístico de uma comunidade como um tipo tolerável de imprecisão de desempenho e não como sujeita a uma descrição sistemática.

A sociolinguística variacionista surge em meados dos anos 60 como uma orientação linguística que vai questionar os conceitos do paradigma estrutural saussureano e criticar alguns pontos desse paradigma como a dicotomia língua x fala, recorte que permitia excluir a fala dos estudos linguísticos, por não se enquadrar no princípio de pertinência, que via a língua como objeto ideal e homogêneo. Dessa maneira, a sociolinguística rompe com essa idéia dicotômica entre língua e fala e passa a estudar a língua em uso, reconhecendo que a variação é inerente a ela. Mais do que simplesmente reconhecer o caráter constitutivo da variação, a sociolinguística veio demonstrar que a variação é regular e sistemática, longe, portanto, de se configurar como “caos” linguístico (TARALO, 1985).

Nesse período, imperava a premissa de que a variação não seria passível de sistematização. A geração que sucedeu Sapir adotou fortemente o axioma da categoricidade como uma hipótese simplificadora, em que os dados, na análise linguística, deviam ser regularizados para se eliminar a variabilidade do mundo real. Entre as muitas afirmações derivadas da adoção desse axioma destaca-se a de Joos (1950) como a mais transparente. Para o autor, a linguística deveria ser um tipo de matemática em que a incoerência é por definição impossível. Joos apresenta a seguinte elaboração:

Técnicas matemáticas habituais caem principalmente em duas classes, a contínua (por exemplo, o cálculo infinitesimal) e a discreta ou descontínua (por exemplo, teoria dos grupos finitos). Agora resulta que a matemática chamada "linguística" pertence à segunda classe. Ela nem mesmo admite qualquer compromisso com a continuidade como a estatística, ou a teoria dos grupos infinitos. A linguística é uma mecânica quântica, no sentido mais extremo. Todas as continuidades, todas as possibilidades de gradação infinitesimal, são impelidas para fora da linguística em uma direção ou outra. (JOOS, 1950, p.701-702 *apud* CHAMBERS, 1996, p.13)¹

¹ Ordinary mathematical techniques fall mostly into two classes, the continuous (e.g. the infinitesimal calculus) and the discrete or discontinuous (e.g. finite group theory). Now it will turn out that the mathematics called "linguistics" belongs to second class. It does not even make any compromise with continuity as statistics does, or infinite-group theory. Linguistics is a quantum mechanics in the most extreme sense. All continuities, all possibilities of infinitesimal gradation, are shoved outside of linguistics in one direction or the other.

Chambers (1996), por seu lado, ressalta que a categoricidade é apenas um postulado que se pode desejar manter ou não; como tal, constitui um verdadeiro axioma que deve ser admitido como uma condição necessária para o desenvolvimento da análise linguística.

Para os pesquisadores estruturalistas, não era difícil ignorar a variação porque, linguisticamente, as sentenças significariam a mesma coisa a despeito de que variante fosse usada. O princípio de pertinência admitia que seriam passíveis de descrição e de análise as unidades que contivessem alguma informação relevante para a língua, considerada um meio de comunicação. Esse princípio de pertinência, que atraiu a atenção dos linguistas, fez com que as variantes fossem tratadas como formas livres ou facultativas, pois, para os estruturalistas, as variantes não passavam de flutuação aleatória. De acordo com Camacho

A procura por um objeto homogêneo levou também o estruturalismo a desviar seu foco de atenção para segmentos cada vez menores de linguagem, o que veio a culminar com a introdução do termo *idioleto* no arcabouço teórico da linguística norte-americana, significando o conjunto de usos de uma mesma língua que um falante faz num determinado momento da interação. (CAMACHO, 2010, p. 6)

Entretanto, Labov (1972) entende que essa busca pela noção do idioleto, como objeto de estudos, acaba por negar um conceito basilar da própria linguística estruturalista, que é o próprio conceito de língua como instituição social.

Outro paradigma com o qual a Sociolinguística rompe na década de 60 é o Gerativismo. Nesse paradigma, a variação linguística foi contornada por meio dos juízos intuitivos do falante-ouvinte sobre a gramaticalidade das construções. Entretanto, a suposta uniformidade dos julgamentos intuitivos é capaz de identificar apenas variáveis linguísticas já implementadas e que recebem correção social aberta (LABOV, 1972).

As críticas levantadas por Labov são dirigidas tanto para o parâmetro metodológico da teoria estruturalista quanto para o da teoria gerativista, mais especificamente para a relação de

adequação entre teoria e dados. Para Chomsky (1975), a teoria linguística é subdeterminada pelos dados, sendo necessário criar medidas de avaliação interna para selecionar a melhor alternativa entre as muitas teorias explanatórias do mesmo fenômeno. Já para Labov, somente o estudo da linguagem no próprio contexto social fornece meios para decidir qual análise é a correta dentre as possíveis. Como forma de sustentar essa posição, é que Labov introduz o conceito da variação nos estudos linguísticos de modo definitivo, tratando-a como propriedade inerente e constitutiva da linguagem.

A sociolinguística variacionista, também denominada de Teoria da Variação, leva sempre em conta um conjunto de pressupostos metodológicos. Um deles, já mencionado anteriormente, dissocia a estrutura linguística de homogeneidade. O procedimento metodológico que essa concepção aciona é a não aceitação da intuição como fonte dos dados linguísticos, já que, na qualidade de objeto social, a investigação da língua leva necessariamente em conta sua inserção no contexto social. Segundo Scherre (1997), esse procedimento implica dar conta da estrutura gramatical do discurso a partir da conversação natural, de narrativas formais, de gêneros escritos diversos etc. Uma consequência desse princípio é a atribuição de importância metodológica à quantificação dos dados em função de restrições linguísticas e não-linguísticas, que operacionalizam a confirmação ou a refutação das hipóteses.

Outro pressuposto teórico relevante é o de que essa concepção de linguagem, que acomoda fenômenos variáveis em relação a restrições linguísticas e não-linguísticas, conduz a descrições mais adequadas e resolve questões ligadas à mudança linguística. Nesse caso, considera-se que a compreensão da mudança, em especial, a da mudança sintática, só é possível se considerada a fronteira entre o linguístico e o extralinguístico. Como consequência direta dessa posição, sobressai a aceitação não só de motivações internas, mas também de

motivações externas, com base no postulado de que a explicação dos fatos linguísticos pode estar fora do sistema linguístico.

De forma geral, as formas em variação recebem o nome de ‘variantes linguísticas’ e as variantes podem constituir metodologicamente o que, em sociolinguística variacionista, denomina-se ‘variável linguística’, concebida como uma classe ordenada de variantes. No caso específico deste trabalho, constitui a variável dependente as variantes ausência e presença das marcas formais de pluralidade no predicativo. A frequência de uso das duas variantes é influenciada por fatores linguísticos, ou internos, e por fatores extralinguísticos, ou externos, que, no final, constituem o conjunto das variáveis independentes ou fatores condicionadores.

O estudo sociolinguístico requer uma descrição quantitativa, estatisticamente fundamentada, de um fenômeno variável, que forneça subsídios para a análise e a sistematização do que representava, na perspectiva estruturalista, o aparente “caos” da língua falada. Essa sistematização tem por principal objetivo mostrar a regularidade estrutural da variável e estabelecer a relação entre processos sincrônicos de variação e uma perspectiva diacrônica, de modo tal que fenômenos em variação num recorte sincrônico podem representar processos de mudança em progresso.

A concepção teórica da sociolinguística variacionista se pauta pela preocupação com o exame da linguagem em seu contexto social, como uma necessidade epistemológica de resolver problemas inerentes à área coberta pela teoria geral da linguagem. A sociolinguística defende a idéia de que respostas a questões relacionadas à estrutura da linguagem só podem ser obtidas mediante uma abordagem que pressuponha a interrelação da linguagem com a estrutura social (LABOV, 1972). De acordo com Camacho (2008),

É a abordagem dos fenômenos assim definidos que constitui o universo da sociolinguística. Trata-se mais especificamente, do estudo da estrutura e da

evolução da linguagem em uso no contexto social da própria comunidade de fala. Os tópicos desse tipo de enfoque recobrem, segundo Labov (1972), a área convencionalmente denominada de “linguística geral”, por lidar com fenômenos decorrentes dos níveis fonológico, morfológico, sintático e semântico, de modo a esclarecer a configuração de regras linguísticas, sua combinação em sistemas, a coexistência de diversos sistemas alternativos, assim como, e principalmente a evolução diacrônica de tais regras e sistemas (CAMACHO, 2008, p.16)

Camacho (2008) observa também que o estruturalismo autônomo (DU BOIS, 1985), que pode ser vinculado aos paradigmas gerativista e estruturalista, busca explicações num objeto construído a partir de um ponto de vista idealista, o que implica ignorar a realidade empírica do fenômeno observado, cuja natureza inerentemente social permanece, segundo o autor, parcialmente obscurecida.

Chomsky (1975), em relação ao estruturalismo saussureano, inaugurou um novo ponto de vista sobre o objeto, modificando a sua natureza. O objeto, de natureza social, foi transformado, em sua teoria, em psicológico. No entanto, a concepção de competência não foi capaz de ultrapassar a noção de um conjunto de regras idênticas por conta da idealização que opera sobre os fenômenos da linguagem, aspecto que a aproxima da noção de *langue* de Saussure (1997). Dessa maneira, a metodologia do estruturalismo autônomo ganha uma resposta limitada, em relação ao funcionamento da linguagem, pois é examinada sob a óptica de sistemas de signos, que é social, ou da competência, que é psicológica, já que representa o conhecimento intuitivo de que dispõe o usuário.

Calvet (1977) consegue exprimir as palavras certas para esse impasse, pois, para o autor, dar uma resposta baseado no estudo do sistema, como faz o estruturalismo, é afirmar que a linguagem funciona como instrumento de comunicação, sob o risco de omitir, de uma só vez, a sociedade, o indivíduo e as múltiplas funções da linguagem. Dar, por outro lado, uma resposta com base no estudo da competência significa acentuar o indivíduo produtor,

mas visto somente da perspectiva do modelo criado de produção, o que torna esse indivíduo independente das determinações sociais e psicológicas (CALVET, 1977, p.62)

De acordo com Camacho (2008), Labov entende o processo de desenvolvimento de concepções monolíticas de linguagem como a desenvolvida pelo paradigma vigente nos anos 60, ou seja, o gerativismo, baseando-se em duas suposições epistemológicas mais ou menos explícitas.

Em primeiro lugar, a estrutura linguística está rigorosamente associada com a homogeneidade e, nesse caso, uma teoria linguística adequada pode ser completamente desenvolvida com base na porção de comportamento linguístico uniforme e homogêneo, que só pode ser assim definido com base numa severa idealização dos fenômenos observados, o que exclui a variação. Em segundo lugar, supõe-se que os usuários de uma língua tenham amplo acesso a suas intuições sobre as regras do sistema e são perfeitamente capazes de relatá-las adequadamente. (CAMACHO, 2008, p.17)

Labov (1972), no entanto, acredita que o estudo dos fenômenos linguísticos mediante a análise das intuições é insuficiente, porque não fornece um modo adequado para descobrir se o modelo postulado está adequado ou menos adequado. O autor alega que é mediante o estudo da linguagem no contexto social que a quantidade de dados se expande em grande medida, de modo a fornecer os meios para avaliar qual explicação é a correta entre um conjunto de alternativas possíveis.

Labov chega a uma alternativa teórica para resolver os problemas da estrutura linguística. O autor inclui a concepção de que o sistema linguístico, inerentemente variável, é constituído pelo conjunto das formas que se manifestam na situação social por falantes em processo real de comunicação. Nesse aspecto, a análise da língua envolve a adoção de fatores não apenas linguísticos, mas também extralinguísticos, como os de natureza social e estilística. Esse modo de ver significa adotar a concepção de que operam na linguagem

motivações em competição, que podem ter natureza interna e externa. Essas motivações participam ativamente na aplicação de uma regra, podendo favorecê-la ou não.

1.1.1 O compromisso da análise sociolinguística com explicações funcionalistas

Observa-se, até este momento, que o objeto de estudo e sua delimitação sempre estão no centro das discussões dos principais paradigmas vigentes, o formalista, praticamente dominado pela Gramática Gerativa, e o não-formalista, que inclui várias disciplinas científicas, como a Sintaxe Funcionalista, a Sociolinguística, a Análise da Conversação, a Sociolinguística Interacional e as várias tendências da Análise do Discurso. Parece relevante refletir sobre as concepções formalista e funcionalista, já que este trabalho objetiva discutir as motivações formais e funcionais no fenômeno de concordância nominal em predicativos.

A natureza e a delimitação do objeto sempre foram focos de grandes debates na atividade científica. Antes de esclarecer as oposições entre essas duas correntes, serão explicitadas as diferenças entre as duas abordagens teóricas que aparecem devido ao modo distinto de se ver a natureza da linguagem.

Primeiramente, o formalismo considera a linguagem um fenômeno mental, enquanto o funcionalismo a considera como um fenômeno social. Na concepção formalista, os universais linguísticos derivam de uma herança linguística geneticamente comum ao ser humano. Para o funcionalismo, por outro lado, esses universais derivam da universalidade dos usos a que a linguagem está sujeita nas sociedades humanas. No que se refere ao processo da aquisição da linguagem, o formalismo a identifica como uma capacidade inata da mente humana, enquanto o funcionalismo, como um desenvolvimento das necessidades e habilidades comunicativas da criança na sociedade (CAMACHO, 2003).

O uso da língua tem um significado muito diferente para a concepção formalista. Essa

corrente estuda a linguagem como um sistema autônomo, centralizado na sintaxe. Em contrapartida, os funcionalistas, mesmo reconhecendo o estatuto da sintaxe, veem-na necessariamente em relação à função que desempenha no contexto social. O formalista acredita que a sintaxe é autônoma em relação à semântica e as duas são igualmente autônomas em relação à pragmática; para eles, as prioridades vão da sintaxe à pragmática, via semântica. Todavia, um funcionalista vê as relações entre os componentes mais visíveis da organização linguística – o pragmático, o semântico e o sintático – de um modo tal que o primeiro é mais abrangente, dentro do qual se devem examinar os outros dois, não havendo espaço para um sintaxe autônoma (DIK, 1997).

A relação entre a estrutura e a função linguística é estabelecida pelo enfoque funcionalista da variação em termos de motivações (CAMACHO, 2008). Dessa maneira, um linguista funcionalmente orientado assume que há motivações sistemáticas, que estão sujeitas a princípios regulares subjacentes à conexão entre funções discursivas e estruturas morfossintáticas relevantes. As motivações são relacionadas ao modo como as estruturas se projetam nos conceitos correspondentes (iconicidade), à frequência e ao grau de acessibilidade dessas estruturas na mente (economia), à frequência e ao grau de acessibilidade de combinações individuais de estruturas e conceitos.

O princípio funcional de economia (HAIMAN, 1983), um dos conceitos base para deste trabalho, representa uma pressão para a simplificação máxima da expressão. A estrutura linguística torna-se, dessa forma, o mais transparente possível, contrapondo-se ao princípio da motivação icônica. Pode-se entender o princípio de economia sintagmática como uma tendência para a redução do comprimento ou da complexidade do enunciado, de maneira que as expressões mais frequentes no uso tendem a reduzir-se fonologicamente e a informação redundante ou recuperável no contexto comunicativo tende a ser suprimida.

Vale lembrar, a esse propósito, que o interesse teórico deste trabalho é verificar se o princípio da Economia é uma motivação em competição com o princípio do Paralelismo Formal, com o objetivo final de mostrar se a marcação de pluralidade no predicativo tem uma explicação metodológica situada na repetição de estruturas ou se a ausência de marcas de pluralidade em fenômenos de CN deve-se a uma explicação metodológica funcional, em função do princípio da Economia.

De acordo com Camacho (2008), situar a sociolinguística nesse quadro panorâmico deixa claro que a abordagem variacionista se inclui mais apropriadamente no espectro das tendências não-formalistas, pois

em sua fase inicial, a sociolinguística variacionista sustenta posições teórico-metodológicas que a distanciam da tendência formalista, inclusive como uma abordagem alternativa. Defende, nessa fase inicial, a covariação sistemática entre fatores sociais e linguísticos, reconhecendo, portanto, a possibilidade teórica de que fatores sistemáticos do contexto social interfiram na estrutura linguística, mediante a implementação, avaliação e transição de variáveis linguísticas. (CAMACHO, 2008, p.19)

Labov, em suas propostas iniciais, concebe a regra variável como um refinamento das regras opcionais mantidas pela teoria gerativa. O formalismo do gerativismo restringe-se apenas ao aspecto notacional da descrição. Por um lado, a sociolinguística variacionista inclui o alargamento da noção de competência, abrangendo, desse modo, a variação; por outro, o uso de amostras de fala real como dado, substituindo as intuições (LABOV, 2008).

A sociolinguística passa a coletar dados de situações reais de comunicação e se depara com o fato de que os membros de uma determinada comunidade de fala se distinguem em gênero, idade, ocupação, classe econômica, escolaridade, etc. O primeiro foco de interesse foi dirigido a esses condicionamentos externos, como os estudos sobre a variedade da ilha de Martha's Vineyard e o da re-inserção de /r/ em posição pós-vocálica na cidade de Nova Iorque, realizados pelo próprio Labov (1972; 2008).

A variável linguística foi considerada uma unidade estrutural da língua, fato que representou uma grande ruptura com a tradição e um momento de renovação teórico-metodológica, pois a unidade linguística, até então, tinha sido invariável, discreta e qualitativa e, a partir desse momento, ela passa a ser variável, contínua e quantitativa. É variável porque se realiza diferentemente em diversas circunstâncias; é contínua porque certas alternativas estão carregadas de significação social conforme a distância ou diferenciação fonética em relação à forma padrão; é quantitativa porque a variável é determinada pela frequência relativa de suas variantes. (CHAMBERS, 1996)

1.1.2 O compromisso da análise sociolinguística com explicações formalistas

A sociolinguística variacionista vai sustentar posições teórico-metodológicas que se distanciam, inicialmente, da tendência formalista, pois as primeiras propostas, esboçadas por Labov, concebem a regra variável como um refinamento das regras opcionais mantidas pela teoria gerativa. Labov (1972), em seus estudos, passa a coletar dados reais de fala, mostrando um interesse pelas diferenças de uma comunidade social, já que os indivíduos se diferenciam em classe social, sexo, idade, ocupação, escolaridade, etc.

Os estudos sociolinguísticos foram primeiramente aplicados a fenômenos fonológicos, em que o problema do significado não é pertinente. No entanto, quando os métodos e técnicas passaram a ser aplicados a casos de variação em sintaxe, houve a primeira crise na metodologia do paradigma variacionista, em razão de forte reação, provocada por Lavandera (1978), Romaine (1981) e García (1985).

Os ataques de Lavandera e García visaram o estatuto teórico da regra variável, quando aplicado à sintaxe. Já Romaine mirou o aspecto indutivista da teoria. Lavandera (1984) mostra, em seu artigo, que é inadequado alargar a outros níveis de análise da variação o

conceito de variável sociolinguística, que foi originariamente desenvolvida com base em dados fonológicos. Segundo a autora²

Os estudos quantitativos da variação que se ocupam de alternâncias morfológicas, sintáticas e léxicas se ressentem da falta de uma teoria bem organizada dos significados. Embora a análise da variação em fonologia, mediante a definição de variáveis fonológicas, possa ser admitida como uma contribuição para uma melhor compreensão dos tipos de informação, que podem comunicar as diferenças formais, deixa de ser reveladora, em muitos casos, a extensão da noção de variável para a variação não-fonológica. (LAVANDERA, 1984, p. 37, tradução nossa)

Para a autora, há diferenças cruciais entre variáveis fonológicas e variáveis não-fonológicas. As primeiras dispõem de significado social e estilístico e não precisam ter significado referencial. No entanto, as variáveis não-fonológicas veiculam significado social e estilístico, mas também têm significado referencial. Lavandera (1984, 1996) destaca a dificuldade metodológica de se lidar com variáveis não-fonológicas, principalmente por envolverem significação referencial. Dessa maneira, a autora propõe relaxar a condição de que o significado referencial deve ser o mesmo para todas as alternativas, introduzindo, assim, uma condição de comparabilidade funcional, já que a equivalência semântica entre duas variáveis sintáticas implicaria numa redução drástica da noção de significado referencial. A autora afirma que

(sua) proposta é enfraquecer a condição de que o significado deva ser o mesmo para todas as formas alternantes, e substituí-la por uma condição de comparabilidade funcional.

Além disso, preferiria não analisar as formas sintáticas alternantes como “variáveis sociolinguísticas”, ao menos que se apliquem outras duas condições a elas: (i) que se possa provar que elas são portadoras de significação social e estilística (...); e (ii) que se possa provar que elas são um tipo de instrumento da linguagem semelhante às variáveis fonológicas,

² Los estudios cuantitativos de variación que se ocupan de alternancias morfológicas, sintáticas y léxicas sufren de la falta de una teoría bien organizada de los significados. Si bien el análisis de la variación en fonología mediante la definición de variables fonológicas, puede aceptarse como una contribución a una mejor comprensión de los tipos de información que pueden comunicar las diferencias formales, en muchos casos la extensión de la noción de variable a la variación no-fonológica deja de ser reveladora.

isto é, elementos cuja propriedade definidora é uma covariação quantificável em relação as quais são as próprias relações de frequência que apontam as diferenças. (LAVANDERA, 1984, p.45-46, tradução nossa)³

Labov (1978), em resposta ao artigo de Lavandera, não concorda com a introdução dessa condição de comparabilidade funcional no modelo variacionista, ressaltando que os métodos de coleta de dados inicialmente fazem retornar uma pergunta a que Lavandera se refere no início de seu trabalho, como “Por que alguém diz alguma coisa?”. Embora Labov alegue que, a maior parte das vezes, a resposta é “para comunicar algum tipo de informação”, ele entende que essa resposta não é muito reveladora, já que gera outra pergunta “comunicar que tipo de informação?”. Nas próprias palavras do autor

Embora a linguística formal reconheça a existência das informações expressivas e afetivas, estas são na prática subordinadas ao que Bühler (1934) chamava por “significado representacional” ou aquilo que denominarei “estados de coisa”. Para ser mais preciso, eu gostaria de afirmar que dois enunciados que se referem ao mesmo estado de coisas têm o mesmo valor de verdade. (...) Em vez de alargar o significado como Lavandera sugere, queremos restringi-lo muito mais de que o faria uma linguística formal. (LABOV, 1978, p. 2, tradução nossa)⁴

O autor nega a dificuldade da expansão das técnicas e métodos da teoria variacionista às variantes não-fonológicas. Segundo Labov (1978), muitas variáveis sintáticas não fornecem qualquer problema para o estabelecimento do mesmo significado representacional.

³ Mi propuesta es debilitar la condición de que el significado deba ser el mismo para todas las formas alternantes, y reemplazarla por una condición de comparabilidad funcional.

Además, preferiría no analizar como “variables sociolingüísticas” formas sintácticas alternantes al menos que se den otras condiciones: (i) que pueda probarse que son las portadoras de significación social y estilística (...), y (ii) que pueda probarse que son un tipo de instrumento del lenguaje semejante a las variables fonológicas, es decir, elementos cuya propiedad definitoria es una covariación cuantificable y con respecto a los cuales las relaciones de frecuencia mismas son las que señalan diferencias.

⁴ Though formal linguistics recognizes the existence of expressive and affective information, these are in practice subordinated to what Bühler (1934) called “representational meaning” or what I will call “states of affairs.” To be more precise, I would like to say that two utterances that refer to same state of affairs have the same truth-value.(...)Instead of extending meaning as Lavandera suggests, we want to limit it much more narrowly than a formal linguistic will do.

Labov também recebeu críticas de Kay e McDaniel (1979). Esses autores se preocuparam com a incompatibilidade entre o modelo gerativista, que trabalha no nível da competência, e a teoria variacionista, que trabalha no nível do desempenho (CAMACHO, 2003). Segundo os autores, a principal falha do enfoque variacionista é o fato de se ter adotado do formalismo apenas o mecanismo descritivo, não considerando, desse modo, suas implicações teóricas.

Camacho (2008) afirma que Labov (1987, 1996), mesmo tendo sido alvo de muitas críticas, radicaliza ainda mais sua posição, em seus trabalhos posteriores, procurando reduzir o alcance das explicações funcionais ao enumerar uma série de argumentos com exemplos que favorecem a autonomia formal. O autor direciona críticas a alguns aspectos dos argumentos funcionais, revelando que

(...) Muitas teorias da linguagem, propostas recentemente, explicariam a estrutura linguística como resultado das intenções do falante para comunicar significado ao ouvinte. Há uma parte do comportamento linguístico que está sujeita ao controle consciente, uma escolha deliberada, ao comportamento decidido e reflexivo. Entretanto, até onde posso ver, não é uma parte principal da faculdade linguagem e tem relativamente pouca influência sobre a evolução da estrutura linguística a longo prazo. Não se trata de negar que nossas intenções possam ser as motivações para nossas ações; o problema é que isso não é tão acessível como alguns linguistas formais e reflexivos gostariam de acreditar (LABOV, 1996, p.909, tradução nossa)⁵

Labov ainda afirma que a necessidade de preservar informação é um argumento relativamente fraco que pode ser sobrepujado por outros fatores, principalmente os fatores formais.

⁵ Muchas teorías del lenguaje propuestas recientemente explicarían la estructura lingüística como resultado de las intenciones del hablante para comunicar significado al oyente. Hay una parte del comportamiento lingüístico que está sujeta a control consciente, a elección deliberada, al comportamiento resuelto y reflexivo. Pero hasta donde alcanzo ver, no es una parte principal de la facultad del lenguaje, y tiene relativamente poca influencia sobre la evolución a largo plazo de la estructura lingüística. No es ello negar que tenemos intenciones, y cuáles puedan ser las motivaciones para nuestras acciones, no es tan accesible como a algunos formales y reflexivos lingüistas les gustaría creer.

1.2 Motivações em competição

A discussão das vertentes formalistas e funcionalistas e suas implicações dentro da teoria variacionista envolve necessariamente discutir a possibilidade de atuarem motivações em competição. Quem primeiramente utilizou essa noção foi Du Bois (1985) numa obra em que o autor discute amplamente a relação entre iconicidade e arbitrariedade. No formalismo do estruturalismo clássico, a iconicidade se revela uma anomalia que operaria contra a visão de que a linguagem é um sistema inteiramente auto-suficiente.

A postulação do princípio da arbitrariedade do signo linguístico feita por Saussure é extremamente importante por restringir a conexão do signo com os fatos positivos do mundo externo, tornando-o regulável por forças completamente internas e autônomas do próprio sistema linguístico em que reinam soberanas as relações opostas e negativas entre os signos no âmbito de relações paradigmáticas.

Em função desse princípio de autonomia do sistema, Saussure separa nitidamente os fatos que devem compor a linguística interna dos fatos que compõem o que ele denomina de linguística externa; esta deveria incluir fenômenos dos tipos mais diversos que estão no escopo de áreas atuais como a Fonética Articulatória, a Etnografia da Linguagem, a Sociolinguística, a Geografia Linguística, a Linguística Antropológica e a Análise do Discurso. Como os princípios de organização que operam num domínio não afetam o outro, a *langue* seria, em princípio, independente das forças estruturadoras que se originam fora dela.

No arcabouço teórico do estruturalismo clássico, que Du Bois (1985) chama de “estruturalismo autônomo”, a iconicidade é claramente considerada uma anomalia por operar contra a visão de que a linguagem é um sistema completamente autônomo com o conceito de arbitrariedade no centro. Dessa forma, nenhum grau de iconicidade é permitido, a não ser os emanados das próprias relações internas ao sistema, nem mesmo os signos onomatopéicos,

que são vistos também como produtos das convenções que dirigem as relações arbitrárias no interior de cada sistema linguístico.

O estruturalismo autônomo estabelece uma forte distinção entre a linguística interna ou linguística propriamente dita, que é o estudo da língua, e a linguística externa, que abarca um conjunto diversificado de aspectos mais ligados à esfera da fala. Entende Du Bois (1985) que essa distinção é necessária, já que é nela que esse autor se apóia para mostrar que certos fenômenos podem ser simultaneamente imotivados a partir do ponto de vista sincrônico e motivados a partir de um ponto de vista metagramatical, que trata as gramáticas como sistemas adaptativos, na medida em que são tanto parcialmente autônomos, e daí *sistemas*, quanto parcialmente suscetíveis a pressões externas ao sistema, daí *adaptativos*. Segundo o autor, essa distinção é frutífera apenas se reconhecermos a existência de motivações em competição e o subsequente desenvolvimento de arcabouço teórico para, inicialmente, descrever e analisar a interação das motivações com contextos específicos e, posteriormente, para prever a resolução da competição entre elas.

Nessa perspectiva chamada de “estruturalismo autônomo” por Du Bois (1985), as únicas forças que organizam a linguagem são as internas ao sistema, consideradas por Saussure (1997) como forças que operam permanente e universalmente em todas as línguas, e nenhum tipo de motivação externa é capaz de atingir o sistema. O principal arcabouço teórico a herdar de Saussure esse forte comprometimento com a autonomia da gramática é a Gramática Gerativa em suas diversas versões.

Em contraste com essa perspectiva, o “funcionalismo transparente” assume o princípio de que fatores sintáticos aparentemente autônomos constituem, na realidade, os resultados transparentes dos objetivos funcionais do falante, de modo que não é necessário postular nenhum mecanismo sintático arbitrário para a língua. Em outros termos: as únicas forças que governam a “sintaxe” são as forças positivas da linguística externa, exemplificadas por Du

Bois como mecanismos de processamento de base biológica, intenções do falante etc. (DU BOIS, 1985).

Na medida em que sejam reconhecidas como teorias, o estruturalismo autônomo e o funcionalismo transparente devem ser entendidos como inerentemente opostos e aparentemente irreconciliáveis; entretanto, estudos recentes em gramaticalização têm posto ênfase nas relações entre discurso e gramática, e tentado integrar parcialmente esses dois enfoques.

A gramaticalização, conforme Du Bois (1985), é um processo de mudança de construções relativamente livres no discurso, cuja forma idiossincrática é motivada somente pelos objetivos do falante em função do evento de fala imediato, para construções relativamente fixas na gramática. Na visão de Du Bois (1985), um tratamento adequado do processo de gramaticalização numa teoria que preveja a relação entre discurso e gramática, terá que forçosamente reconhecer que, embora a gramaticalização esteja envolvida com influências dos fenômenos do discurso externos ao sistema, a própria noção de gramaticalização implica que essas influências interagem com as da língua em si como um sistema gramatical. (DU BOIS, 1985, p. 347).

Um pré-requisito para o enfoque da variação multidimensional de categorias linguísticas, como a marcação de pluralidade, é o princípio de, que nesse processo, interagem motivações em competição. Duas motivações que competem para o controle de um único paradigma linguístico, categoria ou morfema podem ser tanto ambas internas quanto ambas externas ou mesmo pode ocorrer que uma das motivações seja interna enquanto a outra seja externa.

Em primeiro lugar, para duas motivações competirem no mesmo âmbito funcional deve haver o que Du Bois (1985) chama de “bem limitado”, pelo qual elas competem. O

“bem limitado” é o domínio da forma linguística, sobre as quais forças motivadoras múltiplas não podem prevalecer completamente.

Segundo Du Bois, o grau de preparação de que dispomos para aferir o significado da teoria linguística se dá a partir de uma perspectiva que vê as gramáticas como sistemas adaptativos e que, portanto, reconhece, em primeiro lugar, a interação entre forças internas e externas. Quando as forças que se originam de fenômenos externos penetram no domínio da língua, elas encontram e interagem com as forças internas. Os fenômenos internos, reconhecidos como intrinsecamente linguísticos, devem ser, na perspectiva de Du Bois (1985), forças dinâmicas em vez de estruturas fixas, categorias ou entidades de qualquer tipo.

Quando se diz que a linguagem é adaptativa, na medida em que responde a pressões do contexto externo, deve-se de pronto rejeitar a teoria do estruturalismo autônomo, que é limitada na tentativa de explicar a estruturação da linguagem, por reconhecer apenas forças motivadoras internas. Quando se diz, por outro lado, que a linguagem é um sistema, e que, como tal, tem certo grau de continuidade de existência, deve-se rejeitar o arcabouço teórico do funcionalismo transparente, que falha em explicar o fato de que categorias gramaticalizadas são preservadas, por assim dizer, numa forma reificada mais ou menos congelada. Mais importante ainda que isso talvez, ele falha em construir apropriadamente o aspecto interno mais fundamental da gramática, o processo de gramaticalização em si mesmo.

Como alternativa para as versões rigidamente ortodoxas do estruturalismo autônomo e do funcionalismo transparente, Du Bois (1985) propõe a busca de uma síntese entre os dois arcabouços, um novo marco teórico que seja capaz de lidar com a interação de forças externas e internas, assim como com a resolução sistemática da competição que as envolve. A tarefa metodológica inicial é construir um arcabouço para uma linguística real da fala, que seja capaz de incorporar as descobertas dos estudos modernos do discurso.

A compreensão dessas diferenças tem uma importância fundamental para os objetivos

desta pesquisa, pois, em grande medida, foram essas razões teóricas que organizaram os procedimentos metodológicos para o processamento dos dados, ou seja, com base em uma metodologia que favorecesse o enfoque da marcação de pluralidade nas estruturas predicativas a partir de duas diferentes perspectivas, a formal e a funcional. A dimensão formal explica as motivações internas, enquanto a dimensão funcional, acrescida dos fatores extralinguísticos, explica as motivações externas. Nesse caso, o ‘bem limitado’, na visão de Du Bois (1985), é o domínio da variável em busca de resolução seja em direção da categorialidade mediante predomínio de uma das variantes, seja em direção de estabilidade mediante distribuição por diferentes comunidades sociais como indício de identidade cultural.

1.3 A literatura sobre marcação de pluralidade

De acordo com a norma padrão do português brasileiro, o fenômeno de concordância é considerado regra obrigatória. Como a concordância reitera marcas, a variante padrão é necessariamente redundante, já que repete as marcas contendo a mesma informação em pontos diversos da cadeia sintagmática. Traduzindo a tradição gramatical, Scherre afirma o seguinte:

Na concordância dentro do SN, colocam-se marcas explícitas de plural em todos os seus elementos flexionáveis quando o núcleo do sintagma for formalmente plural; na concordância do predicativo com o sujeito, repetem-se marcas formais de plural em todos os elementos flexionáveis dos predicativos quando o sujeito for formalmente plural; e na concordância verbal, colocam-se marcas explícitas de plural no verbo, quando o sujeito for formalmente plural ou quando for composto. (SCHERRE, 1997, p.182)

Por outro lado, é também verdadeiro que a concordância de número é uma variação marcada pela identidade social do falante e por motivações emanadas do contexto social. Uma motivação potencialmente relevante para não se usar a CN é o baixo nível de escolaridade que

impede as pessoas de passar pela pressão normativa da escola; outra ainda é a situação estilística: não se emprega a regra de concordância, em todos os momentos, num contexto social mais informal, como por exemplo, o de um ambiente familiar. Assim, o modo categórico como a tradição gramatical concebe a variação de pluralidade não se aplica de fato ao uso da língua no contexto social.

Como se sabe, à medida que a língua varia, é possível expressar um mesmo enunciado de modo diferente sem mudar o seu significado, como nas sentenças contidas em 6(a-b)

- (6) a *os orelhão (∅) de lá não é igual (∅) aqui (AC029/NE/L.069)*
 b *os orelhões de lá não são iguais aos daqui*

Na alternativa contida em (6a), a marca formal de pluralidade se encontra apenas no determinante, enquanto na contida em (6b) as marcas formais de pluralidade estão presentes no determinante e no núcleo do SN sujeito e do predicativo, além da marcação na cópula verbal. A ausência de marcas não é obstáculo para que, ainda assim, seja assegurada a noção de pluralidade sem provocar nenhuma ambiguidade tendo em vista a marcação de número no determinante.

É possível afirmar, então, que, de um ponto de vista funcional, a explicitação de pluralidade no determinante poderia tornar desnecessárias as marcas nos outros constituintes, por razão de economia, em função da redundância inerente da regra de concordância nominal na variedade padrão. Outra explicação possível seria uma motivação por reiteração de formas em paralelismo, tanto na marcação quanto na ausência de marcação (SCHERRE, 1988). Esse modo variável de marcar número leva a refletir sobre a motivação que estaria em jogo no processo de variação da concordância nominal; ou o paralelismo formal, na medida em que a indicação de pluralidade explícita, ou não explícita, levaria em conta a marcação do mesmo tipo na posição anterior; ou o princípio funcional de distintividade (KIPARSKY 1972), na medida em que a indicação redundante a partir do constituinte marcado acionaria a

desnecessidade de outras marcas subsequentes sem risco de ambiguidade referencial.

Seria possível também supor que poderia haver motivações em competição (DU BOIS, 1985), de natureza interna, ou formal, e externa, ou social. Nesse caso, seria possível supor que os casos alternativos que não apresentam todas as marcas de plural podem representar a escolha de diferentes grupos sociais, de modo que haveria grupos sociais em que predominaria a escolha de SNs nominais com marcas formais de plural e, em contraste, outros grupos sociais para os quais a presença de todas as marcas nem sempre é necessária. Segundo Lemle (1978), essa variação é motivada por fatores sociais diversos, conforme comprovam suas próprias palavras:

De um modo geral, pode-se dizer que os fatores determinantes da heterogeneidade linguística são três: geográfico, responsável pela divergência linguística entre comunidade fisicamente distantes uma da outra; o social, responsável pela divergência linguística entre distintos subgrupos de uma comunidade local, sendo fatores potencialmente distintivos a estratificação social, a faixa etária, o sexo, a ocupação profissional dos falantes, o desejo ou interesse que eles têm em manterem características linguísticas que os demarquem; o registro de uso, ou nível de formalidade atribuído ao encontro pelos interlocutores, numa gama que vai desde o mais coloquial ao mais formal. (LEMLE, 1978, p.61)

As duas variantes estão igualmente presentes na modalidade falada do português, mas a variante padrão, também conservadora, é vista como forma de prestígio e é usada por uma camada mais alta da sociedade, além de estar contida como a única forma correta nas gramáticas normativas e pedagógicas. Já a variante não-padrão, considerada estigmatizada, é, normalmente, utilizada pelas camadas menos privilegiadas socialmente.

Os estudos sociolinguísticos têm lidado, frequentemente, com esse fenômeno de ausência e de presença das marcas formais de plural tanto no Brasil quanto no exterior. Scherre (1988), em sua tese de doutorado, explicitou os fatores linguísticos e não-linguísticos que regem a distribuição sistemática da variação de número, tomando por base duas perspectivas: uma considerando como dado de análise cada constituinte do SN (perspectiva

atomística) e a outra tomando como dado o SN todo (perspectiva não-atomística).

Na primeira perspectiva, a autora examinou todos os itens lexicais passíveis de receber marca de pluralidade, de acordo com as seguintes condições variáveis: uma escala de saliência fônica, contendo três dimensões, e um conjunto de seis outros condicionamentos, de natureza diversa; esse conjunto inclui outros fatores morfológicos, como grau dos substantivos e adjetivos e classe gramatical dos constituintes, traços semânticos dos substantivos, grau de formalidade dos itens lexicais, fatores sintáticos, como função e posição de cada elemento do SN, e fatores fonológicos, relacionados ao contexto seguinte à marca de plural.

Já na segunda perspectiva, a autora considera SNs de no mínimo três constituintes, distinguindo aplicação da regra de não-aplicação, quando todos os elementos flexionáveis do SN receberem marcas formais de pluralidade. Nessa perspectiva, Scherre (1988) apresenta as seguintes variáveis: a estrutura sintagmática do SN; na função textual, a forma de apresentar a informação; a pluralidade do SN; a localização do SN em relação à oração; a pluralidade do contexto, e as influências globais de restrições localizadas, abordadas na perspectiva atomística.

Em sua análise sobre a pluralidade do contexto, a autora ressalta não considerar o verbo na codificação dos dados, enfatizando, desse modo, a pluralidade no sintagma nominal. Além disso, ela inclui todos os SNs, diferentemente de outras abordagens que consideram apenas o SN na função de sujeito. Scherre acredita que a pluralidade do contexto escalado como primeiro grupo de fatores, segundo o ponto de vista estatístico, encontra explicação discursivo-textual, já que há uma tendência de formas semelhantes aparecerem juntas, criando uma harmonia formal, que torna o discurso coeso.

Scherre (1994) afirma que o fenômeno da variação da concordância não é restrito apenas a uma região ou classe social específica, mas caracteriza toda a comunidade de fala

brasileira. A autora afirma, ainda, que é possível descrever e explicar um conjunto de variáveis linguísticas e não-linguísticas que regem a sistematicidade desse fenômeno, tornando evidente a existência de um gerenciamento da variação de acordo com fatores internos, linguísticos, e de acordo com fatores externos, referentes à situação social, com a qual estamos envolvidos.

As investigações no espanhol cubano realizadas por Terrel (1975), sobre as condições funcionais da supressão do marcador de plural, feitas antes do trabalho de Scherre (1988), constataram que os falantes cubanos evitam suprimir todos os traços de indicação morfológica de número, preservando, dessa maneira, a primeira forma da estrutura sentencial. Outros trabalhos, como o de Braga (1978), com cópula coletado na região do Triângulo Mineiro, confirmam que o determinante, que geralmente ocupa a primeira posição no SN, é a categoria em que mais é observada a retenção de /s/.

No entanto, essa alegação foi amplamente reformulada por Scherre (1988), mediante a reanálise do mesmo fenômeno. A autora rejeita a idéia de que posição do constituinte seja o fator determinante da marcação de pluralidade no SN, mostrando que posição e categoria gramatical são condições fortemente entrelaçadas para a determinação de marca de número no sintagma nominal. Segundo Scherre

tomar classe por posição ou posição por classe encobre regularidades linguísticas importantes. É necessário, portanto, que se introduza uma nova abordagem analítica que dê conta da relação entre estas duas variáveis. Esta nova abordagem se torna mais imperiosa ao verificarmos (1) que as classes gramaticais não nucleares que ocorrem na primeira posição são todas antepostas ao núcleo do SN, (2) que as da segunda posição são predominantemente antepostas ao núcleo do SN e (3) que as classes não nucleares da terceira posição são predominantemente pospostas ao núcleo do SN. (SCHERRE, 1998, p. 97)

A autora, assim, rejeita a idéia de que posição do constituinte seja o fator determinante da marcação de plural no SN, com base na postulação de um princípio mais geral, baseado na

constatação de que, no SN, marca leva à marca e zero leva a zero, princípio esse que denominou de Paralelismo Formal. Ao postular esse princípio, Scherre (1988) recusa também a atuação das Condições de Distintividade de Kiparsky (1972), considerada por ela como uma hipótese funcionalista no arcabouço da teoria gerativa. De acordo com Camacho

(...) esse rótulo se justifica plenamente porque Kiparsky reconhece a existência de regularidades fonológicas e morfológicas baseadas em condições de natureza funcional, e um desses aspectos regulares diz respeito a regras opcionais baseadas em frequências variáveis (CAMACHO, 2008, p.14)

Kiparsky se interessa por um fenômeno encontrado em algumas variedades do inglês norte-americano que apaga variavelmente o fonema /t/ em final de palavra. Quando esse fenômeno fonológico se apresenta como marca de pretérito, ele é cancelado somente nos casos de verbos irregulares, cujas formas de presente e de pretérito contêm uma vogal diferente na raiz, como no caso de *keep vs. kep(t)*. Pode-se afirmar que a regra de apagamento é bloqueada ou aplicada em menor grau quando sua aplicação neutraliza a distinção entre a forma do presente e a do pretérito, como em *passed*, e é liberada quando a distinção se mantém devido à alternância vocálica na raiz (KIPARSKY, 1972, p.197-198).

Na tentativa de explicar esse fenômeno variável, o autor postula princípios funcionais, denominados de *Condições de Distintividade*, caracterizados como uma tendência capaz de reter uma informação semanticamente relevante na estrutura superficial.

Poplack (1980) foi uma das primeiras pesquisadoras a incorporar esses princípios aos estudos sociolinguísticos. A autora investiga a influência de condições formais e funcionais competindo pela marcação variável de plural no espanhol porto-riquenho. Poplack (1980) reivindica a superioridade das condições estruturais para a explicação desse fenômeno variável. Sua pesquisa revela evidências de uma tendência local, isto é, de uma tendência para a concordância no nível da sequência: a ausência de um marcador no segmento precedente à

ocorrência favorece a supressão neste elemento, enquanto a presença de uma marca imediatamente precedente favorece a retenção de uma marca na ocorrência. Isso mostra que, se a cadeia manifestar a supressão de marcas nos dois primeiros constituintes, é alta a contribuição do fator “posição” para a probabilidade de supressão no terceiro elemento. Se a sequência for realizada com marcas na primeira e na segunda posição, é bem baixo o efeito da supressão no constituinte da terceira posição. O resultado é, enfim, exatamente oposto do que se descobriu em outros estudos, e é caracterizado por Labov (1987) como um argumento contra-funcional.

Hochberg (1986) retoma a hipótese de Kiparsky no estudo da variedade porto-riquenha do espanhol, cujos resultados seguem uma direção oposta aos de Poplack (1980), confirmando, assim, a atuação das condições de distintividade para a explicação do mesmo fenômeno. Entretanto, em estudo parecido ao de Poplack (1980), Guy (1981) não encontra, nos seus dados sobre o português brasileiro, respaldo para a confirmação da hipótese funcionalista, e acaba postulando que as restrições funcionais atuam apenas quando forma e função são perfeitamente paralelas (GUY 1981, p. 345 *apud* SCHERRE 1988, p. 44).

Lefebvre (1981), estudando o Quéchuá de Cuzco, apresenta também evidências de restrições funcionais e não-funcionais em competição, mostrando algumas situações que são previstas pela hipótese funcionalista de Kiparsky, outras que não o são (LEFEBVRE, 1991 *apud* SCHERRE, 1988, p.43).

A principal tendência de todos esses estudos é fornecer razões para suspeitar de que a explicação para a marcação de pluralidade pode estar num processo estritamente linguístico de natureza mecânica, de modo similar aos resultados obtidos por Poplack (1980) no espanhol porto-riquenho, por meio do princípio de que uma marca leva a outra, princípio defendido também por Scherre (1988) para o português brasileiro.

Recentemente, Scherre e Naro (1993), examinando a regra de concordância verbal, demonstram que o princípio do Paralelismo Formal também se aplica ao português, indicando uma tendência de formas similares ocorrerem juntas no nível oracional e no discurso.

O paralelismo formal foi absorvido como princípio teoricamente explanatório por Labov (1987; 1996) como forma de rejeição simultânea de princípios funcionalistas, um movimento que repercutiu largamente na teoria variacionista. Resta, portanto, saber se procedimentos mecânicos sempre sobrepujam os funcionais que requerem marca de pluralidade quando o processo de comunicação exigir eficiência informativa ou se representam motivações competindo entre si, como Lefebvre (1981 *apud* SCHERRE, 1988) sugere para o Quéchuá. No entanto, antes de entrar nessa questão será detalhada a literatura sobre a concordância nominal no predicativo, foco deste trabalho.

1.3.1 Literatura sobre a marcação de pluralidade no predicativo

Os muitos estudos baseados apenas na concordância de número entre os constituintes do SN mostram que os constituintes sintagmáticos flexionáveis podem apresentar marcas formais de plural ou essa marca pode aparecer restrita ao determinante, como se vê em (7 a-b).

- (7) a *mas as casas ali são bem **antigas**... né?... (AC 120/ DE/L. 266)*
 b *as coisa tava praticamente **resolvida** as pessoas... tava sendo **socorrida** né? (AC 103/NE/L.107-109)*

Vale lembrar que, este trabalho centra o foco no estudo da variação na concordância nominal de estruturas predicativas (adjetivos, nomes e participios passivos). Até este momento, tem-se notícia de quatro relevantes trabalhos sobre esse assunto, que abarcam, no entanto, universos distintos: Scherre (1991) relata um trabalho com falantes urbanos da cidade

do Rio de Janeiro; Dias (1996) revela um panorama dos falantes da região sul do Brasil, também residentes em centros urbanos; Lucchesi (2008) mostra a variação de predicativo em informantes de três comunidades afro-brasileiras isoladas no interior do estado da Bahia; e, finalmente, Antonino (2007) apresenta resultados de informantes de dois municípios do interior do estado da Bahia.

É claro que esses estudos não podem ser absolutamente comparados entre si, mas há semelhanças entre os trabalhos de Scherre (1991) e Dias (1996), visto terem por escopo a fala urbana, além do fato de Dias ter tomado por base os resultados de Scherre. Já os trabalhos de Antonino (2007) e Lucchesi (2008) integram o *Projeto Vertentes* e tratam da fala popular e rural de comunidades mais isoladas do interior da Bahia com forte influência de colonização africana.

1.3.2 A concordância nominal nas estruturas predicativas da fala urbana

Scherre (1991) baseia seu trabalho na fala de 64 informantes do projeto Censo da cidade do Rio de Janeiro. A autora classifica seus informantes em quatro faixas etárias (07 a 14 anos, 16 a 25 anos, 26 a 49 anos e 50 a 71 anos), em três níveis de escolaridade (os antigos primário, ginásio e segundo grau) e por gênero (masculino e feminino). Scherre encontrou um total de 837 ocorrências nas amostras de fala, sendo que 785 delas são de predicativo do sujeito, perfazendo um total de 90%, e apenas 88 ocorrências eram participios passivos, somando apenas 10%. A pesquisadora desconsidera 114 ocorrências por diversas razões que poderiam, de algum modo, comprometer a análise dos dados, como os casos em que a tradição gramatical não necessariamente considera como sujeitos à concordância, como no exemplo (8).

(8) *Eles dois fazem uma boa dupla*

Desse modo, foram utilizadas 759 ocorrências, sendo 678 predicativos e 81 estruturas passivas. Em seu trabalho, Scherre mostra uma frequência geral de presença de marcas explícitas de plural equivalente a 46% (410/873) das ocorrências.

Dias (1996), por seu lado, investiga 72 informantes que fazem parte do banco de dados do *Projeto Varsul*, residentes em três cidades da região sul do Brasil: Florianópolis (SC), Chapecó (SC) e Irati (PR). Diferentemente de Scherre (1991), Dias classifica seus informantes em apenas duas faixas etárias (25 a 49 anos e acima de 50 anos). A autora faz a estratificação social também por gênero e escolaridade, além disso, ela acrescenta a variável *grupo étnico*, compreendida pelos fatores açorianos, italianos e eslavos.

Em seu trabalho sobre a região sul, Dias (1996) trabalha com uma amostra de 938 ocorrências de predicativos e participios passivos e, por razões semelhantes às de Scherre (1991), a autora desconsidera 170 ocorrências, trabalhando, no final, com 768 dados. Sua amostra dispõe de 625 ocorrências de predicativo, correspondentes a 81%, e de 143 ocorrências de participios passivos, correspondentes apenas a 19% do total. Dias (1996) detecta a frequência geral de presença de marcas de plural de aproximadamente 43%, resultado muito parecido ao de Scherre (1991) para a fala urbana do Rio de Janeiro.

As variáveis analisadas pelas autoras foram praticamente as mesmas, pois Dias (1996) utiliza o trabalho de Scherre (1991) como referência. As variáveis analisadas pelas autoras estão descritas a seguir:

Variáveis usadas por Scherre (1991)	Variáveis usadas por Dias (1996)
Paralelismo formal	Paralelismo formal
Características formais do sujeito	Características formais do sujeito
Características formais do verbo	Características formais do verbo
Estrutura do predicativo	Estrutura do predicativo

Processos morfofonológicos de formação de plural	Processos morfofonológicos de formação de plural
Tonicidade dos itens singulares	Tonicidade dos itens singulares
-----	Distância do sujeito correferente
Ordem de elementos na estrutura	Ordem dos elementos na estrutura
Presença ou ausência de material interveniente entre o verbo e o predicativo	Presença ou ausência de material interveniente entre o verbo e o predicativo
Tipo de estrutura	Tipo de estrutura
Gênero	Gênero
Idade	Idade
Escolaridade	Escolaridade
-----	Etnia

Quadro 1: Variáveis linguísticas e extralinguísticas utilizadas nos trabalhos de Scherre (1991) e Dias (1996)

Como as variáveis analisadas são praticamente idênticas, os resultados são expostos conjuntamente, tratando-se primeiramente, das variáveis estatisticamente relevantes para ambos os trabalhos e, posteriormente, depois das variáveis não selecionadas pelo VARBRUL.

As três primeiras variáveis selecionadas, a saber: *paralelismo formal das sequências de predicativos/particípios no discurso*, *características formais do sujeito* e *características formais do verbo*, possibilitam comprovar o princípio do Paralelismo Formal, segundo o qual o emprego de determinadas formas gramaticais num discurso tende a influenciar o uso de formas semelhantes dentro do mesmo discurso (SCHERRE, 1989). O princípio do Paralelismo Formal prevê que, quanto maior for a presença de elementos com marcação explícita de plural, maiores serão as possibilidades de elementos subsequentes também apresentarem marcas formais de pluralidade.

A primeira variável selecionada é *paralelismo formal das sequências de predicativos e particípios passivos no discurso*. Com base nessa variável, as autoras procuram controlar se a forma analisada encontra-se precedida ou seguida de formas marcadas ou não-marcadas. Dessa forma, são estabelecidos cinco fatores descritos a seguir:

- (i) Predicativo/particípio em construção isolada;
- (ii) Predicativo/ particípio primeiro de uma série;
- (iii) Predicativo/particípio precedido de predicativo/ particípio com todas as marcas de plural;
- (iv) Predicativo/particípio precedido de predicativo particípio sem marca(s) de plural;
- (v) Casos Mistos.

Ambos os trabalhos convencionam que, no estabelecimento de uma série, predicativos e participios passivos devem referir-se ao mesmo sujeito, ter a mesma forma, caso se refiram a sujeitos diferentes e, além disso, devem ocorrer a uma distância entre si que não ultrapassasse dez estruturas. Formulados esses critérios, os resultados mostram que as formas precedidas por elementos marcados recebem mais marcas de plural (0.70 em Scherre e 0.85 em Dias) e os predicativos precedidos por elementos não-marcados recebem um grau reduzido de marcas de plural (0.19 em Scherre e 0.15 em Dias), resultado esperado com base em estudos sociolinguísticos recentes sobre a marcação de número.

Nota-se, no entanto, uma diferença na atuação do fator *predicativos primeiro de uma série*: no trabalho de Dias (1996), essa categoria é neutra, enquanto no de Scherre, ela leva à marcação formal de plural no predicativo. Com as formas mistas ocorre o inverso, pois se mostram neutras em Scherre (1991) e favorecedoras de marcas explícitas de pluralidade em Dias (1996).

A segunda variável estudada são as *características formais do sujeito*. As autoras analisam as ocorrências com os seguintes fatores:

- (i) Sujeito zero
- (ii) Sujeito explícito com todos os elementos marcados
- (iii) Sujeito sem a marca no(s) último(s) elemento(s)
- (iv) Sujeito com marca semântica de plural

Tanto Scherre (1991) quanto Dias (1996) mostram que, quando o sujeito traz predominantemente todas as marcas de plural, são muito altas as chances de ocorrer marcação explícita de pluralidade no predicativo, fato que corrobora o princípio de que marcas levam a marcas e zeros levam a zeros. O fator *sujeito zero* em Dias contraria o resultado obtido por Scherre (1991), visto que Dias (1996) não vê o favorecimento da presença de marcas no trabalho de Scherre (PR 0.53). A autora justifica esse resultado com a idéia de “recuperação da informação”, alegando, no entanto, que essa explicação necessita de um estudo mais aprofundado para prevalecer como hipótese explanatória. No trabalho de Dias (1996), o *sujeito zero* é neutro (PR 0.49), o que faz a autora lançar mão de outra variável, a distância do sujeito correferente, para explicar esse resultado.

A análise da terceira variável *características formais do verbo*, pelas autoras, se baseiam nos seguintes fatores:

- (i) Sentenças com zero verbal;
- (ii) Verbo com marca de plural;
- (iii) Verbo sem marca de plural.

Conforme o previsto, a atuação dessa variável nos trabalhos de Scherre (1991) e Dias (1996) reforça o princípio do Paralelismo Formal, pois a frequência de verbos marcados favorece a frequência de concordância nominal no predicativo, enquanto os verbos não-marcados inibem marcação explícita de plural. No âmbito dessa variável, o fator “sentenças com zero verbal” se comporta como o sujeito zero, favorecendo em ambos os trabalhos a marcação de plural.

A próxima variável selecionada *estruturas do predicativo* já não trata mais da questão do paralelismo formal, os resultados obtidos mostram a influência de saliência. A variável foi subdividida em quatro fatores:

- (i) predicativo adjetivo e formas participiais de um elemento
- (ii) predicativo nominal de um elemento
- (iii) predicativos de mais de um elemento
- (iv) predicativos com mais de um elemento com os itens “todo/tudo”

É possível observar que, em ambos os trabalhos, o predicativo desfavorece a aplicação da regra de concordância (PR 0.44 em Scherre e PR 0.48 em Dias). Scherre justifica esse comportamento com base no fato de que a configuração adjetival é formada, em sua maior parte, por palavras paroxítonas, considerado o critério de saliência fônica mais comum para o PB. Além disso, as palavras paroxítonas perdem a marca formal de plural com mais facilidade, por essa razão esse critério é menos saliente que os demais.

Nota-se também que os *predicativos com “tudo/todo”*, no trabalho de Dias (1996), tendem a desfavorecer a marcação de plural (PR 0.48). Quando se trata, contudo, de *predicativos com mais de um elemento*, ocorre justamente um resultado inesperado de favorecimento na aplicação da regra (PR 0.61). Esse dado pode encontrar explicação na influência do contexto antecedente, entendidas as características do verbo ou do sujeito. Alega, Dias, no entanto, que, para essa afirmação, seria necessária uma análise mais detalhada das ocorrências.

A variável *processos morfofonológicos de formação de plural* revela uma diferença relevante entre os dois trabalhos. Apesar de resultados, em geral, semelhantes, essa variável é selecionada no trabalho de Scherre (1991), mas não na pesquisa de Dias para a região sul. Essa variável e a variável rotulada *tonicidade dos itens regulares* tentam comprovar o princípio da saliência fônica⁶ (SCHERRE, 1988), em que os itens mais salientes, isto é, com

⁶ Esse princípio consiste em estabelecer que as formas mais salientes, e por isso mais perceptíveis, são mais prováveis de serem marcadas do que as menos salientes. A saliência fônica é focalizada particularmente nos estudos da concordância verbal, especialmente na dimensão que envolve a quantidade de material fônico na relação singular/plural. Naro (1981) propõe que se estabeleçam duas dimensões no eixo da saliência fônica:

maior diferenciação de material fônico na relação singular/plural, favorecem a presença de marcas formais de pluralidade. Desse modo, quanto maior a saliência fônica dos itens que compõem as estruturas predicativas, maior o favorecimento da presença de marcas explícitas de plural.

Em relação, ainda, à variável *processos morfofonológicos de formação de plural*, Scherre divide a variável em três fatores: plural regular, plural irregular e plural misto. Dias (1996), por sua vez, faz a diferenciação do plural irregular, como os plurais irregulares terminados em <l, r, s, ão> e o plural duplo. Neste trabalho, seguiremos a divisão de Dias (1996).

Já na variável *tonicidade dos itens regulares* se estabelece uma divisão em três fatores tanto no trabalho de Scherre (1991) quanto no de Dias (1996), a saber: itens oxítonos, itens paroxítonos e itens proparoxítonos. Ambas as pesquisas mostram que os predicativos com itens oxítonos, considerados mais salientes fonicamente, apresentam mais marcas explícitas de plural que os com itens paroxítonos e proparoxítonos. Vale destacar, no entanto, que somente no trabalho de Dias (1996) essa variável foi selecionada como estatisticamente relevante.

As outras variáveis descritas a seguir não foram consideradas estatisticamente relevantes pelo pacote VARBRUL. Mesmo assim, contribuem substancialmente para os resultados das pesquisas. A primeira delas é a variável *ordem dos elementos na estrutura* que foi dividida em dois fatores:

- (i) Ordem canônica: sujeito + verbo + predicativo
- (ii) Ordem não-canônica

É possível constatar que a variável ordem canônica tende a desfavorecer a presença de marcas no trabalho de Scherre (1991) em que o peso relativo está abaixo de 0.50. Já o trabalho de Dias (1996) acaba por confirmar a neutralidade desse fator (PR 0.50).

Outra variável não selecionada pelo VARBRUL, nos dois trabalhos, é *presença ou ausência de material interveniente entre verbo e predicativo/passiva*. Essa variável abrange três fatores: (i) ausência de material interveniente, (ii) presença de intensificadores e (iii) demais materiais intervenientes. A variável em destaque é pouco produtiva na pesquisa de Scherre (1991), já que não provoca nem aumento nem diminuição do uso de marcas formais de pluralidade nos predicativos. Em contrapartida, Dias (1996) relata haver um ligeiro desfavorecimento da regra com a presença de material interveniente não intensificador.

A última variável linguística analisada é o *tipo de estrutura*. A intenção das autoras é verificar a existência de diferença na marcação de pluralidade no predicativo em estruturas ativas e estruturas passivas. A pesquisa sobre a fala urbana do Rio de Janeiro mostra que as estruturas passivas desfavorecem a aplicação da regra de CN nos predicativos, embora o peso relativo (0.44) seja considerado pouco expressivo por Scherre (1991). No trabalho de Dias (1996), a estrutura ativa também se revela neutra para a marcação de pluralidade no predicativo (PR 0.48).

O enfoque das variáveis sociais revela um quadro estável de variação. Quanto ao gênero, as mulheres favorecem mais o uso das formas de prestígio; quanto à escolaridade, o nível é diretamente proporcional ao uso das marcas formais de pluralidade. Sobre a faixa etária, os resultados se revelaram curvilíneos no trabalho de Scherre, já que jovens e velhos desfavorecem o uso das formas padrão e os de idade intermediária favorecem o uso de marcas no predicativo. Como já descrito acima, Dias (1996) analisa apenas duas faixas etárias, observando, no final, que os jovens desfavorecem o uso da variante padrão, com peso relativo

de 0.41, enquanto as pessoas mais velhas tendem a favorecer a presença de marcas de plural nos predicativos/passivas (PR 0.59).

De acordo com as características da comunidade de fala do sul do país, Dias (1996) acrescenta uma quarta variável social denominada, por ela, de *etnia* para considerar a influência linguística de açorianos, italianos e eslavos. Dias observa que a regra padrão de concordância nominal é discretamente favorável no grupo étnico italiano (0.57), já os eslavos se encontram numa posição intermediária, com peso relativo (0.48) e os açorianos são os que menos aplicam a regra nos predicativos (0.44). A rigor, como se vê, o grupo *etnia* não se mostra uma variável significativa.

Antes de mencionar o estudo sobre o predicativo na fala do interior da Bahia, pretende-se expor os resultados do trabalho de Salomão (2007) sobre as *motivações formais da marcação de pluralidade nas estruturas predicativas da variedade falada na região de São José do Rio Preto*⁷. Este trabalho de iniciação científica é fruto do Projeto *Motivações Formais e Funcionais da Marcação de Pluralidade do Português Falado na região de São José do Rio Preto*⁸.

O estudo piloto de Salomão (2007) tem por objetivo submeter a um tratamento variacionista, de base quantitativa, dados de marcação variável de plural em contexto predicativo. A proposta deste trabalho é examinar as motivações, linguísticas e extralinguísticas, em competição na indicação de pluralidade das estruturas predicativas na variedade falada em São José do Rio Preto. A principal hipótese teórica deste trabalho é a de confirmar se a ausência de marcas de pluralidade em fenômenos de concordância nominal tinha por explicação metodológica a repetição de estruturas em paralelismo formal.

⁷ Projeto de iniciação científica orientado pelo Prof. Dr. Roberto Gomes Camacho. Este trabalho foi financiado pela FAPESP, processo n° 06/55311-7

⁸ Projeto financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo: n° 2006/53339-1

O universo de pesquisa é composto por ocorrências obtidas em 95 narrativas de experiência das Amostras Censo extraídas do Banco de Dados Iboruna⁹, que também formam parte das amostras para esta dissertação. Esse trabalho restringiu seu escopo às diferenças de gênero, diferenças de escolaridade (2º Ciclo do Ensino Fundamental, Ensino Médio e Ensino Superior) e diferenças de faixa etária (de 16 a 25 anos, de 26 a 35 anos, de 36 a 55 anos e mais de 55 anos).

Como a marcação de pluralidade é uma variável fortemente marcada por pressões normativas, visto que a escola privilegia a variedade padrão da língua, os fatores sociais são capazes de condicionar a presença ou ausência de pluralidade. Com isso, algumas comunidades linguísticas seguem a variedade padrão da língua, devido ao maior contato com ela e outras utilizam a variedade não-padrão, pois a presença de todas as marcas nem sempre é necessária. Dessa forma, foram avaliadas as variáveis independentes de natureza externa com as de natureza interna, para verificar, por um lado, se os diferentes agrupamentos sociais na comunidade riopretense são detentoras de diferentes gramáticas na aplicação da concordância ou, se, por outro lado, a gramática é sempre a mesma, mas cada variante tem diferente frequência de uso conforme a distribuição social dos informantes investigados.

Este estudo-piloto analisou a atuação de oito fatores condicionadores do processo de marcação variável de pluralidade no predicativo, cinco deles linguístico e três extralinguísticos. Os grupos de fatores linguísticos são: (i) paralelismo formal; (ii) paralelismo oracional: marcas do sujeito; (iii) paralelismo sintagmático: marcas do verbo; (iv) configuração estrutural do predicativo/particípio e (v) saliência fônica da oposição singular x plural.

Os grupos de fatores extralinguísticos, por sua vez, são: (i) gênero, (ii) idade e (iii) escolaridade. A partir do processamento estatístico, o VARBRUL indicou os fatores de

⁹ O Banco de Dados Iboruna, fruto do Projeto ALIP, está criteriosamente descrito no capítulo *Procedimentos Metodológicos*.

análise estatisticamente mais significativos, são eles: *escolaridade* e *gênero*, dentre os extralinguísticos; e *paralelismo oracional: marcas do sujeito* e *paralelismo sintagmático: marcas do verbo*, dentre os linguísticos.

O cruzamento dos fatores linguísticos com os extralinguísticos permitiu observar que a escolaridade condiciona a presença de marcas formais de pluralidade no predicativo; à medida que aumenta o nível de instrução formal, aumenta também o uso de marcas; já as mulheres são mais sensíveis às formas de prestígio do que os homens. Considera-se, dessa forma, que os fatores sociais exercem influência significativa no condicionamento de ausência e presença de marcas formais de pluralidade.

O estudo sobre o *cópus Iboruna* manifesta muita concordância com as descobertas de Scherre (1991), realizadas com base numa variedade diferente. É possível notar que a maioria dos índices estatisticamente significativos está muito próxima dos resultados a que essa autora chegou, como se as hipóteses que os resultados dela sugerem se confirmassem nos dados do *Iboruna*. Já os resultados que mostram discordâncias têm a importância científica de concluir que comunidades fisicamente afastadas nem sempre têm um comportamento sociolinguístico idêntico e é justamente essa heterogeneidade dialetal que tem fascinado a pesquisa sociolinguística.

No entanto, o princípio do Paralelismo Formal parece não ser a única motivação para a aplicação ou não aplicação da pluralidade dentro do *cópus Iboruna*. Dessa maneira, não deixa de ser passível de alguma suspeição que alguma outra motivação não-interna possa motivar essa tendência em sentido contrário de marcação de plural em relação à presença e à ausência de marcas. Esse resultado, motivado por esse estudo piloto, fornece, portanto, subsídios relevantes para a hipótese aqui investigada de motivações potenciais em competição.

1.3.3 A concordância nominal das estruturas predicativas e passivas em comunidades afro-brasileiras

As duas pesquisas descritas a seguir fazem parte do *Projeto Vertentes*, coordenado por Dante Lucchesi da UFBA. A primeira delas, publicada em Lucchesi (2008), foi realizada com base em falas vernáculas de moradores de três comunidades afro-brasileiras isoladas no interior da Bahia: Helvécia, Cinzento e Bananal. A segunda (ANTONINO, 2007) investiga as pesquisas sociolinguísticas, nos predicativos e nas construções passivas, a partir da análise da variedade falada em dois municípios do interior do estado da Bahia: Santo Antônio de Jesus e Poções.

Lucchesi (2008) analisa 36 entrevistas, com 12 informantes de cada localidade, categorizadas de acordo com as mesmas variáveis sociais usadas por Scherre (1991) e Dias (1996), isto é, *idade, escolaridade e gênero*. Além disso, o autor se baseia na variável *tempo de estadia fora da comunidade*. Na variável idade, há uma divisão de três faixas etárias (20 a 40 anos, 41 a 60 anos e mais de 60 anos); em relação ao nível de escolaridade, os informantes são classificados como analfabetos ou semi-analfabetos, pois todos fazem parte da base da pirâmide social.

Antonino (2007) faz uma divisão muito parecida com a de Lucchesi. Seu *córpus* está composto pela fala vernácula de 48 informantes das cidades de Santo Antônio de Jesus e Poções, em proporções iguais, ou seja, 24 informantes para cada localidade. Além disso, o *córpus* está dividido por gênero (masculino e feminino) e por três faixas etárias: (i) de 25 a 35 anos, (ii) de 45 a 55 anos e (iii) acima de 65 anos. Assim como Lucchesi (2008), Antonino controlou *o tempo de estadia dos informantes fora da comunidade*. Segundo a autora,

A escolha por esses informantes se deu por acreditar que estes, com pouco ou nenhum acesso à ação normatizadora da instituição escolar, mantêm mais evidentes marcas de um passado em que a situação de contato entre

línguas exerceu influência em sua realidade linguística. (ANTONINO, 2007, p. 41)

Essa afirmação tem relevância porque, diferentemente dos trabalhos sobre predicativo na fala urbana, as pesquisas de Antonino e Lucchesi procuram investigar o processo de transmissão irregular, ou seja, as situações de contato do português com outras línguas, principalmente as de origem africana. De acordo com esse princípio, o contato entre as línguas provoca a formação de uma nova língua, o *pidgin* ou língua crioula, com sintaxe menos desenvolvida, própria para a comunicação. Esses estudos são importantes para a discussão da alegada origem crioula do português brasileiro, mas seus resultados têm pouco a dizer para o tratamento da variedade riopretense que não se envolve diretamente com a questão do contato linguístico.

Lucchesi (2008) mostra que a variação de CN nos predicativos e em estruturas de voz passiva na fala de comunidades isoladas da Bahia é um fenômeno muito acentuado, pois há apenas 1 % de aplicação da regra de CN. De acordo com o autor, pode-se afirmar que a regra de concordância nominal nos predicativos está praticamente ausente nas gramáticas destas comunidades. O autor lembra que a concordância de número dentro do sintagma nominal, no trabalho de Scherre (1988), é de 72 %, ao lado de uma marcação de 50% em estruturas passivas e de predicativo. Segundo Lucchesi (2008), essa diferença está diretamente correlacionada a uma maior coesão estrutural dentro do SN. Lucchesi (2008) acredita haver uma polarização sociolinguística no português brasileiro, pois essa significativa diferença de frequência de uso de aplicação da regra de CN nos predicativos está relacionada a diferenças dialetais motivadas por contato maior ou menor com as línguas africanas na formação do português.

O estudo de Antonino (2007) também registra um baixo índice de aplicação da regra. Em um total de 307 ocorrências, apenas 13 delas recebem marcas de plural, totalizando um

índice de apenas 4% de marcação. Como se pode notar, a produção de sentenças é mais elevada na fala popular urbana, mas também há uma maior utilização das marcas explícitas de plural nas estruturas predicativas. Antonino (2007), assim como Lucchesi (2008), também acredita numa realidade linguística polarizada.

Os estudos que relatam as comunidades do interior da Bahia são extremamente importantes para a composição de um panorama geral da realidade sociolinguística do português do Brasil e é em função dessa relevância dialetológica que eles têm importância comparativa com este trabalho. Por outro lado, o estudo proposto aqui trata de uma comunidade de fala urbana, que também sofre influências de outras variedades como o espanhol, o italiano e o árabe; apesar disso, o foco deste trabalho não se centra nos estudos de transmissão linguística e, por essa razão, a base de comparação serão os trabalhos de Scherre (1991) e Dias (1996).

1.4 O problema da norma e o preconceito linguístico

A linguística brasileira tem dedicado um espaço privilegiado para a discussão do conceito de norma linguística. A expressão traz consigo várias possibilidades de interpretações, gerando uma confusão no imaginário dos falantes. Além disso, pode ocasionar o preconceito linguístico, pois as pessoas tendem a acreditar que a norma mais monitorada seja a própria língua e tudo o que se desvia daquele conjunto de regras esteja errado. Para essa discussão, mencionaremos três interpretações de conceito de norma: norma culta, norma-padrão e norma gramatical.

A expressão norma culta/comum/*standard* designa o conjunto de fenômenos linguísticos que ocorre habitualmente no uso dos falantes alfabetizados em situações monitoradas de fala e de escrita. De acordo com Faraco (2008, p.73) “esse vínculo com os usos monitorados e com as práticas da

cultura escrita leva os falantes a lhe atribuir um valor social positivo, a recobri-la com uma capa de prestígio social.”

Por esse motivo, a norma culta tornou-se objeto privilegiado de registro e estudo sociocultural. Esse processo acaba causando, no imaginário dos falantes, uma representação de norma como se ela fosse uma variedade superior do português, isto é, uma variedade melhor do que todas as outras. Esse fato leva os falantes a confundir o conceito de norma com a própria língua, pois eles acreditam que as outras variedades existentes dentro do PB são degradação da “língua verdadeira.”

No entanto, essa representação imaginária dos falantes não encontra sustentação na realidade, já que as mudanças não alteram a plenitude formal de nenhuma das variedades linguísticas, ou melhor, as variedades passam por reconfigurações, mas nunca perdem seu caráter estruturado e sua capacidade ilimitada de expressão. Além disso, a língua é sempre heterogênea, isto é, constituída por um conjunto de variedades, por um conjunto de normas, fato que inviabiliza o conceito de que as variedades não cultas sejam deturpações. De acordo com Faraco,

a norma dita culta é apenas uma dessas variedades, com funções socioculturais bem específicas. Seu prestígio não decorre de suas propriedades gramaticais, mas de processos sócio-históricos que agregam valores a ela. Em outras palavras, seu prestígio não decorre de suas propriedades intrínsecas (linguísticas propriamente), mas de propriedades extrínsecas (sócio-históricas). (FARACO, 2008, p.74)

Dito isso, as variedades se equivalem do ponto de vista gramatical, pois todas são igualmente organizadas e complexas. No entanto, essa equivalência formal não implica que todas se equivalham socialmente, já que a diferenciação existente acaba por hierarquizá-las de modo tal que algumas recebem avaliação positiva, enquanto outras são desprestigiadas e até mesmo estigmatizadas. Essa valoração não é linguística, mas resulta do modo como se constituem as relações entre os grupos sociais através da história.

A norma-padrão, diferentemente da norma culta, não é uma variedade de língua, mas “um construto sócio-histórico que serve de referência para estimular o processo de uniformização”

(FARACO, 2008, p. 75), a norma-padrão é abstrata e é utilizada como uma baliza extraída do uso real para servir de referência, em sociedades marcadas por uma acentuada dialeção.

De acordo com Faraco (2008), na Europa, a variedade de língua que foi tomada como referência para a construção da norma-padrão difere de um país para o outro. Em alguns lugares a norma-padrão tem uma perspectiva mais conservadora, em outros, mais pragmática

porém, não deixou de estar próxima da norma culta/comum/standard, ou seja, da variedade praticada à época pela aristocracia ou, mais propriamente (considerando que a questão da norma-padrão foi, antes de tudo, um trabalho dos homens letrados), da variedade praticada pelos “barões doutos” - na feliz expressão do erudito português do século XVI, João de Barros, autor, entre outras obras, de uma das primeiras gramáticas do português. (FARACO, 2008, p. 76)

Nesse período, a gramática, por ter adquirido força coercitiva, não foi considerada apenas como um instrumento descrito, como um registro da norma, mas ela se tornou instrumento normatizador, responsável pela fixação de um padrão de comportamento linguístico.

Em decorrência disso, a palavra *norma* tem, em seu uso contemporâneo, dois sentidos: normalidade e normatividade. Nos estudos linguísticos, o conceito de norma está relacionado ao conjunto de fenômenos linguísticos habituais (normais) numa comunidade de fala. Já no funcionamento monitorado da língua, a palavra *norma* tem caráter regulador, isto é, é utilizada para guiar os comportamentos dos falantes em determinadas situações (BAGNO, 2003).

Há uma dificuldade grande, atualmente, em se discutir a questão da norma-padrão entre linguistas, gramáticos, jornalistas e autoridades públicas, já que boa parte dos estudiosos que se envolvem com o tema tem uma visão reducionista da questão, isto é, a norma-padrão acaba se transformando em “um rol congelado de formas ditas ‘corretas’.” (FARACO, 2008, p. 78) Além disso, a função é sempre recriminatória, visto servir para julgar os falantes que não utilizam a norma-culta, como se eles não soubessem falar e escrever o português corretamente, o que acaba gerando o preconceito linguístico. Essa recriminação sobre o uso da língua não é difícil de ser feita, pois,

geralmente, nem mesmo as pessoas que se dizem cultas, estudadas, usam as formas cultuadas pela norma-padrão. De acordo com Faraco

Qualquer debate hoje costuma logo ser abortado por recorrentes acusações da e na mídia de que os linguistas são (perigosamente) relativistas e, portanto, contrários ao ensino de um padrão de língua. (FARACO, 2008, p. 79)

Os linguistas, entretanto, estão combatendo o caráter artificial de padrão, pois é essa concepção que se instala como uma camisa-de-força e causa diversos preconceitos linguísticos. Segundo Faraco, essas questões é que devem constituir o ponto de partida de qualquer debate e não a acusação de relativismo.

No caso brasileiro, diferentemente do que ocorreu na Europa, a norma-padrão não teve como objetivo central responder a uma situação de dialeção profunda. Além disso, o padrão linguístico nunca esteve vinculado à construção de um estado unificado. Nas palavras de Faraco,

O projeto da norma-padrão no Brasil teve como objetivo principal (...) combater as variedades do português popular. Se no século XVIII, com o *Directório dos Índios*, se buscou implantar uma política que visava calar as línguas indígenas, em especial a chamada língua geral, no século XIX, a intenção era calar as variedades rurais e (progressivamente) rurbanas. Nesse afã, os formuladores e defensores da norma-padrão se opuseram com igual furor às características das variedades populares cultas faladas aqui. O excessivo artificialismo do padrão que estipularam impediu, porém, que ele se estabelecesse entre nós. (FARACO, 2008, p.82)

Já a norma gramatical representou uma flexibilização dos juízos normativos, fato que rompeu, pelo menos em parte, a rigidez da tradição conservadora. A norma gramatical é o conjunto de fenômenos apresentados como cultos pelos gramáticos. As gramáticas contemporâneas já acolhem vários fenômenos da norma culta/comum, porém, elas ainda não possuem muitas descrições sistemáticas dessa norma. Por essa razão, podemos afirmar que as gramáticas atuais estão num meio termo, pois veiculam excessos da norma-padrão e poucas descrições da norma comum.

Apesar disso, a norma gramatical é vista como limitadora de usos, pois há uma expectativa de que os usos (a norma culta/comum/*standard*) se conformem ao que está estipulado na norma gramatical. Por essa razão, o uso culto deveria estar subordinado aos instrumentos normativos.

Segundo Faraco

(...) não é fácil, no plano do imaginário social, resolver esta dupla subordinação. Objetivamente, a língua precede sempre os instrumentos e estes devem, portanto, estar sempre em consonância com ela. Culturalmente, porém, dá-se um valor quase sagrado aos instrumentos normativos, como se eles precedessem a língua e garantissem sua preservação. (FARACO, 2008, p. 88)

Essa condição não é problema se a língua fosse homogênea e estática, mas a realidade é outra, já que a língua é heterogênea e está em constante mutação. Como se vê, nem mesmo a difusão institucionalizada da norma-padrão é capaz de homogeneizar normas sociais e de parar as mudanças linguísticas.

Essa rápida discussão leva a refletir sobre as relações entre a chamada regra de concordância nominal, tomada como obrigatória na variedade normativa, e a variável presença/ausência de marca de pluralidade, que de fato ocorre na fala quotidiana. Como a heterogeneidade constitutiva da linguagem não é levada em conta pela tradição normativa, a idéia generalizada é a de que a regra de concordância nominal é obrigatória não apenas no registro culto formal da variedade escrita, mas em todos os registros falados também, mesmo nos mais informais.

Essa concepção gera o preconceito de que a ausência de marcas é estigmatizada socialmente, fato que discrimina socialmente os falantes que só reconhecem a regra da variedade não-padrão. A consequência mais imediata dessa discriminação é a criação de um abismo entre falantes com grau superior de escolaridade e falantes com grau mais baixo de educação formal. Essa relação será enfocada com especial atenção neste trabalho na discussão da variável escolaridade.

Após ampla explanação sobre os pressupostos teóricos deste trabalho, pautados na Teoria da Variação e Mudança Linguística, nos Princípios Funcionalistas de Economia (HAIMAN, 1983), nas

Condições de Distintividade (KIPARSKY, 1972), nas Motivações em Competição (DU BOIS, 1985) e no Princípio do Paralelismo Formal (LABOV, 1996), realizou-se toda uma discussão de como a Sociolinguística Variacionista se relacionou com as teorias formais e funcionais desde seu início até a quebra de seu estatuto teórico, ou seja, a radicalização da posição de Labov (1987), na extensão da variável sociolinguística para os estudos da sintaxe.

Além disso, foi possível discutir, nessa seção, os trabalhos sobre a marcação de pluralidade dentro do SN e a literatura sobre a concordância nominal nos predicativos do português brasileiro, em que foram mostrados quatro trabalhos relevantes, sendo dois relacionados à fala urbana e dois à fala rural. No encerramento da seção I, houve, ainda, uma discussão sobre norma e as causas que geram o preconceito linguístico.

Na próxima seção, trataremos mais especificamente sobre a região¹⁰ de São José do Rio Preto, revelando as suas características sócio-históricas. Além disso, serão apresentados todos os procedimentos metodológicos utilizados neste trabalho e o modo de quantificação dos dados.

¹⁰ Região será tratada, neste trabalho, como cidades circunvizinhas a São José do Rio Preto.

2. UNIVERSO DE PESQUISA E PROCEDIMENTOS METODOLÓGICOS

O presente estudo segue os preceitos metodológicos da sociolinguística variacionista propostos por Labov (1972). Os informantes que compõem o *cópus* da pesquisa pertencem a um grupo social em cuja fala a regra de concordância no predicativo representa um fenômeno variável.

Esta seção se organiza do seguinte modo: a primeira parte apresenta a caracterização histórico-social da região de São José do Rio Preto; a segunda traz um breve histórico do Projeto ALIP; a terceira parte mostra os procedimentos de constituição do Banco de Dados Iboruna; a quarta parte, a composição da subamostra, utilizada como *cópus*; a quinta parte define os contextos variáveis (envelope variacional); a sexta parte descreve os fatores de análise e revela as hipóteses relacionadas a cada variável, e a sétima mostra como foi feita a quantificação dos dados.

2.1 Caracterização histórico-social da região¹¹ de São José do Rio Preto

Nesta subseção, caracterizaremos sócio-historicamente as cidades que fazem parte do Banco de Dados Iboruna, *cópus* utilizado neste trabalho, a saber: São José do Rio Preto, Mirassol, Bady Bassit, Guapiaçu, Cedral e Onda Verde. Todos os municípios são circunvizinhos a São José do Rio Preto, que é a cidade mais importante entre as citadas e que também leva o nome da região administrativa.

A cidade de São José do Rio Preto foi fundada no ano de 1852, por famílias de mineiros que aqui se fixaram e deram início à exploração e à criação de animais domésticos. A aglomeração urbana aconteceu, principalmente, devido à doação de terras dessa região à igreja católica. Particularmente, Luiz Antonio da Silveira doou parte de seu patrimônio a seu santo protetor, São José. Em 19 de julho de 1894, por meio da lei nº 294, São José do Rio

¹¹ Considera-se região, neste trabalho, as cidades que fazem parte do Banco de Dados Iboruna e que são circunvizinhas a São José do Rio Preto. Trata-se de uma porção da região do Noroeste Paulista do Estado de São Paulo.

Preto foi desmembrada da cidade de Jaboticabal e tornou-se um município. A extensão geográfica, com mais de 26 mil Km², era imensa, limitada pelos rios Paraná, Grande, Tietê e Turvo.

No ano de 1904 é criada a comarca de São José do Rio Preto e, a partir de 1906, o nome da cidade se reduz para Rio Preto, para voltar a retomar, em 1945, o nome original de São José do Rio Preto. Já em 1912, com a chegada da Estrada de Ferro Araraquarense (EFA), a cidade assumiu uma importante posição de polo comercial de concentração de mercadorias produzidas no então conhecido “Sertão de Avanhandava” e de irradiação de materiais vindos da capital.

Atualmente, São José do Rio Preto possui uma área de 434,10 Km², compreendida entre as fronteiras dos municípios de Ipiguá e Onda Verde ao norte, Bady Bassitt e Cedral ao sul, Guapiaçu a leste e Mirassol a oeste.

A cidade dista 450 Km da capital do estado pela rodovia Washington Luís (SP-310) e 700 Km de Brasília, pela Rodovia Transbrasiliana (BR-153). São José do Rio Preto é predominantemente urbana, considerando que apenas cerca de 5% de sua população ocupa a zona rural.

Tem uma população de 414.272 habitantes¹², que somada aos habitantes das cidades que compõem a região administrativa do município, composta por 96 municípios, perfaz cerca de 1 milhão e 400 mil habitantes. A região administrativa de São José do Rio Preto se localiza no noroeste do Estado de São Paulo, fazendo divisa com os estados de Minas Gerais e Mato Grosso do Sul, conforme se observa nos mapas 1e 2:

¹² Todas as informações colocadas aqui, a respeito da população das cidades, foram retiradas do Censo de 2000.



Figura 1: Mapa da região administrativa de São José do Rio Preto



Figura 2: Localização de São José do Rio Preto na região Noroeste de São Paulo

A cidade vive um dos momentos mais importantes de sua história: é a 59ª maior economia do Brasil e a 18ª do Estado de São Paulo. Está à frente de várias capitais e municípios maiores. O PIB cresceu 34,2%, totalizando R\$5,2 bilhões¹³.

¹³ Fonte: ftp://ftp.ibge.gov.br/Pib_Municipios/2005/Banco_de_Dados/

São José do Rio Preto vem se constituindo em uma das cidades com melhores condições de vida. Apresenta um índice de desenvolvimento Humano (IDH) que se compara ao dos países desenvolvidos em razão dos bons indicadores de saúde, educação e renda. A qualidade de vida atrai perto de 8.000 novos moradores a cada ano de todas as faixas de renda, advindos de todas as regiões do Brasil.

A cidade é hoje um centro de produção e consumo de bens e serviços de natureza econômica, social e cultural, atingindo uma área de atração de mais de 2.000.000 de consumidores. Essa influência econômica gerou um grande crescimento nas áreas de comércio e serviços especializados de alta tecnologia, medicina avançada e educação.

No setor industrial, o número de micro e pequenas empresas é bem significativo. A cidade conta atualmente com treze minidistritos com mais de mil empresas beneficiadas e três distritos industriais. Dentre as principais atividades das indústrias destacam-se a moveleira, a alimentícia, a de jóias e folhados e os produtos médicos. Além disso, o município conta com alguns centros de referência em educação como a Universidade Estadual Paulista (UNESP), a Faculdade de Medicina (FAMERP) e a Faculdade de Tecnologia (FATEC).

Já a cidade de Mirassol foi fundada em 8 de setembro de 1910, pelo sertanista Joaquim da Costa Penha, mais conhecido como Capitão da Neves, juntamente com o coronel Victor Cândido de Souza, sob a denominação de São Pedro da Mata Una. Mais tarde, a cidade passou a ser chamada de Mirassol, devido à grande quantidade de girassóis encontrada nos arredores da cidade.

Mirassol, situada a oeste de São José do Rio Preto, possui 48.312 habitantes, sendo 23.8 mil homens e 24.4 mil mulheres, de acordo com os dados do IBGE (2000). Trata-se de uma cidade urbana, pois apenas 10% dos habitantes são residentes da zona rural.

O município se caracteriza pela economia moveleira, com cerca de 300 indústrias. Sua posição de polo no setor atrai feiras de grande porte, como a Movinter (Feira de móveis e

fornecedores do Estado de São Paulo) e a TechMóvel (feira da tecnologia e fornecedores da Indústria Moveleira). Ambas movimentam milhões em investimentos, inclusive em outros setores, como o hoteleiro e de prestação de serviços. Mesmo assim, muitos moradores da cidade preferem trabalhar e estudar em São José do Rio Preto, já que a proximidade e o acesso rápido auxiliam o deslocamento das pessoas interessadas por melhores condições de trabalho.

A cidade de Bady Bassit, que fica ao sul de São José do Rio Preto, é cortada pela maior rodovia do país, a BR-153 (Transbrasiliana), que a liga ao polo regional de São José do Rio Preto, um dos maiores do interior de São Paulo. Sua população é estimada em 14.543 habitantes, e a população é predominantemente urbana, pois mais de 11.000 habitantes (cerca de 75%) vivem na cidade. Bady Bassit, assim como Mirassol, possui característica de cidade dormitório, já que grande parte de sua população trabalha e estuda em São José do Rio Preto. A cidade vive basicamente da agropecuária e do comércio.

A cidade de Guapiaçu é servida pela Rodovia Estadual “Assis Chateaubriand” que a liga a São José do Rio Preto (15 km) e Olímpia (35 km). O município conta com 14.086 habitantes. Trata-se também de uma cidade urbana, pois 11.882 habitantes vivem na cidade e apenas 2.204 (18%) vivem no campo. A atividade principal do município é a agropecuária.

Onda Verde e Cedral são as duas menores cidades que fazem parte do Banco de Dados Iboruna. A população de Cedral estima-se em 6.700 habitantes, sendo que 4.980 habitantes residem em área urbana e apenas 1.720 (25%) são habitantes da zona rural. Já Onda Verde tem a sua população estimada em 3.413 habitantes. Como a cidade é menor a distribuição dos habitantes parece mais equilibrada, 2.319 moradores se encontram na cidade e 1.094 (47%) residem na zona rural. Ambos os municípios se destacam pela agropecuária e, em Cedral, nota-se também a produção de cana-de-açúcar.

Fato interessante dos municípios que são circunvizinhos a São José do Rio Preto é que eles dispõem de boa infra-estrutura e índices altos de alfabetização. Ainda assim, comportam-se como cidades dormitórios. Nesses municípios, a taxa de população de meia-idade é bem maior que a taxa de jovens, pois estes vão procurar centros maiores para trabalhar e terminar os estudos, já que Bady Bassit, Cedral, Onda Verde e Guapiaçu não dispõem de instituições de ensino superior.

2.2 O Projeto ALIP – Breve histórico

Como o corpus é composto pelas amostras do Banco de Dados Iboruna, torna-se relevante descrever os procedimentos metodológicos adotados para a sua composição (GONÇALVES, 2008).

Segundo Tenani e Gonçalves (2005), o Banco de Dados Iboruna¹⁴ foi composto pelo Projeto ALIP (Amostra Linguística do Interior Paulista), no período de março/2004 até setembro/2007, em resposta à demanda do Grupo de Pesquisa em Gramática Funcional (GPGF), sediado na UNESP de São José do Rio Preto, que decidiu debruçar-se sobre o estudo da variedade riopretense. A criação do Projeto ALIP foi motivada pela caracterização e consolidação do Grupo de Pesquisa em Gramática Funcional. O compromisso teórico com um investimento empírico requeria que os pesquisadores usassem amostras de falas já sistematizadas, definidas a partir de objetivos pré-estabelecidos, de acordo com o interesse dos grupos que a detinham, como o NURC. A diversidade de critério de levantamento representava um obstáculo a ser superado dentro do grupo, em razão de que muitos de seus

¹⁴ Todas as amostras de fala estão disponíveis em meio eletrônico, na sede do Projeto ALIP, na UNESP de São José do Rio Preto. Encontra-se também à disposição, gratuitamente, no site do projeto (<http://www.iboruna.ibilce.unesp.br>), parte do material integrante do Banco de Dados Iboruna, inclusive o material sonoro.

integrantes procuravam constituir amostra própria que atendesse os interesses mais específicos de pesquisa.

Outro motivo para a proposta do Projeto ALIP refere-se à validade sincrônica, um tanto ultrapassada, de grande parte das amostras disponíveis, que foram completadas há mais de 30 anos, problema que tem sido gradativamente resolvido pela atualização das amostras de fala com as quais os grupos operam. Relevante também é o fato de que os trabalhos de descrição do português não seriam capazes de oferecer uma visão ampla da realidade sociolinguística do Português Brasileiro em função do descarte de um grande número de dialetos/socioletos em proveito de dialetos dos grandes centros metropolitanos do país.

Além dos vários problemas metodológicos encontrados nas amostras de fala, que acabam por restringir o registro de uma dimensão mais abrangente do português, outro fator limitador é a própria qualidade acústica das gravações realizadas no passado, mediante o uso de recursos analógicos, o que acaba dificultando o estudo das variáveis fonético-fonológicas e prosódicas. Além disso, não é possível, às vezes, o acesso às gravações originais e as regravações apresentam perda da qualidade sonora.

De acordo com Gonçalves (2008), o Projeto ALIP teve por objetivo constituir, em primeiro lugar, o primeiro banco de dados que fornecesse suporte para a descrição do português falado no interior paulista. Além de rigor na coleta de dados e no controle de fatores sociais, a amostra abrange o município de São José do Rio Preto e as cidades adjacentes como: Bady Bassitt, Cedral, Guapiaçu, Ipiranga, Mirassol e Onda Verde, o que lhe dá um sabor ao mesmo tempo urbano e rural à amostra. O aspecto interessante desse projeto e o que faz este banco de dados inédito é que a fala do interior paulista é pouco conhecida em bases científicas seja entre os próprios usuários seja entre os próprios linguistas.

O Projeto ALIP utilizou aspectos teórico-metodológicos empregados em outros projetos já consolidados como o NURC, VARSUL, PEUL, VALP, *Gramática & Discurso*,

etc. No entanto, o Grupo de Pesquisa em Gramática Funcional (GPGF), na proposição do Projeto ALIP, não tinha o interesse por algum fenômeno linguístico em particular, marca característica das amostras desses outros projetos, normalmente constituídos para os estudos variacionistas. Trata-se de uma amostra aberta, fato que favorece a utilização por pesquisadores com diferentes interesses linguísticos. Por essa razão, foram usados apenas os princípios teórico-metodológicos dos outros projetos, pois a proposta do Projeto ALIP se embasa em uma idéia mais ampla e mais aberta, guardando a preocupação de captar o máximo possível do dinamismo linguístico do Português Brasileiro usado na região de São José do Rio Preto. (GONÇALVES, 2008)

2.2.1 A constituição do Banco de Dados Iboruna

O Banco de Dados Iboruna¹⁵ foi idealizado para comportar dois tipos de amostras de fala. A primeira delas foi coletada de acordo com os critérios da Sociolinguística Variacionista (LABOV, 1972), já a segunda foi coletada em contextos interacionais livres, sem o controle de qualquer variável, e de modo secreto (RONCARATI, 1996). O primeiro tipo de amostra foi denominada de *Amostra Censo* ou *Amostra de Comunidade* (AC), e o segundo tipo, de *Amostra Interação* (AI). De acordo com Gonçalves (2008),

esses dois tipos de amostras servem a propósitos diferenciados. AC constitui um tipo mais propício para estudos sociolinguísticos e outros que necessitem de comprovação empírica, e AI, para estudos mais voltados para a interface gramática/discurso, uma vez que sob tal abordagem, concebe-se que a codificação linguística do falante é uma decisão que decorre de um modelo de interação verbal construído na interlocução, ou mais precisamente no discurso. (GONÇALVES, 2008, p. 03)

¹⁵ O nome *Iboruna* (= Rio Preto) tem motivação histórica; é um topônimo de origem tupi-guarani (GONÇALVES, 2008) que se pretendeu atribuir a cidade de São José do Rio Preto por ocasião da comemoração do seu cinquentenário. No entanto, houve uma contundente intervenção do episcopado riopretense que acabou impedindo a mudança e conquistou de modo definitivo a denominação primitiva São José do Rio Preto, reduzida a Rio Preto de 1906 a 1944.

Apesar de o Iboruna ser constituído por Amostra Censo (AC) e Amostra de Interação (AI), este trabalho se limitará à primeira somente, já que as variáveis sociais nas Amostras de Interação – escolaridade, gênero, faixa etária - que aparecem explícitas nas Amostras Censo, não foram controladas, por consistirem em gravações anônimas, feitas sem o conhecimento prévio dos informantes; conseqüentemente, os participantes das interações não constituem grupos homogêneos.

Em relação aos aspectos éticos da pesquisa, todos os informantes, que forneceram os dados linguísticos, foram previamente consultados e devidamente esclarecidos sobre os objetivos da pesquisa antes das gravações. A concordância foi formalmente obtida mediante a assinatura de um termo de compromisso.

Antes de enfocar precisamente a subamostra deste trabalho, constituída pela Amostra Censo, é necessário discutir como a amostra é formada, fato amplamente descrito na subseção seguinte.

2.2.2 A Amostra Censo

Para a constituição da Amostra Censo, elegeram-se variáveis sociais relevantes para os estudos sociolinguísticos. Desse modo, foi garantida devidamente a representatividade do censo linguístico. Segue abaixo o quadro com as variáveis e variantes controladas na *Amostra Censo*:

<i>Variável</i>	<i>Variantes</i>
1 - Sexo/ gênero	(1) masculino; (2) feminino
2 - Faixa etária	(1) 07 a 15 anos; (2) 16 a 25 anos; (3) 26 a 35 anos; (4) 36 a 55 anos; (5) mais de 55
3 - Escolaridade	(1) 1º ciclo do Ensino fundamental; (2) 2º ciclo do Ensino fundamental; (3) Ensino Médio; (4) Ensino Superior
4 - Renda	(1) mais de 25 sal. mínimos; (2) 11 a 24 sal. mínimos; (3) 6 a 10 sal. mínimos; (4) até 5 sal. mínimos.

Quadro 2: Variáveis controladas na constituição da Amostra Censo

A partir do cruzamento dos quatro grupos de fatores sociais, expostos acima, constituíram-se os perfis da *Amostra Censo* ou *Amostra de Comunidade (AC)*, composta de 160 células, definindo, desse modo, os informantes contatados. Entretanto, oito células são impossíveis de serem preenchidas, pois demandam faixa etária de *7 a 15 anos e nível superior* concomitantemente. Dessa maneira, estas oito células foram excluídas, restando um total de 152 informantes, como se pode ver no quadro abaixo

RENDA / GÊNERO FAIXA ETÁRIA /ESCOLARIDADE		+ DE 25 SM		DE 11 A 24 SM		DE 6 A 10 SM		ATÉ 5 SM		SUB- TOTAL DE INF.	TOTAL DE INF.
		MAS	FEM	MAS	FEM	MAS	FEM	MAS	FEM		
7 A 15 ANOS	1°.CEF	001	002	003	004	005	006	007	008	8	24
	2°.CEF	009	010	011	012	013	014	015	016	8	
	E. M	017	018	019	020	021	022	023	024	8	
16 A 25 ANOS	1°.C EF	025	026	027	028	029	030	031	032	8	32
	2°. C EF	033	034	035	036	037	038	039	040	8	
	E. M	041	042	043	044	045	046	047	048	8	
	SUP	049	050	051	052	053	054	055	056	8	
26 A 35 ANOS	1°.C EF	057	058	059	060	061	062	063	064	8	32
	2°. C EF	065	066	067	068	069	070	071	072	8	
	E. M	073	074	075	076	077	078	079	080	8	
	SUP.	081	082	083	084	085	086	087	088	8	
36 A 55 ANOS	1°.C EF	089	090	091	092	093	094	095	096	8	32
	2°. C EF	097	098	099	100	101	102	103	104	8	
	E. M	105	106	107	108	109	110	111	112	8	
	SUP	113	114	115	116	117	118	119	120	8	
+ DE 55 ANOS	1°.C EF	121	122	123	124	125	126	127	128	8	32
	2°. C EF	129	130	131	132	133	134	135	136	8	
	E. M	137	138	139	140	141	142	143	144	8	
	SUP.	145	146	147	148	149	150	151	152	8	
SUB-TOTAL DE INF.	1°.C EF	5	5	5	5	5	5	5	5	40	152
	2°. C EF	5	5	5	5	5	5	5	5	40	
	E. M	5	5	5	5	5	5	5	5	40	
	SUP.	4	4	4	4	4	4	4	4	32	
TOTAL DE INFORMANTES		19	19	19	19	19	19	19	19		
		38		38		38		38			
		76				76					

Quadro 3: Distribuição e identificação dos informantes da Amostra Censo por variáveis sociais (RUBIO, 2008)

De acordo com Gonçalves (2008), a manutenção dos 152 informantes não compromete a representatividade da amostra, já que a variação é bastante padronizada, fato

que autoriza generalizações acerca da língua usada numa comunidade como um todo. No entanto, Paiva (1999, p.7) adverte que essas implicações, “embora não possam ser ignoradas, não chegam a comprometer o estudo sociolinguístico” desde que se preste atenção para duas questões essenciais: (i) a necessidade de usar técnicas estatisticamente validas de amostragem e (ii) ter o conhecimento prévio das dimensões relevantes da estratificação, de maneira que seja possível planejar a amostragem corretamente.

Se fosse incluída como variável a *origem geográfica do informante*, definida pelas sete cidades do censo linguístico, e, se fosse atendida a recomendação de cinco informantes por célula, a amostragem final seria equivalente a 5.320 informantes, o que tornaria impossível a proposta de levantamento. Dentro dessas restrições, a definição da Amostra Censo foi pautada pelos seguintes critérios: (1) preenchimento de apenas um informante por célula; (2) aplicação do método aleatório simples (SILVA, 2003), para que fosse possível a distribuição proporcional de 152 informantes ao número de habitantes das regiões geográficas consideradas, conforme o quadro abaixo:

Cidades da Região de São José do Rio Preto	População *	Número de informantes
1. Bady Bassit (12 km ao sul de SJRP)	11.475	04
2. Cedral (14 km, ao sul de SJRP)	6.690	02
3. Guapiaçu (16 km , ao leste de SJRP)	14.049	05
4. Ipiranga (18 km, ao norte de SJRP)	3.461	01
5. Mirassol (14 km, ao sul de SJRP)	48.233	16
6. Onda Verde (25 Km, ao norte de SJRP)	5.407	02
7. São José do Rio Preto ¹⁶	357.705	122
Total da população representada	447.020	152

Quadro 4: Distribuição dos informantes proporcionalmente à densidade populacional das cidades da região¹⁷

De acordo com o trabalho de Gonçalves (2008), na composição do Banco de Dados Iboruna, o método aleatório simples de distribuição dos perfis linguísticos foi aplicado com

¹⁶ Inclui os Distritos de Talhados e Engenheiro Schimidt.

¹⁷ Fonte: IBGE (Censo 2000)

base em quatro procedimentos. O primeiro consistiu na distribuição do total de informantes de modo proporcional ao número de habitantes de cada área. O segundo procedimento definiu a seleção de perfis sociais em relação à identificação das localidades: mediante uso de uma urna com a identificação dos perfis sociais e de outra com a identificação das cidades, procedeu-se a escolha subsequente de um perfil social e de uma cidade. Esse modo de proceder definiu a origem geográfica do informante. Já o terceiro procedimento foi responsável pela reposição na urna da cidade escolhida até que o total de seus informantes fosse definido, garantindo que todos os fatores sociais tivessem a mesma possibilidade de ocorrer em qualquer uma das cidades. No último passo, os procedimentos foram repetidos até que todos os perfis sociais estivessem distribuídos.

Os informantes, que foram entrevistados para a constituição das amostras, contribuíram com cinco tipos de texto orais diferentes, com base na metodologia de Votre e Oliveira (1995). A única exigência feita, para a constituição do banco de dados, era que os informantes residissem nas cidades abrangidas e tivessem nascido nela e que ali morassem desde os cinco anos de idade

2.3 Composição da amostra

Inicialmente, a amostra deste trabalho seria composta por 95 dos 152 inquéritos do Banco de Dados Iboruna, com exclusão da faixa etária de 7 a 15 anos e do primeiro ciclo do ensino fundamental. Decidimos, no entanto, ampliá-la, inserindo o primeiro ciclo do ensino fundamental, a fim de deixá-la mais bem distribuída socialmente, com a presença de todos os níveis de escolaridade de que o banco de dados dispõe. Não consideramos, neste trabalho, a faixa etária de 7 a 15 anos, pois preferimos trabalhar com uma apreensão cognitivamente mais consolidada da norma de prestígio.

Desse modo, servem de base para as investigações os cinco tipos de texto de 128 Amostras Censo. Como se trata de uma amostra aleatória estratificada, haveria a necessidade de somente 32 inquéritos, já que com essa quantidade é possível contar com dois gêneros, quatro faixas etárias e quatro níveis de escolaridade. No entanto, o acréscimo de inquéritos é necessário, porque, por um lado, estruturas predicativas são relativamente escassas e, por outro, como esta pesquisa não acolhe predicativos com mais de um constituinte, a escassez de dados acabaria se agravando. Constituída a amostra, realizamos uma audição rigorosa de todas as entrevistas para confirmar se a concordância de número realmente ocorre ou não nas estruturas predicativas, atividade que foi, em geral, facilitada pela boa condição do áudio das gravações.

O universo de pesquisa é composto, portanto, por um recorte do Banco de Dados Iboruna, contendo 128 de 152 gravações das Amostras Censo. Cada Amostra é composta por cinco gêneros textuais: *narrativas de experiência pessoal*, *relato de narrativa recontada*, *relato descritivo*, *relato de procedimento* e *relato de opinião*. A manutenção do cópuz com os cinco gêneros textuais possibilita um número considerável de ocorrências, o que propicia condições satisfatórias em termos quantitativos para a análise de todas as variáveis dependentes.

Os tipos de texto se mostraram muito relevantes para a obtenção de ocorrências com estruturas predicativas. Em estudo anterior, trabalhamos apenas com as narrativas de experiência pessoal, resultando apenas 152 dados. Esse número baixo de ocorrências impossibilitou a análise de alguns grupos de fatores como a *Saliência Fônica*. O estudo realizado apenas com as narrativas de experiência (NE) deveu-se ao fato de elas serem mais próximas de uma conversa natural, na tentativa de eliminar o paradoxo do observador (LABOV, 2008), já que se trata de uma história pessoal do informante, em que o documentador não aparece ou tem uma participação discreta.

Conseqüentemente, a incorporação da totalidade possível da amostra é fundamental para esta pesquisa, primeiramente pela própria questão estatística e, em segundo lugar, pelo fato de os outros tipos de textos se mostrarem mais produtivos que a própria narrativa de experiência. Por exemplo, os relatos de opinião (RO) se mostraram mais relevantes que as primeiras. Nesse tipo de narrativa, o documentador introduz um assunto de sabor bem contemporâneo, isto é, informações com as quais os informantes possam ter contato facilmente. Na ocasião, os assuntos mais abordados foram o “mensalão”, os problemas da educação brasileira e a máfia do apito ocorrida no campeonato brasileiro de 2007. Aparentemente, os informantes se mostram mais receptivos ao tema proposto pelo documentador, com predisposição para discutir o assunto. Quando o informante não tem domínio do tema proposto, o documentador logo sugere outro e a conversa se estabelece rapidamente.

As *narrativas recontadas*, os *relatos de procedimento* e os *relatos de descrição* apresentam uma quantidade considerável de ocorrências de predicativos também, mas não comparável à incidência de casos nas *narrativas de opinião* e nas *narrativas de experiência*. Vale ressaltar que as *narrativas de procedimento* apresentam, por sua vez, um menor número de dados, já que se trata da discussão de “como fazer algo”, ou seja, o informante normalmente discorre sobre como se faz uma receita, ou como é a sua rotina no trabalho. Como esse tipo de texto é marcado por tarefas que devem ser cumpridas, ele é menos propício ao uso de adjetivos ou participios passivos. Conseqüentemente, o número de ocorrências mostra-se bem menor.

Entretanto, apesar de os tipos de textos apresentarem diferenças de frequência no uso da marcação formal de plural, eles não serão considerados como um grupo de fatores, pois a intenção desta pesquisa é a de examinar se as motivações funcionais e formais estão em competição ou qual delas se mostra mais relevante para a explicação da CN nos predicativos.

Por essa razão, *tipo de texto* e outras variáveis, que poderiam ser amplamente discutidas, como *contexto fonológico seguinte*, não fazem parte do escopo deste trabalho, já que o objetivo principal não é a variação em si, mas as motivações que a estão regendo. No entanto, são grupos de fatores a serem utilizados em outros trabalhos e que provavelmente podem mostrar resultados expressivos para a pesquisa sociolinguística.

Conforme discutido, a constituição da amostra provém da gravação de 152 informantes, cujo perfil é determinado pela combinação de fatores sociais: gênero/sexo, escolaridade, faixa etária e classe social (renda familiar). Desse modo, os grupos de fatores sociais têm a seguinte constituição:

<i>Variável</i>	<i>Variantes</i>
1 - Sexo/ gênero	(1) masculino; (2) feminino
2 - Faixa etária	(1) 16 a 25 anos; (2) de 26 a 35 anos; (3) de 36 a 55 anos; (4) mais de 55 anos
3 - Escolaridade	(1)1º ciclo do Ensino fundamental; 2º ciclo do Ensino fundamental; (2) Ensino Médio; (3) Ensino Superior

Quadro 5: Variáveis sociais controladas na constituição da subamostra

Vale observar que as diferenças de renda familiar foram descartadas em razão da forte correlação entre escolaridade e renda, aspecto que se verifica também nos dados do Iboruna. A inclusão de distinções de sexo/gênero deve-se ao fato de a literatura sociolinguística ser fortemente comprometida com o princípio de que mulheres em geral são mais sensíveis que os homens ao valor social de variáveis linguísticas; finalmente, a inclusão de quatro faixas etárias objetiva verificar se marcação ou supressão de pluralidade é uma variação estável ou se está sujeita à mudança em função da correlação possível entre o caráter inovador da variante conservadora e o aumento gradual de idade, conhecida por “distribuição em tempo aparente” na literatura sociolinguística (LABOV, 1972).

2.3.1 Da seleção das ocorrências

A partir da análise dos 128 inquéritos, procedemos ao levantamento de todas as ocorrências possíveis de manifestação de concordância nominal em estruturas predicativas, em que a regra formal de concordância pode ser aplicada, como se vê em (9-10) e não aplicada, como se observa em (11-12):

- (9) *a sala de aula de terce(i)ra e quarta séries são salas onde não existe alfabeto... porque os alunos já estão **alfabetizados**...* (AC 88/DE/L 335-336)
- (10) *sessenta por cento dos brasile(i)ros são **católicos**...Desses sessenta só cinco por cento são **praticantes**...né?* (AC023//L.544-545)
- (11) *aí ela falo(u) pra mim – “não nós tá **brigado** tal cê a::cha todo mundo sabe que nós tá **brigado**...”* (AC019/NE/L.10-11)
- (12) *minha vó num queria o namoro porque na época éh...as coisa era muito **rígida**...* (AC034/NR/L.42)

Por uma questão de recorte metodológico, foram descartados os casos de hipercorreção, caracterizados como desvios da concordância nominal. Foram excluídas também ocorrências que não apresentam ordem SVO para preservar rigorosamente o princípio de comparabilidade. Como o alvo é o predicativo, interessa verificar se a pluralidade desse constituinte é afetada por marcas precedentes, o que não seria possível detectar nos casos de inversão. Além disso, os casos de inversão, como demonstram os exemplos (13-15), são mais bem adequadamente tratados no âmbito de outra variável – a concordância verbal – que está fora do escopo do presente trabalho.

- (13) *(...) aí ela disse que lá onde ela foi:: ... acho que era:: muito **forte** as ondas... a areia era completamente diferente (...)* (AC70/NR/L. 69-70)
- (14) *(...) **SÓ** que o padrasto dela batia na mãe e ela via porque eles moravam em dois cômodos... entendeu?... então eu não sei nem como eram **divididos** esses dois cômodos... não sei nem onde a M. L. dormia (...)* (AC86/RO/L. 653-656)
- (15) *(...) eles tinham um po(u)co de vergonha e num tinha... esse costume de dançá(r)... e foi **feito** os ensa::ios né? foi trabalhado tudo direitinho (...)* (AC88/NE/L. 25-27)

2.3.2 Da definição da variável dependente

A investigação da marcação de pluralidade na variedade falada na região de São José do Rio Preto se restringe, nesta pesquisa, à análise de um envelope de variação constituído pela variável binária “presença ou ausência de marcação de pluralidade nas estruturas predicativas”, formalizada como <s>, representando o arquifonema /S/. Essa categoria representa, por sua vez, as variantes presença ou ausência de qualquer realização fonética, com alterações morfofonêmicas, da marca de pluralidade em todos os constituintes do SN. O contexto de análise dessa variável é representado pela posição sintática de predicativo, constituído de um único elemento (adjetivo, substantivo ou particípio) que pode ser acompanhado de palavras invariáveis ou não, como se pode ver nos exemplos (16), (17) e (18).

- (16) *a sala de aula de terce(i)ra e quarta séries são salas onde não existe alfabeto... porque os alunos já estão **alfabetizados**...* (AC 88/ L 335-336/ DE)
- (17) *(...) mas as casas ali são bem **antigas**... né?...* (AC 120/ L 266/ DE)
- (18) *(...) que nós falávamos que a S. num era boa pesso::a... que num era boa compnhi::a... aí diz que o marido dela respondia assim – “imagina eles são muito **sistemático** eles tão com inVEJA dela”* (AC 050/ L 134-136/ NR)

Trata-se de um caso mais geral de mudança fonológica afetando, na maior parte das vezes, sílabas átonas finais, que se caracterizam como posições de enfraquecimento regular de segmentos fonológicos do português. Essa regra é aplicada tanto em palavras monomorfêmicas (como *lápiz*, *menos*) quanto em palavras bimorfêmicas (como *casas*, *patos*, *meninas*). Mesmo representando um morfema de plural, a redundância de marcação no SN leva a uma supressão maior da fricativa em contextos de palavras bimorfêmicas do que em palavras monomorfêmicas.

Há redundância na marcação de pluralidade do português, porque em todas as palavras numa dada sentença deve haver a marca formal explícita de plural, no entanto, isso só acontece no português padrão escrito e não sistematicamente no português vernacular. Naro e Scherre (1997) afirmam que

O português vernacular do Brasil, diferentemente do português de Portugal, apresenta variação sistemática nos processos de concordância de número, exibindo variantes explícitas e variante zero de plural em elementos verbais e nominais. (Naro; Scherre, 1997, p.93)

No português brasileiro, presença e ausência de plural se alternam em posição final de palavra, contexto que apresenta grande potencial de variação em função de um processo de erosão fonológica, por distensão dos órgãos fonadores em posição átona. Apresentamos abaixo, em (19-20), alguns exemplos, extraídos do cópuz em análise, que mostram a presença e a ausência de marcação formal de pluralidade em contexto de predicativo, que constitui, portanto, neste trabalho, a variável dependente.

- (19) *eu tinha medo... PAI sempre é mai/éh num sei alguns são mais **rigoroso** então eu tinha me::do eu num saía de ca::sa...* (AC38/NE/L. 37-39)
- (20) *(...) as/os/ avenidas as avenidas são **largas** sabe bem espaçosa é um lugar limpo...* (AC45/DE/L.251-252)

Selecionadas as ocorrências, cada uma delas foi analisada segundo os critérios definidos nos grupos de fatores relacionados a seguir, os quais constituem os contextos variáveis definidos para esta pesquisa.

2.4 As variáveis linguísticas

Cada grupo foi representado por seus respectivos fatores, constituindo ao todo um total de oito variáveis independentes de natureza linguística ou interna, a saber: (i) paralelismo formal no discurso; (ii) características formais do sujeito; (iii) características formais do verbo; (iv) configuração estrutural do predicativo/particípio; (v) saliência fônica da oposição singular *versus* plural; (vi) número de marcas precedentes no sujeito e no verbo; (vii) tipo estrutural do sujeito e (viii) material interveniente. Além disso, recorreremos a três variáveis independentes de natureza externa ou social: (i) gênero; (ii) idade e (iii) escolaridade; conforme amplamente discutidas na literatura recente, a partir dos trabalhos de Scherre (1991) e de Dias (1996).

Por “paralelismo formal no discurso” entende-se aqui a recorrência no contexto discursivo de predicativos com a mesma configuração formal, em termos de marcação explícita ou não de plural; a hipótese que se quer testar é a de que um predicativo com marcas de plural explícitas ou não aciona a ocorrência, no texto subsequente, de outro predicativo formalmente idêntico (SCHERRE, 1991). Nesse grupo, adotamos os fatores utilizados por Scherre (1991), todos eles estão devidamente exemplificados em (21a-d) respectivamente:

- i. predicativo isolado;
- ii. predicativo primeiro de uma série;
- iii. predicativo precedido de predicativo com plural explícito;
- iv. predicativo precedido de predicativo com plural zero;

- (21) a (...) *é bonito as praias são **bonitas** mas a cidade assim o centro assim... é muito... é muito poluído assim...* (AC 044/DE/L 164-165)
- b (...) *sessenta por cento dos brasile(i)ros são **católicos**...Desses sessenta só cinco por cento são **praticantes**...né?* (AC023//L.544-545)

- c (...) *eu acho que as crianças tão muito assim... éh:: **desorienta::das** éh éh só pensam nisso malemá estão fican(d)o **mocinhos** já ...já tão atrás disso...* (AC38/RO/L. 253-255)
- d (...) *ai ela falo(u) pra mim – “não nós tá **brigado** tal cê a::cha todo mundo sabe que nós tá **brigado**...”* (AC019/NE/L.10-11)

Com relação ao grupo de fatores “características formais do sujeito”, pretende-se verificar a marcação de plural no predicativo em correlação com a mesma configuração flexional no sujeito, para testar a hipótese formal de que a presença ou ausência de marcas conduz a uma configuração recorrente no predicativo; em caso de situação inversa, é possível rejeitar essa hipótese e, nesse caso, verificar se há outro princípio funcional implícito como a Condição de Distintividade de Kiparsky (1972) já mencionada. Nesse grupo, denomina-se “sujeito não-explicito” todos os tipos de sujeito considerados por Scherre (1988, 1991, 1994) como sujeito nulo mais o sujeito desinencial. Esse grupo de fatores aparece exemplificado em (22a-e) respectivamente:

i. sujeito não-explicito;

ii. sujeito explícito (i) com todos os elementos nominais flexionáveis marcados (ii) com os últimos elementos com marca formal de plural e (iii) ou com a última marca neutralizada por contexto fonológico seguinte;

iii. sujeito explícito sem a(s) última(s) marca(s) formal(is) explícitas de plural ou sujeito coordenado singular;

iv. sujeito explícito com marca semântica de plural (*nós* e numerais isolados);

v. sujeito explícito com marca formal de plural totalmente neutralizada.

- (22) a *então nós casamos e:: a gente trabalhava muito né?... mas **éramos** felizes passeávamos... e viajavamos muito* (AC 138/NE/L.15-16)
- b ***meus pais** são bastante são evangélicos né?... (AC035/NR/L.202)
as avenidas **as avenidas são** largas sabe bem espaçosa é um lugar limpo...* (AC045/DE/L.251-252)

- c *minha vó num queria o namoro porque na época éh...as coisa era muito rígida... (AC034/NR/ L.42)*
num tinha muita informação dentro de casa... porque meu pai e minha mãe sempre foi muito crente... (AC 036/ L 229/ NR)
- d *ái surgiu a briga... mas hoje nós é colega normal (AC065/NE/L.29)*
- e *então é bem gostoso... e lá também... éh:: eles são bem unidos então... tem sempre... um pessoal da nossa idade... (AC46/DE/L.276-277)*

A variável independente “características formais do verbo” teria um efeito similar ao da anterior. Sua análise requer, portanto, a realização de um cruzamento entre os dois grupos de fatores, para o exame do efeito possível da interação entre sujeito e verbo com ou sem marcação de plural sobre o predicativo. Consequentemente, foram definidos os seguintes fatores baseados em Scherre (2003), conforme ilustrados em (23a-c), respectivamente:

- i. zero verbal;
- ii. verbo com marca explícita de plural;
- iii. verbo sem marca de plural.

- (23) a *esse pessoal eles não são dono... são case(i)ro lá... esse pessoal é primo do meu marido... são um amor de pessoas [Ø] super hospitaleiro (AC028/NE/L.54-56)*
- b *porque:: é aquela coisa né? a maioria das pessoas são preconceituosas quan/... quanto a pessoas crentes né?... (AC45/NE/L.19-20)*
- c *minha vó num queria o namoro porque na época éh...as coisa era muito rígida... (AC034/NR/ L.42)*

Por “configuração estrutural do predicativo/particípio” entende-se aqui meramente a classe gramatical do predicativo, que está restrita à (i) adjetivos, (ii) participios passivos, e (iii) substantivos isolados atuando como fatores, conforme se vê em (24a-c), respectivamente. Pode haver diferença estatística entre as classes dos predicativos propriamente ditos, que são, por definição, os adjetivos e os demais. Entende-se, além disso, que os participios passivos

fazem parte de uma classe mais integrada com a cópula do que os demais, o que sugere relevância desse tipo de distinção formal. No entanto, diferentemente de outros trabalhos, como Scherre (1991) e Dias (1996), vale repetir que não pretendemos considerar predicativos de dois ou mais constituintes, devido à dificuldade de medir saliência fônica, que ficaria restrita ao núcleo.

- (24) a *pra robá(r) eles são bem **inteligentes** né?... (AC45/RO/l. 361-363)*
- b *ái minha mãe tam(b)ém ficô(u) meia tris::te brigava comi::go tu::do... aí tam(b)ém foi no outro boteco mas não no mesmo bar que eles foram... aí chegô(u) isso era umas quatro hora da tarde os dois tava **embriagado**... (AC39/NE/L.12-14)*
- c *ele foi brincá(r) assim com ela... –“ai vamo(s) ficá(r) **noivo** vamo(s) ficá(r) **noivo(s)**”-brincan(d)o e ela levô(u) a sério e falô(u) assim –“ai vamo(s)” – ficô(u) toda empolgada – “ai vamo(s) ficá(r) **noivo(s)**” - ... aí ela encheu o saco então teve que colocá(r) aliança né?... (AC44/NR/L.139-142)*

A hipótese que rege a variável “saliência fônica da oposição singular *versus* plural” é a de que o grau progressivamente mais elevado de saliência fônica entre as flexões de singular e de plural é diretamente proporcional à incidência de marca de plural no constituinte em questão, conforme amplamente discutido na literatura sociolinguística (SCHERRE, 1988, 1991, 1994; DIAS, 1996). Os subgrupos definidos estão exemplificados em (25a-f), respectivamente.

- i. plural duplo (novo/novos);
- ii. plural irregular em <l> (fácil/fáceis);
- iii. plural irregular em <s,r> (feliz/felizes; cor/cores);
- iv. plural irregular em <ão>;
- v. plural regular monossílabo (bom/bons);
- vi. plural regular de base paroxítona (bonito/bonitos);
- vii. plural regular de base proparoxítona (hipótese/hipóteses).

- (25)
- a *a gente tenta convencê(r) que num é sério mas se eles...têm certeza se eles são vamo(s) dizê(r) teimosos para isso de(i)xa eles... (AC023/ RO/ L. 534)*
 - b *graças a Deus somo(s) saudável... (AC028/RO/L.185)*
 - c *mas graças a Deus somos muito feliz (AC131/NE/L.054)
eu achei que ela melhorô(u) bastante depois disso daí não que as outras escolas seja piores... mas o povo tem reclamado bastante né? (AC 067/RO/L 361-363)*
 - d *já pode mexê(r) quem gosta... éh:: éh:: e as lan houses tam(b)em é bom pra quem gosta de computador pode í(r) lá... (AC37/NE/L. 344-345)*
 - e *as praia ... e::... ah e o povo de lá que são bastante bonito né?... tudo do do tipo o Brasil... tem morena loira ruiva... (AC39/DE/L. 168-169)*
 - f *ele tem um po(u)co de ciúmes dela... que as meninas são bem próximas da dona N. e a... que trabalha lá... então eu acho que é isso... (AC47/NE/L.54-55)*

O grupo “número de marcas precedentes no sujeito e no verbo” refere-se ao fato de que, se não houver marca formal de plural em todos os constituintes do SN ou no verbo, os outros SNs podem ser não-marcados também, inclusive o predicativo. A hipótese que se pretende comprovar é a de que, pelo princípio da Economia e pela aplicação da condição de distintividade de Kiparsky (1972), a existência de um elemento flexionável no SN, passível de receber marca formal de plural, pode desfavorecer a marcação no predicativo, por motivo de redundância. Esse grupo não se encontra em outros trabalhos, já que normalmente verbo e sujeito são tratados em grupos de fatores diferentes. O grupo de fatores está exemplificado em (26a-i), respectivamente:

- i. sujeito com uma marca e verbo nenhuma;
- ii. sujeito com uma marca e verbo com marca;
- iii. sujeito com duas ou mais marcas e verbo nenhuma;
- iv. sujeito e verbo com todas as marcas;
- v. sujeito não-explicito e verbo com marca;
- vi. sujeito composto e verbo com marca;
- vii. sujeito composto e verbo sem marca;

viii. sujeito pronominal e verbo com marca;

ix. sujeito pronominal e verbo sem marca.

- (26) a *que vida elas vão tê(r)?... pára com tudo... nem estudá(r) num consegue mais né? então elas fica perdida...* (AC140/RO/L.455-456)
- b *aquilo pra gente foi assim... tanto até os menino ficaram super feliz... de vim uma irmãzinha pra eles né?* (AC 102/NE/L 70)
- c *a gente sabe de causo que aconteceu a pessoa sobe em lugar errado né?... então esses passos da segurança é muito importante isso aí* (AC139/RP/L.375-376)
- d *éh meus tios ainda são vivos* (AC/NR/L.154)
- e *o clube que a gente frequenta muito é o Palestra um clube antigo também somo(s) sócio lá...* (AC028/RO/L.164)
- f *dama e valete e reis tam(b)ém num são fortes...* (AC 029/RP/L.151)
- g *o carro da mensagem ao vivo tocan(d)o a música já chegô(u) filman(d)o tudo... a minha irmã e o meu cunhado ficô(u) assustado que/ num tava esperando né?* (AC 066/ L 274-275/ DE)
- h *pediu pra eles e eles ficaram muito quieto... e:: acabô(u) levan(d)o... umas coisa(s) de pertence...* (AC 067/ L 96-97/ NR)
- i *então tinha que chegá(r) se os cara... chegava e ficava com/ que fi/... queria que eles ficava eles era moleque que eles ficava bejan(d)o a mão dos cara aquelas coisa de antigo né?* (AC 119/ L 107-109/ NR)

Essa variável vai ao encontro da hipótese principal deste trabalho em termos da análise da atuação dos princípios funcionais, de que a marcação de plural no SN sujeito ou no verbo pode provocar a ausência de marcas no próprio predicativo.

Incluiu-se, ainda, o grupo de fatores “tipo estrutural do sujeito” neste trabalho, em função da possível influência que esse grupo pode exercer na concordância dos predicativos, com base no fato de que alguns tipos de sujeito aparecem de modo mais frequente que outros. Esse grupo de fatores, que é frequentemente aplicado aos estudos da concordância verbal, pode ser de grande utilidade para o exame de concordância verbal (RUBIO, 2008), de modo que alguns tipos de sujeito tendem a influenciar a marcação formal de pluralidade e outros

tipos não. Foram definidos os seguintes fatores, devidamente ilustrados em (27a-e), respectivamente:

- i. SN pleno simples;
- ii. SN pleno genérico (desprovido de determinantes e modificadores restritivos);
- iii. SN pleno composto;
- iv. Pronome pessoal;
- v. Quantificador;

- (27) a *os olhos são castanho amendoim...* (AC35/DE/L.290)
- b *vai acontecê(r) a mesma coisa que acontece com a droga... **drogas** são proibidas? são...* (AC050/RO/L.393-394)
- c *o pai dela e a mãe dela são muito... assim... rígidos né? pra namoro essas coisas...* (AC47/DE/L.109-110)
- d *a gente tenta convencê(r) que num é sério mas se eles... têm certeza se **eles** são vamo(s) dizê(r) teimosos para isso de(i)xa eles...* (AC023/ RO/ L. 534)
- e *meus amigos?... **alguns** são legais o(u)tros são fa/muito falsos...* (AC 029/RO/ L.179-180)

Esse grupo de fatores apresentava também os fatores *pronome indefinido*, *pronome demonstrativo*, além de *pronome relativo*, que será discutido no fator material interveniente. Os *pronomes demonstrativos*, os *pronomes indefinidos* e os *quantificadores* foram unidos em apenas um fator denominado *quantificador*, pois os pronomes indefinidos cumprem a função de fornecer uma quantificação indeterminada. Os pronomes encontrados foram “alguns”, “todos” e “outras”, como se pode ver nos exemplos abaixo:

- (28) *meus amigos?... **alguns** são legais outros são fa/ muito falsos* (AC 029/ RO/L.179-180)
- (29) *(...) a gente tava com um po(u)co de medo de voltá porque **todos** tavam meio **altos** já e a estrada era simples né?* (AC 049/ NE/L 50-51)

Já os *pronomes demonstrativos* foram unidos juntamente aos *indefinidos* e aos *quantificadores*, porque dispõem da mesma estrutura sintática, isto é, os *pronomes demonstrativos*, *indefinidos* e os *quantificadores* aparecem num SN sujeito com núcleo nominal vazio, vejamos os exemplos (30), (31) e (32):

- (30) *foi o que o:: técnico Bernardinho fez agora que ele tá trabalhando com esses jogadores há:: muito tempo assim... alguns Ø são bem velhos já de carre(i)ra... alguns são mais novos* (AC 054/RO/ L 312-314) (*pronome indefinido*)
- (31) *porque Deus que dá a oportunidade... pras pessoas ouvirem a palavra Dele... aqueles Ø que são escolhido Dele... os que são escolhidos Dele vão ouvi(r) a palavra Dele* (AC 106/NE/ L 187-188) (*pronome demonstrativo*)
- (32) *(...) que:: na época é a:: vaga era pra uma pessoa... mas das três pessoas... duas Ø foram consideradas aptas* (AC 114/ NE/L 55-56) (*quantificador*)

A variável “tipo estrutural do sujeito” não foi utilizada nos trabalhos de Scherre (1988, 1991) e de Dias (1996). Primeiramente, como o principal interesse teórico da tese de Scherre (1988) é fornecer uma reanálise da concordância nominal entre os elementos flexionáveis do SN, a autora realiza a análise atomística e não-atomística desses elementos; em segundo lugar, seu objeto de estudo não é a CN apenas em estruturas predicativas. Essa variável também não foi considerada no trabalho de Dias (1996), mas é frequente nos estudos da concordância verbal. Por essa razão, a variável causou interesse na análise da CN em predicativos, para verificar se o seu comportamento é similar ao dos estudos sobre CV.

Ainda sobre a variável “tipo estruturas do sujeito”, operou-se uma mudança de nomenclatura em um dos fatores depois de analisar os dados qualitativamente. A análise do fator *SN pleno nu* permitiu constatar que todas as ocorrências se identificam com sujeitos genéricos, por essa razão, a nomenclatura foi mudada para “*SN genérico*”, conforme mostram os exemplos de (33-34):

- (33) *vai acontecê a mesma coisa que acontece com a droga **drogas** são proibidas? são... as armas são proibidas? são mas todo mundo vai sabê onde comprá...* (AC 050/RO/ L 393-395)
- (34) *eu sô(u) contra... esse tipo de aborto... por outro lado... tem **mulheres** que ficam grávidas de um estrupo éh:: é um gravidez no/ éh:: num pensa na criança rejeitada ele que tem uma::... começa a rejeitá(r) essa criança* (AC 085/RO/L 447-449)

Outro critério adotado, que se relaciona com o grupo de fatores “tipo estrutural do sujeito”, foi operado em relação às estruturas complexas. Como se trata de um tipo de estrutura que apareceu em apenas nove ocorrências, as estruturas complexas foram consideradas como um SN simples, isto é, a expressão “a maioria de” foi excluída da análise. Outro ponto interessante é que oito das nove ocorrências aparecem com todas as marcas formais de pluralidade, conforme exemplificado em (35), (36) e (37).

- (35) *(...) **a maioria das pessoas** são preconceituosas quan/... quanto a pessoas crentes né?...* (AC 045/NE/L 20)
- (36) *(...) a gente brinca porque a gente fala que é cidade de VELHO... porque **a maioria das pessoas** são mais velhas cidade pequenininha* (AC 046/DE/L 265-266)
- (37) *a expresso Itamarati... a maioria dos motoristas são crentes... porque **a maioria dos crentes** não bebe.* (AC 114/RO/L 654-655)

Por fim, descreve-se o último grupo de fatores que integra este trabalho. O grupo de fatores “material interveniente” tem o objetivo de revelar se a presença de material linguístico, principalmente entre o verbo e o predicativo pode condicionar a não-marcação de pluralidade no predicativo (SCHERRE, 1991). Esse grupo é constituído apenas por dois fatores, (i) presença e (ii) ausência de material interveniente, conforme aparecem exemplificados em (38a-b) respectivamente.

- (38) a *as praia ... e::... ah e o povo de lá que são **bastante** bonito né?... tudo do do tipo o Brasil... tem morena loira ruiva...* (AC39/DE/L. 168-169)

- b (...) *a gente tenta convencê(r) que num é isso mas se eles... têm certeza se eles são vamo(s) dizê(r) teimosos para isso de(i)xá (eles)...* (AC 023/RO/L 534)

Primeiramente, material interveniente entende-se aqui apenas o material linguístico presente entre o verbo e o predicativo, pois o objetivo inicial era observar se a presença de material afetaria a utilização de marcas formais de plural no predicativo. Numa primeira rodada das ocorrências no *Goldvarb*, esse grupo de fatores não se mostrou relevante, pois as diferenças de peso relativo são mínimas entre presença e ausência de material.

Decidiu-se, então, em um segundo momento, incluir como material interveniente, também, as orações cujo sujeito é o *pronome relativo*, que não seria visto como um tipo estrutural de sujeito, mas como material interveniente entre o sujeito e o verbo, já que todas essas ocorrências preservam o SN sujeito que o pronome relativo representa. Essas ocorrências seguem sempre a mesma ordem, ou seja, o “que” é antecedido pelo SN que ele representa, isto é, o sujeito da oração, como se pode ver nos exemplos (39), (40) e (41).

- (39) (...) *porque no tempo delas... os casamentos que eram feitos no povoado... depois ia pá/ pro sítio ou a fazenda... num tinha como transportá(r) todo mundo.* (AC 80/NR/L 86-89)
- (40) (...) *no meu... éh:: doutorado né? eu construo diagramas de fases pra pa entendê(r)... tipos de estruturas que são formadas éh:: com mistura de compostos... diferentes?....* (AC 083/RP/ L 254-255)
- (41) (...) *lá tem umas plantas né? umas árvores que:: é muito::... tradicional lá na cidade... é o plátano...* (AC 094/DE/L 118-119)

O sujeito é analisado, nesses dados, como um *SN pleno simples* ou *SN pleno composto*. Outro critério para o pronome relativo ser considerado apenas material interveniente é o fato de não influenciar a marcação de plural, pois é o SN sujeito, que governa a presença ou não de marcas formais no verbo e no predicativo, conforme mostram os exemplos contidos em (42), (43) e (44):

- (42) (...) e passô(u) um período bastante difícil... como todo **os outros professores...** que:: foram afastados pela Faculdade de Filosofia (AC 146/DE/ L 205-206)
- (43) (...) a memória coletiva... porque há **alguns fatos** que são marcantes na vida de uma comunidade (AC 146/RO/ L 404-405)
- (44) (...) e como **os medicamento** que ele tomava... que era muito forte... ele não dormia... (AC 132/NE/ L 28-29)

2.5 As variáveis extralinguísticas

A hipótese que subjaz à seleção de *gênero* como fator de análise tem a ver com a constatação, já plenamente reconhecida na literatura variacionista, de serem as mulheres mais sensíveis que os homens à atribuição de prestígio, de acordo com os estudos de Scherre (1994, 1996), o que implica em apresentarem frequência menor que os homens para os casos de ausência explícita de marcação de plural. Scherre (1996) mostra que o grupo feminino se utiliza mais frequentemente que o masculino da forma padrão no que tange o processo de concordância nominal.

O princípio geral é o de que as diferenças meramente biológicas, ou de sexo, devem ser tratadas como diferenças sociais, ou, mais apropriadamente, gênero, uma vez que a consistência entre uso do padrão e conservadorismo feminino emerge principalmente de estudos em comunidades de fala ocidentais, que compartilham valores estabelecidos pela natureza das relações sócio-culturais.

Já o grupo de fatores *idade* tem vínculo estreito com processos de variação submetidos à mudança. Naro (2003) afirma que as mudanças linguísticas processam-se gradualmente e em várias dimensões; na dimensão social, por exemplo, os falantes mais velhos preservam formas conservadoras, o que pode ocorrer com camadas sociais de maior prestígio, ou de agrupamentos sociais que exercem atividades que exigem boa apresentação pública.

O eixo das idades define uma escala em ‘tempo aparente’, obtida com base no estudo de falantes de diferentes faixas etárias. A literatura sobre concordância nominal e verbal mostra que essas duas regras coincidem com processos variáveis, mas estáveis; nesse caso, podem exercer a função de identificação social dos falantes de diversas variedades, caracterizando fronteiras entre diferentes tipos de comunidades sociais. É necessário verificar, portanto, se essa hipótese se confirma ou não nos dados do Iboruna.

Quanto ao grupo de fatores *escolaridade*, sua relevância se justifica plenamente no fato de que a escola é capaz de produzir mudanças na fala e na escrita das pessoas que a frequentam, principalmente por atuar como principal instância de preservação das formas de prestígio diante das diversas tendências de mudança em curso na comunidade. Além dos processos conhecidos de atribuição de estigmatização e prestígio com base na variedade das pessoas que desfrutam de maior prestígio social e econômico, reconhece-se que a escola controla alguns fenômenos de mudança nos quais se incluem processos regulares de concordância nominal e verbal.

Além da análise dos fenômenos aqui enfocados de acordo com as três variáveis acima discutidas, é relevante estabelecer uma interação entre os grupos de fatores extralinguísticos, para ser possível mostrar aspectos relevantes da distribuição social das variantes examinadas, o que justifica plenamente a inclusão na análise. É, além disso, absolutamente pertinente examinar a relação das variáveis independentes de natureza externa com as de natureza interna ou linguística propriamente dita. O objetivo desse procedimento é o de avaliar, por um lado, se os diferentes agrupamentos sociais na comunidade riopretense são detentores de diferentes gramáticas na regência da concordância ou, se, por outro, a gramática é sempre a mesma, mas cada variante tem diferente frequência de uso conforme a distribuição social dos informantes investigados.

Outro motivo importante para se examinar a relação, por meio de cruzamentos, entre as variáveis internas e externas (estatisticamente significativas) é o fato de Scherre (1991) mostrar apenas o efeito isolado das variáveis externas. Como o princípio do Paralelismo Formal também se aplica aos dados do Iboruna, é importante verificar se há uma correlação social que evidencie maior frequência de marcação para o agrupamento de maior escolaridade e, vice-versa, maior frequência de não-marcação no agrupamento de menor escolaridade. Nossa hipótese é de que haja uma fronteira social entre paralelismo positivo e o paralelismo negativo, respectivamente, em que presença de marcas conduziria a presença de marcas no agrupamento mais escolarizado, enquanto ausência de marcas conduziria a ausência de marcas no agrupamento de menor escolaridade.

É necessário saber quais são os fatores externos e internos que levam os falantes a fazerem a concordância ou não dentro do predicativo, se há uma regra geral com base na qual a heterogeneidade linguística possa ser sistematizada. Outro ponto que se pode pensar nos estudos sobre presença e ausência de marcas formais de pluralidade nos SN, de forma geral, é que há uma redundância da marca morfológica de pluralidade, assinalada no português-padrão, que envolve toda a sentença. Esse traço é um espaço privilegiado para as motivações em competição, seja em contextos em que a informação deve ser recuperada, como os sujeitos nulos, seja em contexto em que a informação deve ser suprimida, como nos casos em que a informação é redundante quando assinalada em uma multiplicidade de formas.

2.6 Da quantificação e da análise dos dados

O processamento dos dados, nesta pesquisa, é eletrônico, empregando-se, para tal, o pacote estatístico *Goldvarb X* (TAGLIAMONTE *et. al.* 2001), idealizado pela sociolinguística quantitativa, para a análise de fenômenos variáveis. Esse pacote estatístico auxilia o

pesquisador a entender quais fatores linguísticos e extralinguísticos favorecem ou desfavorecem o uso de variantes linguísticas mediante índices quantitativos que extraem frequências e pesos relativos dos fatores linguísticos e sociais em relação à variável dependente.

As ocorrências foram selecionadas no cópús e codificadas, de acordo com códigos mnemônicos atribuídos a cada um dos fatores que constituem os contextos variáveis. A fase precedente à quantificação dos dados é também importante, porque pressupõe, de antemão, uma análise qualitativa dos dados, que permite, após os resultados frequenciais, compreender e explicar as estatísticas numéricas oferecidas pelo programa.

Como se trata de uma variável binária, o processamento quantitativo foi realizado mediante o uso do programa Ivarb, aplicável a variáveis binárias. Esse programa é suscetível a um processamento em duas dimensões, o “step down” e o “bottom up”, que permitem avaliar, com maior segurança estatística, o grau de significação quantitativa dos grupos de fatores, na medida da complementaridade. Por se tratar também de uma variável binária, os dados serão tão significativos quanto forem superiores a 0.5; abaixo desse limite, os pesos relativos apontam para uma relação pouco significativa estatisticamente entre a variável dependente e os grupos de fatores.

A análise qualitativa permite avaliar a marcação de pluralidade no predicativo da variedade riopretense e comparar os resultados com os de outras pesquisas. Além disso, a análise qualitativa fornece indícios seguros sobre a atuação dos processos alternativos de economia e redundância em busca de generalizações sobre a aplicação formal e funcional da variação linguística.

Após toda a discussão em relação aos grupos de fatores que operaram a análise, sobre o universo de pesquisa deste trabalho e os procedimentos metodológicos que foram utilizados, além da explicitação dos dois tipos de análises realizadas, quantitativa e qualitativa, passemos, agora, a análise geral dos resultados.

3. ANÁLISE DOS RESULTADOS

Esta seção descreve os resultados obtidos com a análise quantitativa da CN nos predicativos e participios passivos da variedade falada na região de São José do Rio Preto. A presente exposição pretende mostrar (e discutir) os dados quantitativos para cada um dos grupos de fatores, devidamente acompanhado de uma análise qualitativa.

Os resultados obtidos apresentam-se em tabelas e gráficos. As tabelas estão organizadas em frequências brutas de casos (Freq.), em médias percentuais (%) e em pesos relativos (PR). As médias percentuais constituem um dado significativo, mas são os pesos relativos que revelam o efeito que cada um dos fatores tem simultaneamente sobre o fenômeno linguístico analisado (a variável dependente). Esta seção apresenta também alguns resultados obtidos através do cruzamento de dois fatores relevantes para o estudo da concordância nominal nos predicativos.

Nas subseções seguintes, aparecem os resultados quantitativos específicos da marcação de plural em todos os grupos de fatores separados conforme o estatuto deles: linguísticos ou internos, extralinguísticos ou externos. Além disso, apresentam-se, primeiramente, os grupos de fatores com relevância estatística, ou seja, os selecionados pelo pacote estatístico Goldvarb e, posteriormente, os grupos de fatores não selecionados.

3.1 Análise geral dos resultados

Como já antecipado, a amostra constitui-se de um total de 526 ocorrências com estruturas predicativas e participios passivos, que revelam a aplicação da regra de CN, dentre os quais 58% (310/526) apresentam marcas de plural, enquanto 42% (216/526) não apresentam aplicação da regra de CN. Na tabela e no gráfico a seguir, observam-se a distribuição desses percentuais.

Tabela 1: Número de ocorrências analisadas e percentual de aplicação e não-aplicação da regra de concordância nominal nas estruturas predicativas e nos participios passivos.

Presença de CN	Ausência de CN	Total
58% (310/526)	42% (216/526)	100% (526)

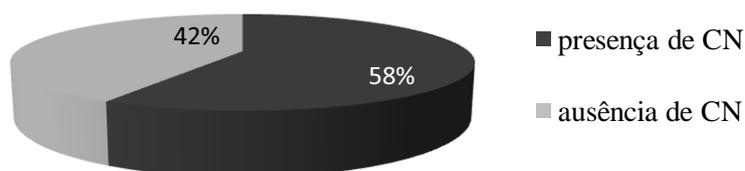


Gráfico 1: Percentual geral de presença/ausência de CN nos predicativos e participios passivos

O programa *Goldvarb* selecionou cinco fatores que, portanto, dispõem de maior relevância estatística para este trabalho, dentre os quais três fatores linguísticos e dois sociais. Os fatores linguísticos selecionados são: (i) número de marcas precedentes no sujeito e no verbo, (ii) paralelismo formal no discurso e (iii) características formais do verbo; os fatores sociais selecionados são: (i) escolaridade e (ii) idade. Os gráficos foram elaborados a partir dos pesos relativos de cada fator e se apresentam de acordo com a ordem de importância fornecida pela seleção do *Goldvarb X*. As tabelas e os gráficos são apresentados a partir do valor da aplicação da regra, ou seja, da presença de marcas formais nas estruturas predicativas. Passemos agora a análise dos fatores linguísticos com significância estatística.

3.2 Análise dos fatores linguísticos estatisticamente relevantes

Esta subseção trata da apresentação dos fatores internos estatisticamente relevantes, selecionados pelo pacote estatístico *Goldvarb X*. A primeira variável selecionada foi *número de marcas precedentes no sujeito e no verbo*. Esse grupo de fatores busca comprovar que, se

não houver marca formal de plural em todos os constituintes do SN e no verbo, provavelmente o predicativo também não é marcado. Essa variável tenta confirmar a hipótese da atuação do princípio da Economia e das Condições de Distintividade de Kiparsky (1972). Vejamos como ela se comporta para o fenômeno estudado, na tabela a seguir.

Tabela 2: Relação entre a variável *número de marcas precedentes no sujeito e no verbo* e a *marcação de pluralidade*

Fatores	Freq.	(%)	PR
Sujeito com duas ou mais marcas e verbo nenhuma	6/22	27,3	0.85
Sujeito e verbo com todas as marcas	164/209	78,5	0.64
Sujeito composto e verbo com marca	4/5	80	0.62
Sujeito não-explicito e verbo com marca	40/58	69	0.51
Sujeito pronominal e verbo com marca	76/135	56,3	0.39
Sujeito com uma marca e verbo nenhuma	1/37	2,7	0.35
Sujeito contendo pelo menos uma marca e zero verbal	14/20	70	0.20
Sujeito com uma marca e verbo com marca	4/20	20	0.14
Sujeito pronominal e verbo sem marca	1/16	6,2	0.02
Total	310/522	59.4	0.586(input)¹⁸

Significância: 0.000

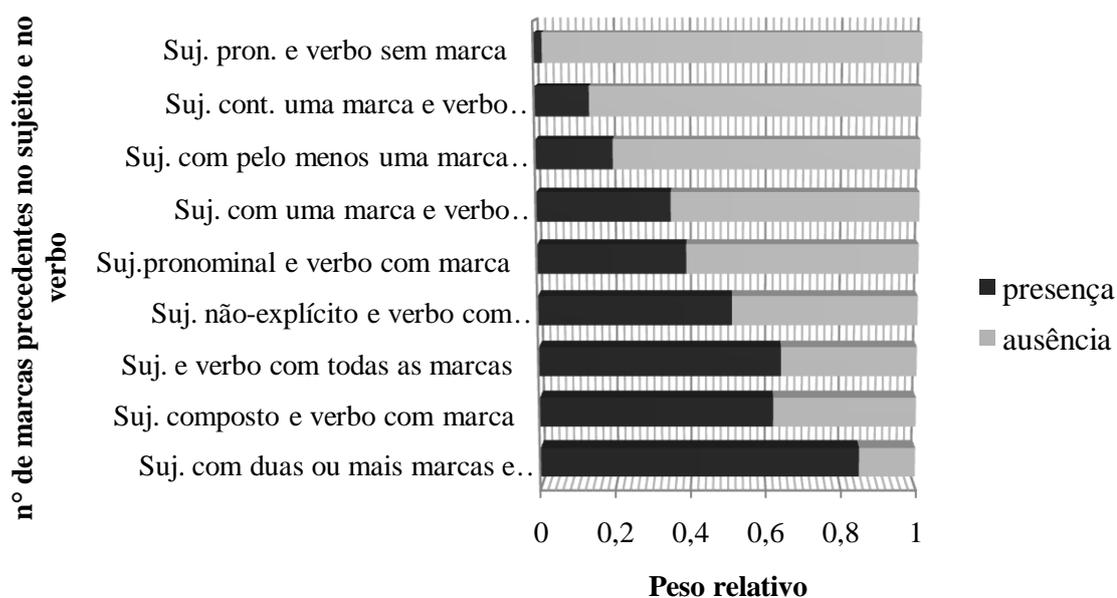


Gráfico 2: Peso relativo da relação entre *número de marcas precedentes no sujeito e no verbo* e a *marcação de pluralidade*

¹⁸ O input representa o valor total independente; e o grau de significância mostra que quanto mais próximo 0,0 for o índice, maior o grau de relevância estatístico dos dados.

Em relação ao grupo de fatores *Número de marcas precedentes no sujeito e no verbo* é possível observar que há indícios de motivações em competição, isto é, alguns fatores como *sujeito e verbo com todas as marcas (45a)*, com PR 0.64; *sujeito composto e verbo com marca (45b)*, com um PR 0.62, revelam a presença do princípio do Paralelismo Formal, pois, como se pode notar nos exemplos abaixo, uma marca formal condiciona outra marca formal, caracterizando um processo de marcação redundante.

- (45) a (...) *mas as casas ali são bem antigas... né?... (AC 120/DE/L. 266)*
 b (...) *dama valete e reis tam(b)ém num são tão fortes (AC 029/RP/ L 151)*

No entanto, os fatores *sujeito com uma marca e verbo nenhuma (46a)*, com PR 0.35; *sujeito com pelo menos uma marca e zero verbal (46b)*, com PR 0.20, *sujeito pronominal e verbo sem marca (46c)*, com PR 0.02, revelam a atuação do princípio funcional da Economia (HAIMAN, 1983), ou seja, a partir do momento que a sentença contenha pelo menos uma marca no sujeito, no nome, no pronome ou no determinante, o falante tem a intuição de que uma única marca já evita a ambiguidade referencial, não sendo mais necessário marcar o verbo e o predicativo. Como se pode ver nos exemplos abaixo:

- (46) a *teve uma época que a Globo tava passan(d)o aqui muito filme assim de sacanagem né?... éh em horário assim que as criança tava acordada (AC 104/RO/ L. 314)*
 b (...) *saíram pá fazê(r) uma boa pescaria... mas o... rio tava um po(u)co agita::do as onda um po(u)co alta e:: acho que eles... com po(u)ca experiên::cia... num determinado momento... o barco viROU. (AC 141/ NR/L. 41-43)*
 c *elas ficô(u) fechada dentro do cemitério (AC 105/NE/ L 143)*

Entretanto, chama a atenção, nesse grupo, o fator *sujeito com duas ou mais marcas e verbo nenhuma*, que apresenta um significativo PR (0.85) de aplicação da regra de concordância nas estruturas predicativas. A questão é que a sentença que se enquadra nesse

fator dispõe de duas ou mais marcas no sujeito, verbo sem marca formal e marca no predicativo, fato que não caracteriza nem o princípio do Paralelismo Formal nem o princípio da Economia (HAIMAN, 1983) ou das Condições de Distintividade apresentadas por Kiparsky (1972).

No entanto, apesar de haver poucos dados, a escolaridade pode ser importante para esse fator, pois 14 das 22 ocorrências foram emitidas por informantes com nível médio ou superior, razão pela qual os predicativos também apresentam marcas de pluralidade. Além disso, pode-se também explicar a falta de marcas de plural no verbo devido à baixa saliência fônica deles, como é possível observar nos exemplos contidos em (47a-c)

- (47) a *e éh::... normalmente alguns alunos dessa classe... éh **ficava incumbido** de í(r) éh:: regá(r) essa planta uma ou duas vezes por semana (AC 149/ L 235-237/ RP)*
- b *os pés de alface **fica enormes** na época do frio... agora dá alface mas os/... pés são bem **menores**... (AC 114/ L 595-596/ RP)*
- c *(...) as coisa **tava praticamente resolvida** as pessoas... tava sendo **socorrida** né? (AC 103/ L 107-109/ NE)*

O baixo grau de saliência fônica da oposição singular x plural não é, contudo, fator unânime no desfavorecimento da regra, já que no caso, o verbo *ser* aparece sem flexão com predicativo marcado, embora seja justamente o tipo de verbo que apresenta alto grau de saliência fônica. O Princípio da Saliência Fônica, portanto, é pertinente, mas não é determinante. Desse modo, a variável externa *escolaridade* pode atuar de forma relevante nesse tipo de regra. Em função da pressão normativa que os informantes possam ter sofrido, é possível que procurem marcar todos os constituintes nominais da sentença, dando preferência à característica da regra de concordância nominal em português. A tendência por aplicar a regra padrão nos SNs, entretanto, também não impede a falta de marcação de plural, mesmo em pessoas com máximo de escolaridade.

Outro fator desse grupo que mostra um resultado pertinente é *sujeito com uma marca e verbo com marcas* (PR 0.14). O peso relativo indica uma forte tendência para a não-aplicação da regra. Esse fator não permite confirmar nem a influência do princípio do Paralelismo Formal nem o de Economia. A atuação do princípio do Paralelismo Formal favoreceria a presença das marcas de plural no predicativo, o que não ocorre. A atuação do princípio de Economia favoreceria a necessidade apenas de uma marca no SN-sujeito para não tornar a sentença ambígua. Vejamos alguns exemplos em (48a-d)

- (48) a *então você vai daqui pra lá... você olha a:: as comporta tão aberta... você vê cá(r) no rio...* (AC 132/ L 222-223/ DE)
- b *às vezes certas administração são ruim mais a empresa é boa.* (AC 139/ L 106-107/ NE)
- c *a pia de cinco metro de comprimento... as porta são VERde as janela da cozinha são verde...* (AC 125/ L 99-100/ DE)
- d *no começo as coisa são difícil... mas ho::je depois de muito tempo tal de quase quarenta ano de casal casa/ de casamento né?... e:: a gente:: tem uma vida boa graças a Deus... as filhas são muito boa meus gen::ro... tenho sete netos né?* (AC 127/ L 50-53/ NE)

O fator *sujeito com uma marca e verbo com marca* apresenta, na maioria dos casos, predicativos marcados pelo seu baixo grau de saliência fônica, como se pode observar nos exemplos (48a) e (48b). Há, com efeito, somente três casos de plurais irregulares nas estruturas predicativas, como em (48d). Além disso, se observa, nas ocorrências, um elevado grau de neutralização entre o núcleo do SN e o verbo, e a falta de marcas no predicativo, como em (48c).

Na análise dos fatores dessa variável, é possível refletir sobre o princípio defendido por Lucchesi (2008) de que, por apresentar maior coesão estrutural, o SN dispõe de mais marcas formais de plural que o predicativo.

Essa diferença pode ser explicada por uma coesão estrutural que se observa no interior do SN, sobretudo no plano da relação Especificador-Núcleo,

respectivamente ocupada por um determinante (artigo, demonstrativo, etc.) e um nome. Entretanto, quando o SN exibe uma estrutura do tipo Determinante + Nome + Modificador, o adjetivo, que está em adjunção à direita do nome como um modificador, participa da relação de concordância no mesmo nível que um predicativo ou um particípio passado em uma estrutura passiva. (LUCCHESI, 2008, p.153)

Entretanto, a verdade dessas afirmações só poderá ser confirmada com base nos cruzamentos dos grupos de fatores e de uma análise ligeira entre a diferença de marcação de pluralidade no SN e nas estruturas predicativas para ver se esse princípio de coesão estrutural se aplica também ao Iboruna. Essa análise se estende também ao sujeito e verbo contendo todas as marcas.

Já o resultado de *sujeito não-explicito e verbo marcado* (PR 0.51) não é o esperado, já que como a frequência mostra uma tendência de 69% para a aplicação da regra de concordância, esperava-se também que o peso relativo demonstrasse o favorecimento da aplicação da regra de CN, em razão de recuperabilidade de informação. Isso, no entanto, não ocorreu, e o fator revelou-se neutro para a CN nas estruturas predicativas.

Os dados da tabela 2 permitem observar que o Princípio do Paralelismo Formal e o Princípio Funcional da Economia competem entre si dentro da comunidade de fala investigada. Du Bois (1985) defende que para que duas motivações possam competir no mesmo âmbito funcional deve haver um “bem limitado”. O “bem limitado” é o domínio da forma linguística, em que forças motivadoras múltiplas não podem prevalecer completamente. No caso desta pesquisa, o “bem limitado” é a indicação de pluralidade que, pelo visto, está sujeita tanto às motivações derivadas de redundância quanto às derivadas de ambiguidade. Entretanto, apenas os cruzamentos dos fatores permitirá confirmar qual motivação é mais forte para os informantes do Iboruna.

O segundo grupo de fatores selecionado pelo pacote estatístico *Goldvarb* foi o *paralelismo formal no discurso*. A expectativa em torno dessa variável era a de que contextos

em que os predicativos precedentes fossem marcados com o plural favorecessem a aplicação da regra nas estruturas predicativas subsequentes, e o inverso também seria aplicável, isto é, predicativos antecidos de predicativos sem marcas formais de plural refreariam a aplicação da CN nos predicativos subsequentes. Vejamos os resultados da tabela 3:

Tabela 3: Relação entre a variável *paralelismo formal no discurso* e a marcação de pluralidade

Fatores	Freq.	(%)	PR
Predicativo/particípio precedido de predicativo/particípio com plural explícito	62/67	92,5	0.85
Predicativo/particípio passivo primeiro de uma série	55/87	63,2	0.56
Predicativo/particípio passivo isolado	187/327	57,2	0.46
Predicativo/particípio passivo precedido de predicativo/particípio passivo com plural zero	6/41	14,6	0.10
Total	310/522	59,4	0.586 (input)

Significância: 0.042

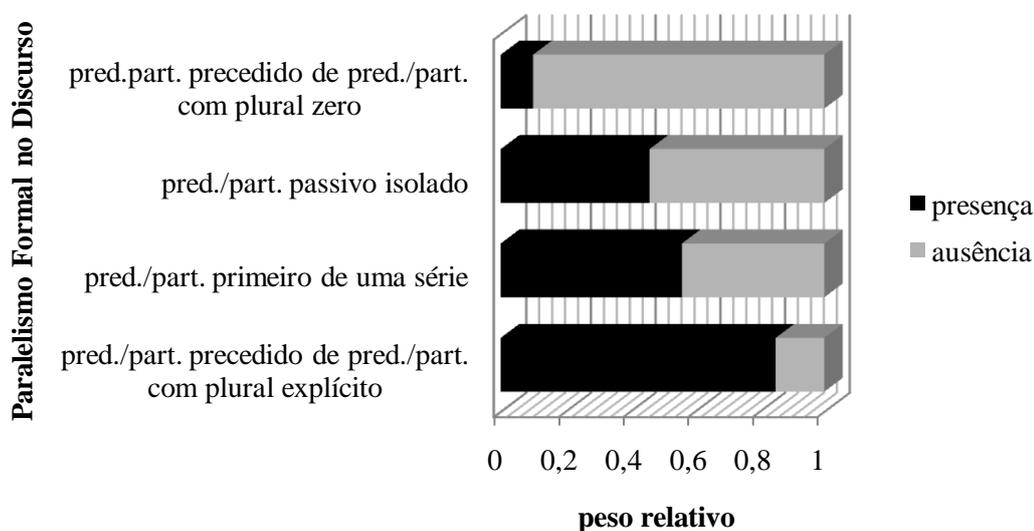


Gráfico 3: Peso Relativo da relação entre a variável *paralelismo formal no discurso* e marcação de plural

Observa-se que são o primeiro e o último fator da tabela 3 que revelam maior diferença entre presença e ausência de marcas explícitas de plural. O gráfico 3 mostra que, na série, quando um predicativo é precedido de outro com marcas de plural, ele, muito provavelmente, será marcado também (PR 0.85). No entanto, se o predicativo vier precedido

por outro sem marcas formais, ele terá grande probabilidade de ser não-marcado, como mostra o peso relativo de 0.10. Portanto, a aplicação da regra de CN é severamente condicionada por esse fator. É possível concluir daí que o plural explícito com marcas textuais precedentes motiva a pluralidade no predicativo seguinte e que a falta de pluralidade no predicativo precedente inibe a pluralidade no predicativo subsequente. Esses resultados foram encontrados também nos trabalhos de Dias (1996) e de Scherre (1991) com pesos relativos muito parecidos e indicado como fator estatisticamente relevante.

Quando o predicativo for o primeiro de uma série em outras construções com predicativo, há certa tendência para a marcação de pluralidade; como, no entanto, o peso relativo desse fator é próximo de 0.50, é neutra a probabilidade de aplicação da regra. O mesmo se aplica aos casos em que o predicativo estiver isolado no discurso: a tendência para a ausência de CN nas estruturas predicativas é visível no PR de 0.46, que, todavia, está próximo da neutralidade.

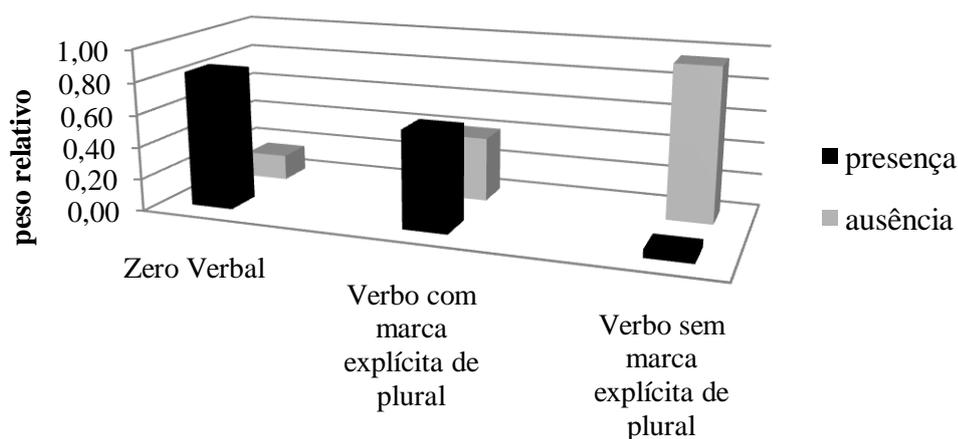
Esses resultados revelam haver uma forte influência do princípio do Paralelismo Formal, já que, por um lado, uma marca de pluralidade leva a outra marca e, por outro lado, a ausência de uma marca formal conduz também a não aplicação da regra (SCHERRE, 1991, 1994). No entanto, é necessário ressaltar que o grau de escolaridade pode estar relacionado a esses dados, reduzindo, se for o caso, o alcance desse princípio, principalmente porque escolaridade constitui uma variável independente selecionada pelo programa estatístico juntamente com as variáveis internas. Entretanto, para comprovar essa interferência é necessário verificar a atuação dos cruzamentos entre as variáveis linguísticas, com a variável escolaridade, o que será discutido mais à frente neste trabalho.

Passemos agora a verificar as *características formais do verbo*, a relação entre presença ou ausência de marcação de plural e a configuração flexional do verbo, como exposto na tabela 4.

Tabela 4: Relação entre a variável *características formais do verbo* e a marcação de pluralidade

Fatores	Freq.	(%)	PR
Zero Verbal	15/22	68,2	0.84
Verbo com marca explícita de plural	288/426	67,6	0.60
Verbo sem marca de plural	7/74	9,5	0.05
Total	310/522	59,4	0.586 (input)

Significância: 0.042



Características formais do verbo

Gráfico 4: Peso Relativo da relação entre às *características formais do verbo* e marcação de pluralidade

A tabela 4 mostra haver influência das marcas formais do verbo nas estruturas predicativas, isto é, o fato de o verbo conter marcas de plural motiva a probabilidade de elas aparecerem, também, no predicativo, como indica o PR de 0.60. O contrário também é verdadeiro: a ausência de marca no verbo é inversamente proporcional à aplicação de concordância no predicativo como se pode ver no PR 0.05. É necessário observar, no entanto, que o PR referente à presença de marcas no predicativo com verbo flexionado não é muito alto, fato que nos leva a pensar em outra motivação (formal ou funcional) agindo na CN nas estruturas predicativas. Fato interessante também é que, nos casos de zero verbal, a aplicação de regra atinge seu ápice, com um PR de 0.84.

A maior parte das ocorrências com verbo sem marca explícita de plural dispõe de uma ou duas marcas de plural no sujeito, verbo não flexionado com conseqüente ausência de

marcas explícitas de plural no predicativo. Parece, então, que, sendo a informação redundante, ela não precisa ser repetida. Além disso, evita-se ambiguidade referencial, uma vez que a pluralidade fica assegurada em pelo menos uma marcação, como se pode ver nos exemplos contidos em (49a-b).

- (49) a (...) *minha vó num queria o namoro porque na época é as coisa era muito **rígida**...* (AC 034/NR/L.42)
- b (...) *porque:: naquela época num era que nem hoje que os bastardo era **reconhecido** cê tinha que brigar muito* (AC 100/NR/L.119)

Há a possibilidade de estabelecer uma correlação entre frequência dos verbos com marcas e sem marcas de plural com o *sujeito explícito com todos os elementos flexionados* e com o *sujeito explícito sem as marcas formais explícitas* - fator não selecionado pelo pacote estatístico. Essa correlação se estabelece da seguinte maneira: quando um sujeito dispõe de todas as posições flexionáveis marcadas, normalmente, o verbo dessa oração também é flexionado.

Observa-se, portanto, que tanto verbo com marca explícita de plural quanto o sujeito com todos os elementos flexionáveis favorecem a presença de plural no SN predicativo. Da mesma forma, verbo sem marca de plural e sujeito explícito sem as últimas marcas formais acionam ausência de plural SN predicativo, como se pode ver nos exemplos contidos em (50a-b).

- (50) a *Os meus dentes estavam todos tortos* (AC55/NE/L.55)
- b *As moça aqui em Mirassol ficou tudo abismada* (AC129/NE/L.71)

Vale a pena salientar que o princípio de Paralelismo Formal, que se deduz da comparação entre os fatores, não tem uma aplicação categórica, isto é, não é aplicável a toda comunidade de fala, mas representa apenas uma tendência estatisticamente significativa em

favor da aplicação positiva de concordância no predicativo. Na variedade em análise, não se trata de regra unânime que permita postular uma correlação positiva entre o sujeito com marcas e verbo também flexionado. O que se deve reter aqui é que a maior parte dos casos da variedade estudada revela que a presença de marcas no sujeito aciona a pluralidade no verbo; similarmente, quando o sujeito não dispuser de marcas formais, é frequente o uso de verbos também não-flexionados. Essa correlação entre as marcas do sujeito e do verbo acaba por provocar efeito paralelo no predicativo.

Observa-se, finalmente, que a ausência de verbos também pode ser associada com a presença de marcas de plural, como comprova o peso relativo significativo de 0.84 para a aplicação da regra. A posição em justaposição do sujeito e do predicativo parece reforçar a marcação de natureza estilisticamente formal da construção, fato que deve favorecer a marcação de pluralidade. No entanto, é necessário analisar qualitativamente estes dados, porque a variável escolaridade pode influenciar diretamente os resultados do fator zero verbal. Dessa maneira, afirmar que a presença de marcas no predicativo, quando não houver verbo na sentença, seja apenas o resultado da atuação do Paralelismo Formal pode não ser ainda a resposta mais apropriada.

3.3 Análise dos fatores linguísticos não selecionados

Esta subseção trata da apresentação dos fatores internos não selecionados pelo *Goldvarb* (TAGLIAMONTE *et al.*, 2001) como estatisticamente significativos. Apesar de terem menor importância estatística do que os fatores discutidos anteriormente, vale a pena examinar o comportamento da variável dependente em relação a eles, já que alguns grupos de fatores normalmente selecionados, nos estudos sociolinguísticos, não se mostram

estatisticamente relevantes neste trabalho. Vejamos inicialmente como se dá essa relação com a variável independente *material interveniente*, exposta na tabela 5

Tabela 5: Relação entre o *material interveniente* e a marcação de pluralidade

Fatores	Freq.	(%)
Presença de material interveniente	115/185	62.2
Ausência de material interveniente	195/337	57.9
Total	310/522	59.4

O grupo de fatores *material interveniente* pretende confirmar a hipótese de que a presença ou a ausência de material interveniente entre verbo e predicativo pode afetar a aplicação de marcas formais de plural no predicativo. No entanto, como descrito na seção anterior, nas orações cujo sujeito é o próprio pronome relativo, este não foi analisado como um tipo estrutural de sujeito, mas como um tipo de material interveniente entre o sujeito e o verbo, já que todas essas ocorrências preservam o SN sujeito que o pronome relativo representa estruturalmente; essas construções estão mais próximas das construções simples com material interveniente.

Esperava-se, neste grupo, uma situação inversamente proporcional, ou seja, a presença de material desfavoreceria a marcação de pluralidade no predicativo, enquanto a ausência favoreceria a presença de marca formal. No entanto, não é esse quadro que se pode vislumbrar nos dados da tabela 5. Vejamos alguns exemplos com presença de material interveniente:

- (51) a. *“ah é muito bom é muito maravilhoso vou experimentá(r)” ... e as consequências depois não são **muito** agradáveis... pra eles (AC 056/RO/L 344-345)*
- b. *porque as minhas crianças eles e/ eles são **totalmente** inocentes quando eles/ eu recebo eles... Ele não... já sabia falá(r) de tudo quanto é tipo de palavrão (AC 086/RO/L 757-759)*
- c. *o aluno no particular é uma coisa no estado é outra... muda um pouco... porque eles tão/ muitos né? são **mais bem** alimentados eu acho que eles têm uma... rapidez de raciocínio (AC 116/ L 275-276/ RO)*

A presença de *material interveniente* não afeta necessariamente a presença de marca flexional de plural no predicativo. É necessário observar, no entanto, que a diferença estatística entre presença e ausência de material interveniente é muito reduzida, é justamente devido a essa proximidade estatística que esse grupo não foi selecionado como pertinente.

Outra variável independente não-selecionada se refere à influência sobre a marcação de plural no predicativo das marcas formais no sujeito. Vejamos como se comporta a variável *características formais do sujeito*, na tabela a seguir.

Tabela 6: Relação entre a variável características formais do sujeito e a marcação de pluralidade

Fatores	Freq.	(%)
Sujeito não-explicito (desinencial e nulo)	39/56	69,6
Sujeito explícito com marca formal de plural totalmente neutralizada	30/45	66,7
Sujeito explícito: com todos os elementos nominais flexionáveis marcados; com os últimos elementos com marca formal de plural; ou com a última marca neutralizada por contexto fonológico seguinte	219/330	66,4
Sujeito explícito com marca semântica de plural (nós e numerais isolados)	12/33	36,4
Sujeito explícito sem a(s) última(s) marca(s) formal(is) explícitas de plural ou sujeito coordenado singular	10/58	17,2
Total	310/522	59,4

Os resultados indicam que, para esse fator, a hipótese do paralelismo formal, defendida por Scherre (1991), se aplica aos dados do Iboruna. Com efeito, se o sujeito contiver todas as marcas formais de plural, é alta a probabilidade de marcação positiva no predicativo, como aponta a frequência 66,4 %, que favorece a marcação. Esse resultado também pode ser observado na relação entre sujeitos explícitos com marca formal de plural totalmente neutralizada e marcação de pluralidade no predicativo, que dispõe de uma frequência percentual de 66,7 %. Como as marcas estão neutralizadas, principalmente por

causa da proximidade com o verbo *ser*, não é possível descrever, sem um aparato acústico adequado, se o informante emite ou não a marca de pluralidade. Desse modo, é como se o sintagma nominal apresentasse todas as marcas, numa análise simples, de oitiva.

No entanto, se o sujeito se apresentar predominantemente sem as últimas marcas formais de plural, a chance de o predicativo aparecer marcado decresce para 17,2%; nesse caso, esse fator tende a favorecer a ausência de marcação. Comportam-se, semelhantemente, os casos em que o sujeito explícito contiver marca semântica de plural, como pronomes de primeira pessoa do plural e numerais, que apresentam frequência de 36,4%; desse modo, sujeitos com pluralidade semântica tendem a ser inibidores de marca explícita de plural no predicativo.

É, por outro lado, curioso observar que a maior frequência (69,6%) se aplica aos casos em que o sujeito é não-explícito, isto é, desinencial ou nulo. Curioso, porque contraria o Princípio do Paralelismo Formal. Se esse princípio se aplicasse de modo constante e regular, seria de esperar que sujeitos zeros, no caso dos nulos, acionassem ausência de marcas no predicativo e, inclusive, até menos marcas que os casos submetidos a outros fatores. Todavia, a impressão obtida por esse comportamento dos informantes é a de que, na ausência de marca no sujeito, alguma pluralidade tem que ser marcada em algum outro lugar da estrutura da oração e esse lugar acaba sendo o predicativo, o que trabalha em favor da preservação de informação.

Todos os resultados aqui relacionados são exatamente paralelos aos encontrados por Scherre (1991) em relação aos dados da Amostra Censo do Rio de Janeiro, com algumas diferenças pouco significativas nas médias percentuais e nos pesos relativos. Essa autora menciona a possibilidade de haver duas motivações em competição no processamento da informação de plural: o paralelismo formal agindo nos casos em que marca leva à marca e ausência leva à ausência e um processo de recuperação de plural não-explícito no sujeito, que

é, nesse caso, zero. Se essas motivações estão em competição na variedade carioca, o mesmo é verdadeiro para os dados do Iboruna. Entretanto, não deixa de ser passível de alguma suspeição que alguma outra motivação de natureza externa possa motivar essa tendência em sentido contrário de marcação de plural em relação à presença e à ausência de marcas.

O grupo de fatores, a seguir, *tipo estrutural do sujeito*, foi controlado porque se julgava que algumas estruturas que ocupam a função de sujeito poderiam influenciar a aplicação ou não da regra de CN nos predicativos. Vejamos como se comporta a variável na tabela abaixo.

Tabela 7: Relação entre *tipo estrutural do sujeito* e a marcação de pluralidade

Fatores	Freq.	(%)
SN genérico (pleno nu)	12/13	92.3
SN pleno composto	5/6	83.3
Quantificador	18/26	69,2
SN pleno simples	165/276	59.8
Pronome pessoal	71/144	49.3
Total	273/467	58.5

A expectativa sobre essa variável era a de que quando o sujeito fosse um pronome pessoal, a tendência geral seria de que houvesse maior probabilidade de marcação de plural nos predicativos, devido à correlação existente entre o tipo de sujeito e o paralelismo formal. No entanto, o resultado foi completamente distinto, já que o fator *pronome pessoal* apresenta frequência equivalente a 49,3%, isto é, esse fator tende a não favorecer, nem inibir a aplicação da regra de concordância nominal em estruturas predicativas.

Além disso, esperava-se que o sujeito pleno composto apresentasse frequência mais baixa, pois haveria uma tendência para a ausência de marcas nos últimos elementos do SN pleno composto. Entretanto, o resultado foi contrário ao esperado, pois o SN pleno composto, na variedade investigada, tende a favorecer a aplicação de regras de CN nos predicativos e

particípios passivos. No entanto, como se trata de poucos dados, não se pode pensar em uma regra categórica, mas apenas numa tendência favorável para a marcação de pluralidade nos predicativos.

O resultado encontrado aqui é totalmente diferente dos resultados obtidos no estudo da concordância verbal na variedade falada em São José do Rio Preto. Rubio (2008) afirma que o SN pleno composto desencadeia baixa frequência de aplicação da regra de CV (38% de frequência de plural), resultado que, segundo o autor, pode ser explicado a partir do cruzamento desse fator com outros, como paralelismo formal de nível oracional e posição do sujeito em relação ao verbo. Ocioso dizer que essa variável é específica para os estudos da CV.

É necessário ressaltar, no entanto, que não há muitas ocorrências de SN compostos na análise da CN nos predicativos, e é talvez por essa razão que a diferença de frequências entre o estudo de CN e de CV seja tão distinta. Essas afirmações não passam, contudo, de suposições, já que não há número de ocorrências para uma relevância estatística sobre esse fator.

Com relação ao SN genérico, já era esperado que a sua frequência fosse alta, visto que se trata de um SN com um único elemento, de base nominal nu. Desse modo, o elemento do SN-sujeito deve ser marcado para evitar o problema da ambiguidade referencial.

Conforme já mencionado, o fator *quantificador* engloba também os pronomes indefinidos e os demonstrativos. Os pronomes indefinidos exercem uma função de quantificação indeterminada. Já os pronomes demonstrativos dispõem da mesma estrutura sintática dos demais, ou seja, os pronomes demonstrativos, indefinidos e quantificadores estão num SN sujeito sem núcleo nominal. Separadamente, os pronomes demonstrativos, indefinidos e os quantificadores não demonstram relevância estatística, devido à escassez de dados. Entretanto, o amalgamento deles num único fator passa a dar-lhes maior importância

estatística, que revela uma tendência para o favorecimento da aplicação da regra de concordância nominal nas estruturas predicativas.

Em relação ao SN pleno simples, esperava-se uma frequência de aplicação mais baixa, por conta de ser sua estrutura formada por determinante + núcleo. A tendência de o determinante dispor de marcas formais e de não ser núcleo marcado desencadearia uma menor incidência de aplicação da regra no predicativo. Entretanto, essa hipótese não se confirma, já que o SN pleno simples tende a favorecer a aplicação da regra de concordância, com uma frequência baixa, mas positiva de marcação de pluralidade (59,8%).

Uma variável independente que visa confirmar a atuação do princípio da Saliência Fônica (SCHERRE, 1989) postula que os itens mais salientes, isto é, com maior diferenciação de material fônico na relação singular/plural, favoreceriam a presença de marcas formais de pluralidade no predicativo. Desse modo, quanto maior a saliência fônica dos itens que compõem as estruturas predicativas, maior o favorecimento da presença de marcas explícitas de plural. Scherre dividiu a variável em três fatores: plural regular, plural não-regular e plural misto, enquanto Dias (1996) estabeleceu uma diferenciação entre plurais irregulares, separando os plurais irregulares terminados em <l, r, s, ão> e o plural duplo. Vejamos a tabela 8, que exhibe o resultado do grupo de fatores *saliência fônica*.

Tabela 8: Relação entre saliência fônica e marcação de pluralidade

Fatores	Freq.	(%)
Plural irregular em <r>:	8/9	88.9
Plural duplo	7/8	87.5
Plural regular de base proparoxítona	20/26	76.9
Monossílabos regulares	11/15	73.3
Plural irregular em <s>:	2/3	66.7
Plural irregular em <l>	17/26	65.4
Plural regular de base paroxítona	243/432	56.4
Plural em <ão>:	1/2	50

Plural regular de base oxítone	1/2	50
Total	310/522	59.4

Os resultados da tabela 8 mostram que os plurais irregulares em <r, s, l> e o plural duplo revelam uma tendência para aplicação da regra de concordância nominal em estruturas predicativas. No entanto, o número de dados é reduzido nesses fatores, por essa razão, eles não têm significação estatística.

Os resultados aqui obtidos são próximos dos de Dias (1996), em que o fator *saliência fônica* não foi selecionado pelo pacote estatístico *Varbrul* e não se mostrou relevante para a pesquisa. Contudo, os plurais irregulares confirmam a aplicação do princípio de *saliência fônica*. É digno de nota que plurais regulares de base proparoxítone favorecem a marcação de plural, resultado surpreendente em face dos apresentados por Scherre (1991). Embora baseados em apenas 26 dados, é surpreendente esse favorecimento de marcas porque, geralmente, os plurais das palavras proparoxítonas têm um grau menor de *saliência fônica*, em função da posição mais à esquerda da sílaba tônica.

A hipótese que se esperava confirmar é a de que os constituintes mais salientes dispusessem de maior incidência de marcas de plural do que os menos salientes. No entanto, é possível que a baixa representatividade dos índices tenha prejudicado um pouco esses resultados, que acabam não sendo conclusivos. Um exemplo claro disso é o do plural em <ão>, que tem certo grau de *saliência*; como, todavia, a incidência de casos é reduzida, o fenômeno acaba sendo irrelevante estatisticamente. Em suma, os resultados obtidos neste trabalho mostram que o princípio de *Saliência Fônica* pode ser considerado um fator não condicionador da presença de marcas formais de pluralidade em estruturas predicativas nos dados.

Em relação à variável *configuração estrutural do predicativo/particípio*, entende-se poder haver diferença estatística entre as classes dos predicativos propriamente ditos, que são, por definição, os adjetivos, e as demais classes. Note-se, além disso, que os *particípios*

passivos fazem parte de uma classe mais integrada com a cópula do que os demais, o que sugere relevância desse tipo de distinção formal. Vejamos a tabela a seguir.

Tabela 9: Relação entre configuração estrutural do predicativo/particípio e marcação de pluralidade

Fatores	Freq.	(%)
Particípio	133/218	61
Adjetivo	156/257	60.7
Substantivo	21/47	44.7
Total	310/522	59.4

A variável *configuração estrutural do predicativo*, apesar de não ter sido selecionada entre as mais importantes, mostra que predicativos representados por substantivos acionam menos marcas formais de plural; no entanto, os substantivos identificam a configuração estrutural menos recorrente no conjunto de dados em análise. Já adjetivos e particípios tendem a favorecer a aplicação da regra da CN nas estruturas predicativas. Esses resultados mostram que a natureza da categoria gramatical em função predicativa pode atuar como motivação para a explicitação ou não de marca formal de plural. Há realmente diferença de frequência entre substantivos, particípios e adjetivos, entretanto, as frequências de particípios e de adjetivos são praticamente as mesmas. Desse modo, não é possível analisar os particípios passivos como uma classe mais integrada à cópula, em função de não haver diferença estatística relevante para essa conclusão. O que se pode concluir é que adjetivos e particípios passivos estão mais próximos da função predicativa que substantivos, categoria que atua tanto na evocação de um Subato Referencial quanto na evocação de um Subato Atributivo (HENGEVELD; MACKENZIE, 2008).

3. 4 Análise das variáveis extralinguísticas

A marcação de pluralidade no SN predicativo e no SN não-predicativo é uma variável fortemente marcada por pressões normativas. A normatividade, que é um traço definidor das gramáticas tradicionais, pode ser entendida como “um ideal definido por juízos de valores e pela presença de um elemento consciente da parte das pessoas concernidas”. (ALEÓNG, 2001, *apud* BAGNO, 2003). A língua representada nas gramáticas tradicionais é considerada padrão para grande parte da população que constitui a elite. No entanto, aquilo que está representado na gramática normativa não é exatamente a realidade da língua falada em uso na comunidade, como já discutido na primeira seção deste trabalho. Os exemplos usados para configurar as regras nas gramáticas são comumente extraídos de textos literários muito distantes da realidade da língua falada pela população.

Esse modo de se ver a língua instaura um sentimento de preconceito linguístico, muito comum em relação à variedade não-padrão. Há estigmatização social dentro de uma determinada comunidade linguística, visto que as regras da gramática normativa sempre são formuladas por pessoas de alta classe social e geralmente os falantes das variedades não-padrão são as pessoas de classe social desprestigiada, com pequeno contato com a escola.

Por essa razão, é de grande importância analisar também o efeito das variáveis sociais, principalmente porque a diferença de escolaridade e de idade são fatores que podem condicionar fortemente a presença ou a ausência de plural no predicativo, como já mostramos em uma etapa inicial desta pesquisa por meio de um projeto-piloto. Já a diferença de gênero não se mostra muito significativa para o condicionamento da variável investigada. A análise dos dados submetidos à diferença de idade representa, conforme já mencionado, a necessidade de verificar se presença ou ausência de pluralidade é uma variação estável ou se

está sujeita à mudança, o que seria detectável na correlação possível entre o uso da variante inovadora e da variante de prestígio.

Inicia-se o percurso da análise pelos resultados das variáveis sociais *escolaridade* e *idade*, que foram selecionadas pelo pacote estatístico *Goldvarb* como as mais significativas estatisticamente. Apesar de gênero não ter sido caracterizada como uma variável relevante, discutiremos suas implicações para os resultados desta pesquisa. A tabela 10 mostra a relação entre *escolaridade* e marcação de pluralidade.

Tabela 10: Relação entre *escolaridade* e a marcação de pluralidade

Fatores	Freq.	(%)	PR
Ensino Superior	114/152	75	0.64
Ensino Médio	118/184	64,1	0.54
2° Ciclo do Ensino Fundamental	53/109	48,6	0.37
1° Ciclo do Ensino Fundamental	25/77	32,5	0.31
Total	310/522	59,4	0.586 (<i>input</i>)

Significância: 0.042

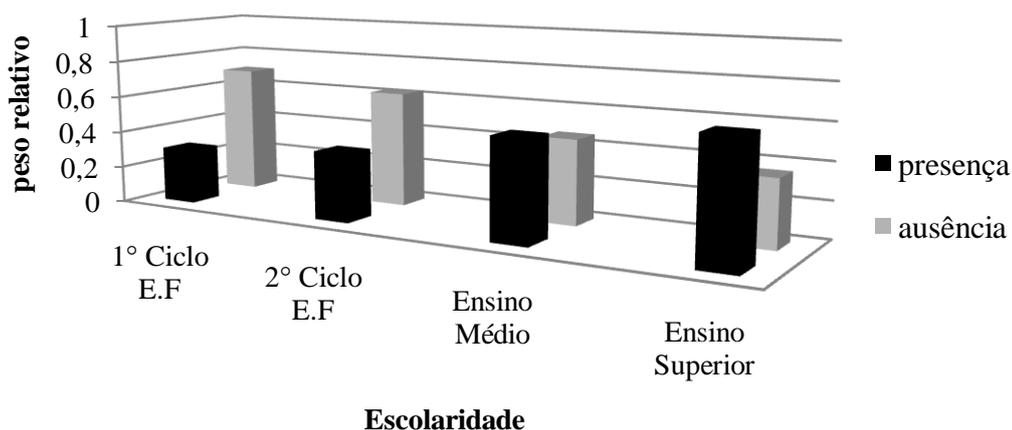


Gráfico 5: Peso relativo da relação entre *escolaridade* e a marcação de pluralidade

Os resultados da tabela 10 permitem concluir que a incidência de marcas no predicativo é diretamente proporcional ao acréscimo de grau de escolaridade. Além disso, o

gráfico permite verificar, também, que estar no primeiro e no segundo ciclo do ensino fundamental praticamente desfavorece o uso de marcas de plural. Há, com efeito, uma fronteira nítida entre informantes do ensino médio e superior e informantes do ensino fundamental de 1º e de 2º ciclos em termos de probabilidade de marcação positiva de plural.

A utilização das marcas nos ensinos médio e superior é maior certamente porque há um contato maior dos falantes com a atividade normativa imposta pelo sistema escolar que se aplica de forma rigorosa no processo pedagógico. Deve-se ressaltar que, embora a média percentual seja elevada para a presença de marcas, o PR de 0.64 para os informantes de nível superior mostra claramente que a regras de CN nos predicativos é variável mesmo para os que têm o mais alto grau de escolaridade.

É necessário refletir sobre a influência da variável *escolaridade* na aplicação ou não da regra de concordância nominal. O princípio que se estabeleceu para os estudos da CN e da CV, denominado Paralelismo Formal, defende que “marca leva à marca e zero leva a zero”, ou seja, o princípio estaria mais ligado aos fatores internos da língua e a concordância nominal passaria, então, a ser mecânica e determinada por mecanismos puramente estruturais. No entanto, nota-se que a escolaridade influencia diretamente a aplicação ou não da regra seja no predicativo ou no SN. Dessa forma, pode-se dizer que as motivações linguísticas competem com as extralinguísticas e o Paralelismo Formal, nos termos de Scherre (1988), pode ter uma atuação menos relevante. Estudos anteriores se debruçaram sobre a análise das variáveis sociais separadas, sem estabelecer o cruzamento delas com as variáveis linguísticas. Por essa razão, o Paralelismo Formal se fortalece como princípio geral explanatório nos estudos variacionistas. Nossa proposta é testar esse princípio, com as variáveis independentes de natureza externa, como *escolaridade* e *idade*, com as variáveis independentes de natureza interna, que foram selecionadas pelo pacote estatístico, para ver como o Paralelismo Formal

atua e se ele é a única explicação para o processo de variação investigado. Trataremos dessa relação mais a frente.

Com relação à variável independente *idade*, esperava-se que os falantes mais velhos apresentassem maior probabilidade de realizar a CN nas estruturas predicativas que os mais jovens, já que são menos propensos a aceitar mudanças no sistema linguístico (NARO, 2003). No entanto, não foi isso que aconteceu nas ocorrências analisadas, como se pode ver na tabela 11.

Tabela 31: Relação entre a idade do informante e a marcação de pluralidade

Fatores	Freq.	(%)	P.R
26 a 35 anos	70/111	63,1	0.59
16 a 25 anos	77/114	67.5	0.52
36 a 55 anos	73/144	50,7	0.37
Mais de 55 anos	90/153	58,8	0.53
Total	310/522	59,4	0.586 (input)

Significância: 0.042

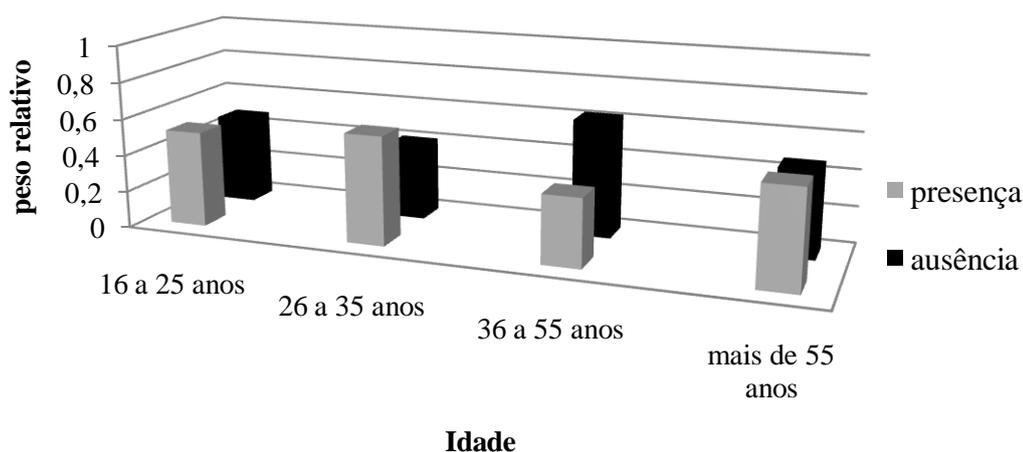


Gráfico 6: Peso relativo da relação entre *idade* e a marcação de pluralidade

A análise do gráfico 6 mostra não haver regularidade na aplicação da regra de plural nos predicativos ao longo das faixas etárias estratificadas como acontece com os níveis de escolaridade. Não é possível postular nenhum tipo de correlação entre o grau maior ou menor de maturidade biológica e maior ou menor frequência na aplicação da regra. O gráfico revela um crescente uso das marcas formais nas duas primeiras faixas etárias, com o grupo *de 26 a 35 anos* apresentando o maior peso relativo. No entanto, há uma queda considerável na aplicação positiva da regra na faixa etária seguinte e a CN no predicativo volta a elevar-se no grupo de *mais de 55 anos*, mostrando uma situação de equilíbrio com as duas primeiras faixas etárias em relação ao PR. Em vez da disparidade esperada, o que a diferença observada indica é mais similaridade de comportamento entre os sujeitos de diferentes faixas etárias.

O fato de a maior aplicação da regra incidir sobre os informantes de *26 a 35 anos* é plenamente justificável na variedade falada em São José do Rio Preto, já que é nessa faixa que as pessoas estão num rito de passagem de aluno de graduação a profissional no mercado de trabalho, o que também se aplica à faixa etária anterior. Inserção no mercado ocupacional pode ser uma explicação plausível para a aplicação positiva da regra pelos informantes dessa faixa etária.

A tabela 12, a seguir, exhibe a distribuição de dados entre a variável *gênero* e o uso de marcas formais de plural no predicativo. Apesar de não se caracterizar como uma variável estatisticamente relevante, ela também é importante para os estudos da CN, principalmente no que diz respeito à análise qualitativa dos dados.

Tabela 12: Relação entre a variável *gênero do informante* e a marcação de pluralidade

Fatores	Freq.	(%)
Feminino	187/286	65.4
Masculino	123/236	52.1
Total	310/522	59.4

O resultado obtido na variável *gênero* contraria um pouco nossas expectativas, uma vez que julgávamos que ela fosse selecionada como estatisticamente relevante, em vez de *idade*, de acordo com um estudo-piloto contido em Salomão (2007). No entanto, apesar de não ter sido selecionada, a variável mostra a mesma tendência que os estudos de Scherre (1991) e Dias (1996) e outros estudos variacionistas, já que as mulheres tendem a utilizar mais marcas formais que os homens. De acordo com Scherre, “as mulheres evidenciam que elas são mais sensíveis a atuação da escola do que os homens, no sentido de favorecer o uso da forma socialmente prestigiada”. (SCHERRE, 1996, p.254). A diferença de resultado é que normalmente, nos estudos citados anteriormente, os homens desfavorecem a marcação de pluralidade, seja no sintagma nominal, seja no predicativo. No entanto, neste trabalho, nota-se que eles se encontram numa zona mais neutra, isto é, se não a favorecem também não a desfavorecem.

3.5 Cruzamento entre fatores extralinguísticos

A análise das três variáveis externas permite concluir que os resultados deste trabalho apresentam uma configuração normalmente interpretada como um padrão de variação estável, detectável no eixo das idades, com base na distribuição por tempo aparente (LABOV, 1972).

Na análise das variáveis sociais é possível afirmar que as mulheres favorecem mais as formas de prestígio, devido a diversos fatores como a sensibilidade à atuação da escola, por exemplo. Outro fato importante é que a presença de formas de prestígio, como a marcação formal de pluralidade, é diretamente proporcional aos anos de escolaridade dos falantes. Além disso, a variável idade aponta um caráter curvilíneo, já que o efeito desse grupo de fatores não é sistemático em todas as faixas etárias.

Em função dessa relevância, efetuou-se um cruzamento entre as variáveis *idade* e *escolaridade*, para que se possa entender melhor a atuação delas como fatores condicionadores de presença de marcas formais de pluralidade, que aparece exposto na tabela 13.

Tabela 13: Cruzamento das variáveis *Idade* x *Escolaridade* e a marcação de pluralidade

Idade	de 16 a 25 anos			de 26 a 35 anos			de 36 a 55 anos			mais de 55 anos		
	F	%	PR	F	%	PR	F	%	PR	F	%	PR
1° ciclo EF	0/7	0	n ¹⁹	03/16	18	.12	43313	44	.33	14/36	38	.28
2° ciclo EF	15/27	55	.43	16/31	51	.39	43252	33	.23	16/33	48	.36
E. M.	36/51	70	.59	29/39	74	.64	27/57	47	.35	26/37	70	.57
E. S.	26/29	89	.84	22/25	88	.82	32/51	62	.51	34/47	72	.61

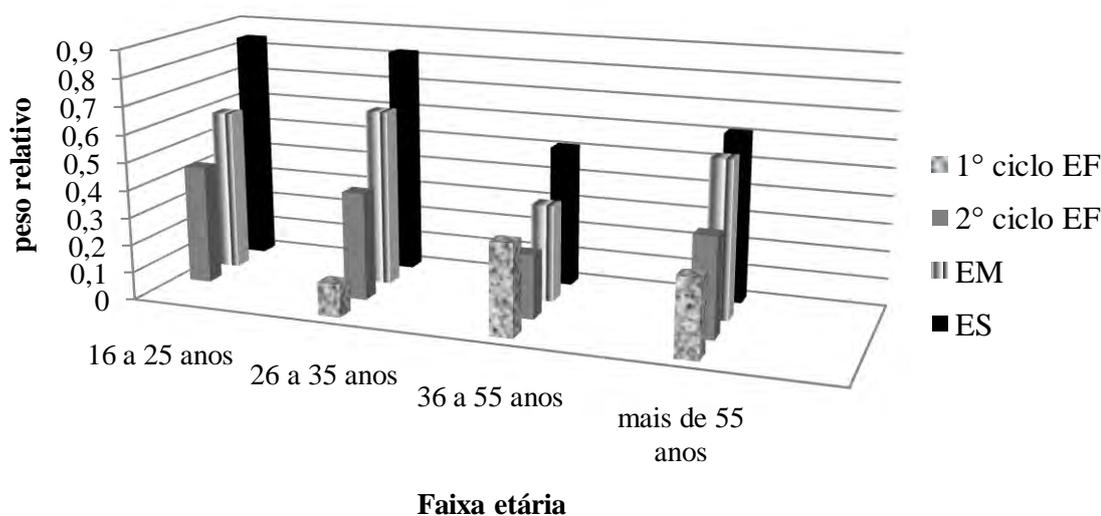


Gráfico 7: Peso Relativo do cruzamento das variáveis *Idade* e *Escolaridade* e a marcação de pluralidade

A tabela 13 confirma que a presença de marcas formais nos predicativos e participios passivos é reconhecida como a variante de prestígio na comunidade investigada. A região de São José do Rio Preto é altamente escolarizada e, dessa forma, a variante-padrão é utilizada com frequência entre os falantes que estão nas universidades ou que estão tentando se inserir no mercado de trabalho. A variedade-padrão pode representar, ainda, uma condição para a

¹⁹ Nas tabelas dos cruzamentos de fatores, internos ou externos, a letra **n** representa um Knockout.

ascensão social do indivíduo no mercado de trabalho. Percebe-se claramente uma linha divisória demarcando os ciclos do Ensino Fundamental e o Ensino Médio e Superior. Essa linha indica que os informantes dos dois ciclos do Ensino Fundamental, em todas as faixas etárias cruzadas, mostram uma tendência para a ausência de marcas explícitas de pluralidade nas estruturas predicativas. É exatamente oposto o comportamento dos informantes de EM e ES, que demonstram uma tendência crescente para a aplicação positiva da regra.

É importante salientar, porém, que os pesos relativos são mais elevados nos grupos de *16 a 25 anos* e *de 26 a 35 anos*, com uma queda de aplicação da regra no grupo de *36 a 55 anos*, que se mostra neutra no ES e com uma forte tendência para a ausência de marcação no EM. Mesmo assim, nota-se ainda uma tendência crescente para o uso de marcas formais à medida que a escolaridade aumenta. Já o grupo *com mais de 55 anos* revela um acréscimo na aplicação da regra de concordância. No entanto, esse acréscimo progressivo se mostra tão significativo quanto o das duas faixas etárias iniciais.

Esse resultado não é o esperado, pois, considerando ser a marcação positiva uma variante prestigiada, a última faixa etária deveria preservar marcas explícitas, diferentemente das faixas iniciais, que utilizariam com menor frequência a concordância nominal. Os resultados da tabela 13 encontram justificativas no alto índice de escolarização dos grupos de *16 a 25 anos* e *de 26 a 35*, visto compreenderem nessas faixas o período de realização do ensino superior e mesmo de pós-graduação, em alguns casos. Além disso, a faixa de *16 a 35 anos*, compreende também a entrada e a possível confirmação do indivíduo no mercado de trabalho. Por essa razão, devem mostrar sensibilidade muito maior à variante de prestígio, cultivadas pelo ensino formal, que constitui, então, a instituição mais preservadora da variedade padrão.

A linha divisória que é possível entrever na tabela 13 acaba por reforçar a correlação estável, diretamente proporcional entre o grau de escolaridade e o uso da variedade padrão.

3.6 Cruzamentos entre fatores linguísticos

Após a análise dos resultados das variáveis linguísticas, passaremos a discutir o cruzamento dos fatores linguísticos entre si, a fim de detectar o grau em que fatores internos exercem influência direta na marcação de pluralidade. Além disso, procuramos discutir se o princípio de Paralelismo Formal, como processo estritamente mecânico, predomina como princípio explanatório, ou se tem seu poder explanatório mitigado por variáveis extralinguísticas, como escolaridade, por exemplo. Vejamos o cruzamento entre *as características formais do verbo* e *paralelismo formal*, expostos na tabela e gráfico a seguir:

Tabela 14: Cruzamento das variáveis *Características formais do verbo* e *Paralelismo Formal*

Paral. sintagmático	Zero verbal			Verbo com marca de plural			Verbo sem marca de plural		
	Ap.	%	PR	Ap.	%	PR	Ap.	%	PR
Predicativo isolado	6/11	54	.42	177/266	66	.55	4/50	8	.05
Predicativo 1° de uma série	5/5	100	n	47/70	67	.55	3/12	25	.17
Pred.preced.de Pred.c/plural explícito	3/4	75	.64	49/62	95	.92	0/1	0	n
Pred.preced. de pred.c/ plural zero	1/2	50	.38	5/28	17	.11	0/11	0	n

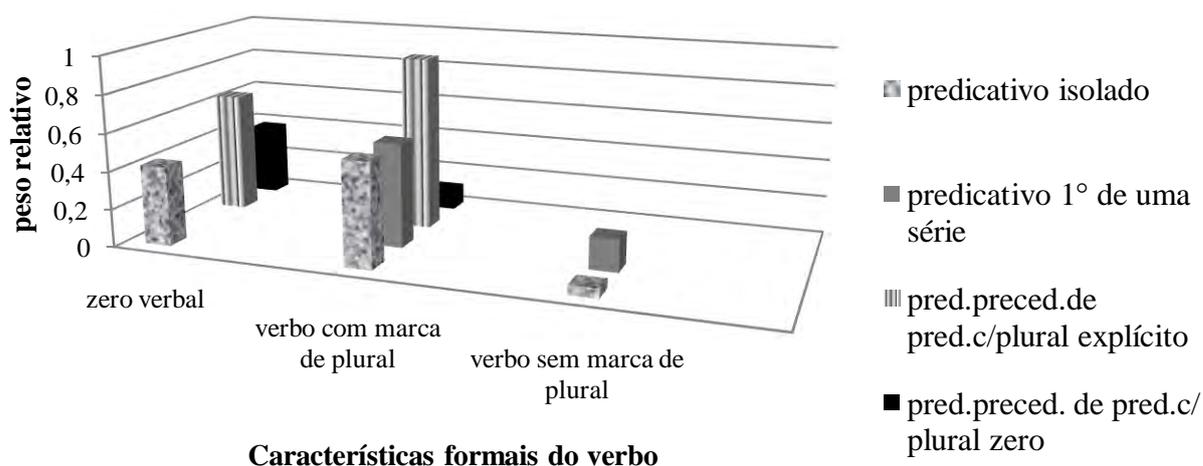


Gráfico 8: Peso relativo do cruzamento entre as variáveis *Características formais do verbo* e *Paralelismo Formal no Discurso*

Os resultados obtidos com base nesse cruzamento são muito interessantes para a discussão que se pretende empreender neste trabalho. Primeiramente, a análise isolada dos fatores *predicativo/particípio passivo precedido de predicativo/particípio passivo com plural explícito* e *predicativo/particípio passivo precedido de predicativo/particípio passivo com plural zero* mostra uma forte influência do Princípio do Paralelismo Formal, já que presença de marca explícita no predicativo precedente desencadeia marca explícita no predicativo subsequente e o inverso é verdadeiro também: a ausência de marcas no predicativo precedente aciona ausência no predicativo seguinte; a principal causa dessas relações está, em princípio, nos fatores mecânicos, como condicionamento fonético e a simples repetição de estruturas (LABOV, 1996).

Apesar de o Paralelismo Formal ser confirmado como princípio de alto poder explanatório nos estudos sobre variação do português brasileiro, é necessário considerar algumas questões. Uma delas é o grau em que este princípio se refere de fato à repetição de estruturas paralelas. Outro aspecto relevante é discutir a interação desse princípio interno com as variáveis externas ou sociais.

Aparentemente, na análise quantitativa, os dois fatores, *predicativo/particípio passivo precedido de predicativo/particípio com plural explícito* e *predicativo/particípio passivo precedido de predicativo/particípio passivo com plural zero*, cruzados com as *características formais do verbo*, principalmente para o fator *verbo com marca de plural*, mostram de fato o poder explanatório do princípio do Paralelismo Formal, nos termos de Scherre (1988), já que um /s/ tende a produzir um /s/ e um zero tende a produzir um zero, como se pode notar nos exemplos a seguir:

- (52) a *as paredes são todas **brancas**...(…). os bancos são **marrons**...* (AC 126/ DE/L 81-83)
- b *“as mesa tá tudo **limpinha** olha que is/... que is/ crianças organizadas” – qué dizê a mãe dá educação em casa... mas não aqui num é assim aqui ta/ as criança é tudo **desorganizada**...* (AC 068/RO/ L 187-189)

Operamos uma rodada estatística dessas ocorrências com a finalidade de verificar se Paralelismo Formal não sofre influência da variável externa *escolaridade* e encontramos alguns resultados relevantes, principalmente depois de estabelecermos uma disposição das ocorrências nas quatro faixas etárias do Banco de Dados Iboruna, como se pode ver na tabela 15.

Tabela 15: Distribuição dos predicativos com plural explícito nos quatro níveis de escolaridade

Ocorrências de predicativos precedidos de predicativos com plural explícito			
1º ciclo do E.F	2º ciclo do E. F.	E. M.	E. S.
7/67 dados	13/67 dados	22/67 dados	24/67 dados
10,44%	19,40%	32,83%	35,82%

Os resultados mostram haver um acréscimo progressivo no número de ocorrências e de frequências à medida que aumenta a escolaridade. Novamente há uma linha nitidamente demarcatória, isolando os dois ciclos do Ensino Fundamental do Ensino Médio e Superior. Dessa maneira, é possível conjecturar se o princípio do Paralelismo Formal se trataria tão somente de simples repetição de estrutura precedente, nos termos propostos por Labov

(1996). A escolaridade, na análise qualitativa, é a variável social que parece exercer maior grau de influência sobre o Paralelismo Formal. Por outro lado, é também possível aplicar o mesmo procedimento em relação à variável *gênero*, a despeito mesmo de seu menor grau de relevância. Os resultados mostram que a maior parte dos dados de aplicação *com predicativos precedidos de predicativos com plural explícito* é fornecida pelo grupo feminino, como se vê na tabela 16:

Tabela 4: Distribuição dos predicativos com plural explícito em relação ao gênero

Ocorrência de predicativos precedidos de predicativos com plural explícito	
Masculino	Feminino
21/67 dados 31,34%	46/67 dados 68,65%

Por outro lado, a análise qualitativa relativa ao fator *predicativo/particípio passivo precedido de predicativo/particípio passivo com plural zero* mostrou um resultado diferente ao do fator *plural explícito*, já que a escolaridade não é uma variável determinante para favorecer a ausência de marcas. Como se pode ver, pelos resultados da tabela 17, a maior parte dos *predicativos precedidos de predicativos com plural zero* encontra-se no 1º ciclo do ensino fundamental, como esperávamos. Esperava-se, também, uma pequena incidência de zeros nos dados fornecidos pelos grupos do Ensino Médio e Superior, mesmo quando a estrutura antecedente não fosse marcada. Vejamos a disposição na tabela 17:

Tabela 17: Distribuição dos predicativos com plural zero nos quatro níveis de escolaridade

Ocorrência de predicativos precedidos de predicativos com plural zero			
1º ciclo E. F.	2º ciclo E. F.	E. M.	E. S.
12/41 dados 29,26%	08/41 dados 19,50%	11/41 dados 26,82%	10/41 dados 24,39%

A variável *idade*, também selecionada pelo programa *Goldvarb*, se mostra mais produtiva, na análise qualitativa, já que os dois grupos das faixas etárias iniciais apresentam

um número menor de ocorrências de *predicativos/particípios passivos precedidos de predicativos/particípios passivos com plural zero*. Os informantes com 16 a 25 anos e com 26 a 35 anos revelaram na análise isolada, uma maior incidência de marcas explícitas de plural e a explicação dada, então, é a de que eles são marcados pela inserção no mercado de trabalho, além de representar o período de permanecerem na graduação e pós-graduação. A tabela, a seguir, revela a disposição das ocorrências de acordo com a faixa etária.

Tabela 18: Distribuição dos *predicativos com plural zero* em quatro faixas etárias do Iboruna

Ocorrência de predicativos precedidos de predicativos com plural zero							
de 16 a 25 anos		de 26 a 35 anos		de 36 a 55 anos		mais de 55 anos	
Apl.	%	Apl.	%	Apl.	%	Apl.	%
10/41	24,39	07/41	17,07	10/41	24,39	14/41	34,14
41,5% (total)				58,5% (total)			

A soma das ocorrências dos informantes das duas faixas etárias iniciais resulta numa frequência de 41,5 %. Esse índice revela uma tendência para um menor índice de aplicação do paralelismo formal nessas faixas iniciais. No entanto, a soma dos resultados das duas faixas etárias subsequentes, *de 36 a 55 anos* e *mais de 55 anos*, resulta justamente num efeito contrário, visto que a tendência esperada seria a de que um predicativo precedente sem marcas formais antecedentes favorecesse a ausência de marcas formais no predicativo subsequente dentro do discurso.

Há ainda outro aspecto relevante que chama a atenção na tabela 18. O grupo de *mais de 55 anos*, em sua análise quantitativa individual, mostra um peso relativo de (0.53), evidenciando uma tendência para a utilização das marcas formais de pluralidade. Essa variável indica a maior presença de *predicativos precedidos de predicativos com plural zero* no discurso. Vejamos alguns exemplos:

- (53) a *e a família pode brigá(r) pode fazê(r) o que quisé(r)... mas eles são assim eles são humilde... eles são humilde... acho que eles se vê que eles tão errado mas o organismo*

deles... pede isso que que vai fazê(r)? (AC 152/ L 556-558/ RO)

- b *a pia de cinco metro de comprimento... as porta são VERde as janela da cozinha são verde...* (AC 125/ L 99-100/ DE)

Podem-se explorar alguns fatos adicionais na interpretação qualitativa da disposição dos *predicativos/particípios passivos precedidos de predicativos/particípios passivos com plural zero* no grupo de *mais de 55 anos*. Primeiramente, é possível considerar a baixa saliência fônica das estruturas predicativas na oposição singular/plural. Como se observa, nos exemplos acima, trata-se de plural regular de base paroxítona na maior parte dos casos, um tipo de plural que não favorece a retenção de marcas, por causa do baixo grau de saliência. Outro ponto a destacar é haver, em quase todas as ocorrências, a presença de neutralização da marca formal entre o núcleo do sintagma nominal e o verbo. Por conseguinte, não é possível detectar se a marca foi realmente emitida. Além disso, grande parte das ocorrências dispõe de *verbo ser*, que tem o maior grau de saliência fônica, e é em razão disso que o verbo é mais comumente flexionado. Entretanto, o que parece fazer mais sentido é o fato de que das 14 ocorrências fornecidas pelos informantes *com mais de 55 anos*, oito delas (57%) são emitidas por indivíduos com grau fundamental de ensino, enquanto seis (43%) ocorrências estão associadas a informantes com nível médio e superior de escolaridade. Além disso, há nove dados (64%) fornecidos por informantes de *mais de 55 anos*, que também são do sexo *masculino*. É possível deduzir, portanto, uma confluência de variáveis sociais que acabam acarretando a ausência de marcas formais de plural nesta faixa etária.

Com relação ao fator *predicativo primeiro de uma série* e ao fator *predicativo/particípio isolado*, é possível afirmar que o cruzamento deles com o fator *características formais do verbo* demonstram favorecimento do uso de marcas formais de plural com o verbo marcado. Quando, por outro lado, o verbo não tem marcas explícitas, em

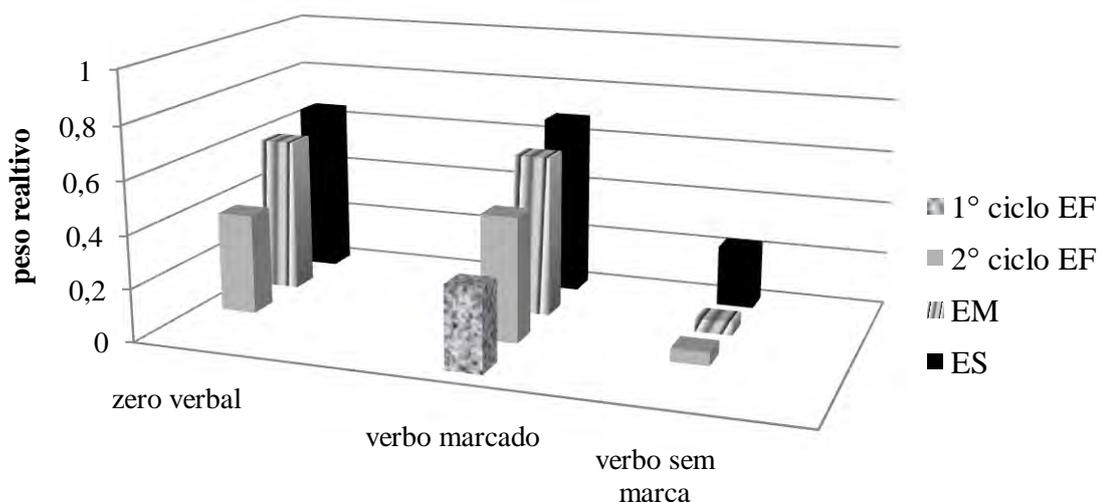
caso de *predicativo primeiro de uma série* e de *predicativo/particípio isolado*, acaba havendo a supressão das marcas formais de pluralidade.

3.7 Cruzamento entre fatores linguísticos e fatores extralinguísticos

Nesta subseção, procuraremos relacionar os fatores linguísticos aos extralinguísticos. É possível prever que, em determinadas estruturas linguísticas e em determinadas situações sociais, os falantes são propensos a colocar ou não todas as marcas formais de plural nos elementos flexionáveis. Isso ficou mais evidente após a checagem de todas as tabelas anteriores, com base na quais é possível chegar a algumas conclusões. Além disso, o cruzamento de fatores internos com externos tem a finalidade de verificar a possível influência de escolaridade sobre o princípio de Paralelismo Formal. Vejamos o cruzamento entre *características formais do verbo* e *escolaridade*, expostos na tabela e no gráfico abaixo.

Tabela 19: Cruzamento das variáveis Características formais do verbo e Escolaridade

Características formais do verbo	Zero verbal			Verbo com marca de plural			Verbo sem marca de plural		
	Ap.	%	PR	Ap.	%	PR	Ap.	%	PR
1º ciclo EF	2/2	100	n	23/54	42	.31	0/21	0	n
2º ciclo EF	3/6	50	.38	48/81	59	.47	2/22	9	.05
Ensino Médio	7/10	70	.59	109/152	71	.61	2/22	9	.05
Ensino Superior	3/4	75	.65	108/139	77	.68	3/12	25	.15



Características formais do verbo

Gráfico 9: Peso relativo do cruzamento entre as variáveis *Características formais do verbo* e *Escolaridade*

Os resultados acima confirmam a influência das variáveis externas sobre o princípio de Paralelismo Formal, pois mostram uma relação diretamente proporcional entre escolaridade e presença de marcas no verbo. Pode-se notar que, visto sob uma orientação vertical de leitura, os fatores *zero verbal* e *verbo com marca de plural* mostram uma linha divisória nítida, em relação aos ciclos do Ensino Fundamental e ao Ensino Médio e Superior. Os informantes de Ensino Fundamental apresentam um uso reduzido de marcas formais no predicativo e participios passivos, enquanto os de Ensino Médio e Superior apresentam comportamento inverso, com uma incidência elevada de marcação de plural nas estruturas predicativas. Os exemplos (54) e (55) mostram casos de predicativos antecidos de verbo com marca de plural empregados por falantes nos primeiros níveis de escolaridade (1º e 2º ciclos do E. F.) e nos ciclos finais (médio e superior), respectivamente.

(54) a *eu acho que hoje as meninas ficam **mocinha** muito depressa né?* (AC 096/RO/L 403)

b *(...) Eles que num num são **realista** num são **verdade(i)ro...** mas são **muito falso*** (AC 029/RO/L 185)

- (55) a (...) a sala de aula de terce(i)ra e quarta séries são salas onde não existe alfabeto... porque os alunos já estão **alfabetizados**... (AC 88/DE/L 335-336)
- b (...) que:: na época é a:: vaga era pra uma pessoa... mas das três pessoas... duas foram consideradas **aptas** (AC 114/NE L 55-56)

O fator *verbo sem marca de plural* revela um resultado um pouco diferente. Era de esperar que o cruzamento de *verbo sem marca* e *ensino superior* conduzisse a um peso relativo inferior ao do ensino médio ou pelo menos compatível com ele. No entanto, é preciso ponderar alguns aspectos adicionais. Primeiramente, há um número reduzido de ocorrências nesse cruzamento, fato que pode mascarar o resultado quantitativo. Em segundo lugar, é necessário considerar a relação do fator *verbo sem marca*, quando disposto nas quatro faixas etárias do Iboruna, a fim de verificar se maior ou menor idade também relativiza o efeito do princípio do Paralelismo Formal, como já vimos antes. Se isso ocorrer de fato, o PR obtido do Ensino Superior não passa de uma máscara do resultado real, causado pela influência determinística da análise quantitativa. Desse modo, a análise qualitativa revela-se mais uma vez de importância superior para análise do cruzamento *verbo sem marca* x *escolaridade*. Vejamos a tabela a seguir:

Tabela 20: Distribuição dos verbos sem marca de pluralidade do Ensino Superior em quatro faixas etárias

Ocorrência de verbos sem marcas de pluralidade no Ensino Superior							
de 16 a 25 anos		de 26 a 35 anos		de 36 a 55 anos		mais de 55 anos	
Apl.	%	Apl.	%	Apl.	%	Apl.	%
0/0	0	2/12	16,6	8/12	66,6	2/12	16,6

As ocorrências de verbos não-marcados no Ensino Superior, dispostos nas quatro faixas etárias do Iboruna, permitem identificar uma concentração majoritária delas na faixa *de 36 a 55 anos*. É justamente esse grupo etário que apresenta PR desfavorável ao uso de marcas de plural nos predicativos em sua análise individual (PR 0.37). Apesar de se tratar de poucos dados, esses resultados ajudam a comprovar que a variável social idade também pode exercer

influência direta na relativização do efeito do Paralelismo Formal. Essa influência é, neste caso, aparentemente, mais forte que a da escolaridade.

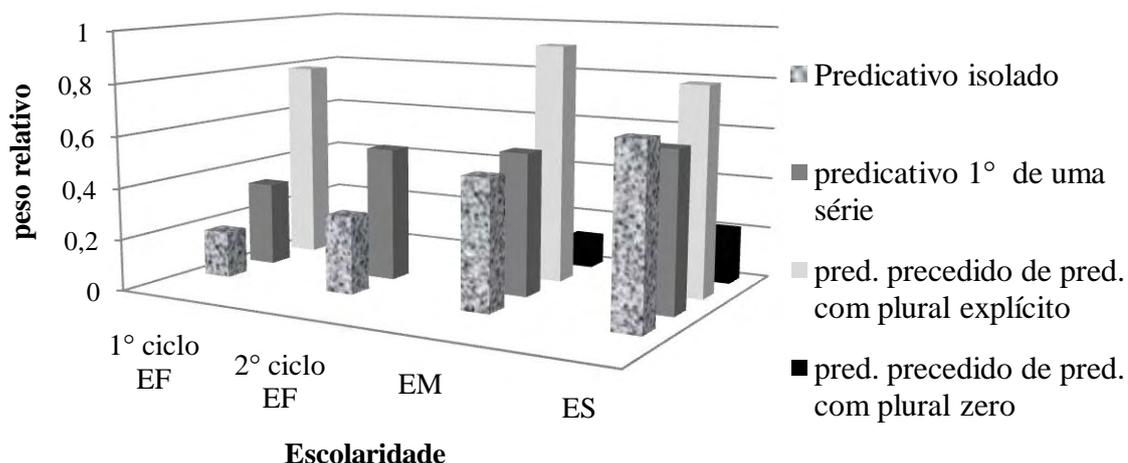
O mesmo caso acontece com o resultado do fator Ensino Médio, que também deveria apresentar peso relativo inferior se comparado aos resultados dos informantes do segundo ciclo do Ensino Fundamental, mostrados na tabela 19. No entanto, das 22 ocorrências de verbos sem marca no Ensino Médio, 11 delas estão na faixa etária de *36 a 55 anos*. Por essa razão, os PR mais elevados da relação entre o Ensino Médio e Superior são provocados pelo baixo número de ocorrências, dado estatístico que acaba causando o mascaramento dos resultados, já que é o grupo *de 36 a 55 anos*, o responsável pela baixa incidência de marcação de plural nos predicativos e participios.

Além disso, sob uma perspectiva horizontal, os dados revelam uma incidência significativa de marcas formais explícitas de plural nas estruturas predicativas em sentenças com verbos que também dispõem de marca de plural e uma significativa redução do uso das marcas de pluralidade nos predicativos cujos verbos são não-marcados. É possível concluir assim, que a variável *escolaridade* exerce influência direta sobre o Paralelismo Formal, na direção de uma relativização de seus efeitos; a variável *idade* também é relevante nesse processo.

Passemos a tratar, agora, dos resultados do cruzamento entre as variáveis *Paralelismo Formal* e *Escolaridade*, expostos na tabela 21. Conforme a orientação da análise qualitativa, esses resultados permitiram corroborar uma forte influência das variáveis sociais sobre a distribuição das marcas em paralelismo estrutural.

Tabela 21: Cruzamento entre as variáveis *Paralelismo Formal no discurso e Escolaridade*

Escolaridade Paral. Formal. no discurso	1º ciclo EF			2º ciclo EF			Ensino Médio			Ensino Superior		
	Ap.	%	PR	Ap.	%	PR	Ap.	%	PR	Ap.	%	PR
Predicativo isolado	11/42	26	.18	28/41	40	.30	76/123	61	.50	72/93	77	.68
Predicativo 1º de uma série	7/16	43	.33	12/19	63	.52	18/27	66	.55	18/25	72	.61
Pred. preced. pred. c/ plural explícito	6/7	85	.79	13/13	100	n	22/23	95	.93	21/24	87	.81
Pred. precedido de pred.c/ plural zero	1/12	8	.05	0/8	0	n	2/11	18	.12	3/10	30	.21

Gráfico 10: Peso relativo do cruzamento entre as variáveis *Paralelismo Formal no Discurso e Escolaridade*

A análise do cruzamento dos fatores *predicativo isolado* e *predicativo primeiro de uma série* mostra que os informantes dos ciclos do Ensino Fundamental fazem uso reduzido de marcas formais, enquanto os do Ensino Médio e Superior apresentam uma forte tendência para a marcação. Enfim, estar no Ensino Superior é condição favorável para o emprego da flexão de plural nesse tipo de estrutura. Um fato interessante a observar aqui é que, contrariamente aos informantes de outros níveis de escolaridade, os do Ensino Superior marcam mais quando o caso é de *predicativo isolado* do que quando é de *predicativo primeiro de uma série*. É interessante porque a estrutura frasal dos dois tipos de predicativos é

a mesma, isto é, constituída por SN, verbo e predicativo, como se vê nos exemplos (56a-b), com predicativos isolados, e (57a-b), com predicativos primeiro de uma série, respectivamente.

- (56) a *a sala de aula de terce(i)ra e quarta séries são salas onde não existe alfabeto... porque os alunos já estão **alfabetizados**...* (AC 88/DE/L 335-336)
- b *(...) e ela abandonô(u) ela foi embora não aguentô(u)... veja bem o quanto que... que esses filhos e esse marido foi **ruim** pra ela...* (AC 110/NR/L 162-163)
- (57) a *(...) ele tem::... os bancos dele são diferentes também... foram **trocados**... não são originais... pra ficá dá um aspecto mais esporte...* (AC 049/DE/L 132-133)
- b *minhas tia são muito **católica** em José Bonifácio mas são assim beata mesmo sabe?... se soubé(r) de um negócio desse ela me mata...* (AC 100/RO/L 380)

De uma perspectiva horizontal, em relação aos fatores *predicativo isolado* e *predicativo primeiro de uma série*, a análise da tabela 21 revela que, quanto mais a escolaridade aumenta, tanto maior é o uso de marcas formais nas estruturas predicativas.

Outro aspecto relevante são os resultados obtidos dos fatores *predicativo precedido de predicativo/particípio passivo com plural explícito* e o *predicativo precedido de predicativo/particípio passivo com plural zero*. Para fornecer uma análise mais profunda, realizamos uma separação de todas as ocorrências em que estes tipos de predicativos aparecem no córpus Iboruna para realizarmos uma análise qualitativa. Essa análise é importante porque, visto de uma perspectiva horizontal, os dados mostram um aumento gradual de marcação nos *predicativos precedidos de predicativos com plural explícito* nos ciclos do ensino fundamental, acréscimo que, julgávamos, se mantivesse entre os informantes de ensino médio e de ensino superior, o que de fato não acontece. Por outro lado, nos *predicativos precedidos de predicativos/particípios passivos precedidos de plural zero* ocorre situação semelhante, já que os informantes dos dois ciclos do Ensino Fundamental praticamente não marcam formalmente os predicativos, mas, nos informantes de nível Médio e Superior, nota-se um acréscimo dos pesos relativos e não a sua diminuição. Esses resultados

não comprometem a análise, pois é muito reduzida a incidência de marcas de pluralidade nos predicativos precedidos de predicativos com plural zero nos informantes com ensino médio e superior.

Como há uma diferença nos pesos relativos não esperada, decidimos fazer uma análise qualitativa mais refinada para ver se alguma variável social motiva esse resultado. A tabela 22 exhibe o cruzamento do fator *predicativos precedidos de predicativos com plural explícito* com o fator *ensino médio* dispostos nas quatro faixas etárias do Iboruna.

Tabela 22: Distribuição dos *predicativos precedidos de predicativos com plural explícito* do Ensino Médio em quatro faixas etárias do Iboruna

Ocorrência de predicativo preced. de pred. c/ plural explícito x ensino médio							
de 16 a 25 anos		de 26 a 35 anos		de 36 a 55 anos		mais de 55 anos	
Apl.	%	Apl.	%	Apl.	%	Apl.	%
8/23	34,78	6/23	26,08	3/23	13,04	6/23	26,08

A distribuição das ocorrências em relação à *idade* não é muito reveladora, já que segue a mesma tendência revelada pela análise individual dessa variável, ou seja, os informantes das duas primeiras faixas etárias marcam com maior frequência o plural formal; em sequência, há uma queda na frequência da terceira faixa etária e um acréscimo de marcação de pluralidade na última faixa etária. Interessante é o fato de a maior parte das ocorrências estar em faixas etárias com marcação positiva de pluralidade nas estruturas predicativas; desse modo, a variável *idade* não exerce influência direta sobre a diferença de PR dos predicativos precedidos de *predicativos/particípios passivos com plural explícito* no Ensino Médio.

Distribuímos os dados dos *predicativos precedidos por predicativos/particípios passivos com plural explícito* com os resultados do fator *ensino médio* em relação à variável *gênero do informante*, como se observa na tabela 23.

Tabela 23: Distribuição dos predicativos precedidos de predicativos com plural explícito do Ensino Médio em relação à variável gênero

Ocorrência de predicativo preced. de pred. c/ plural explícito x ensino médio			
Masculino		Feminino	
Aplicação	%	Aplicação	%
12/23	52,17	11/23	47,82

Apesar de a variável *gênero* ter-se mostrado relevante na análise qualitativa de outras variáveis, não é possível considerá-la aqui uma variável significativa, visto que é reduzida a diferença de ocorrências entre homens e mulheres. Além disso, as frequências também não apresentam uma diferença muito acentuada. Conquanto seja possível observar que o desempenho dos homens incide significativamente sobre a maior parte das ocorrências (12/23), esse resultado é praticamente neutro, considerando que a análise individual mostra um peso relativo do gênero masculino muito próximo a 0.50, com alguma tendência à marcação de pluralidade nos predicativos.

Passemos, agora, à análise da tabela 24, que mostra o cruzamento dos dados do fator *predicativo precedido de predicativo com plural explícito* com o fator *ensino superior* em relação à variável independente *idade*.

Tabela 24: Distribuição dos predicativos precedidos de predicativos com plural explícito do Ensino Superior em quatro faixas etárias do Iboruna

Ocorrência de predicativo preced. de pred. c/ plural explícito x ensino superior							
de 16 a 25 anos		de 26 a 35 anos		de 36 a 55 anos		mais de 55 anos	
Apl.	%	Apl.	%	Apl.	%	Apl.	%
7/24	29,16	3/24	12,5	6/24	25	8/24	33,33

Essa análise qualitativa revela uma disposição de dados diferente da revelada pelos informantes de Nível Médio. A tabela 24 mostra que os informantes *de 26 a 35 anos* apresentam a menor incidência relativa de ocorrências com marcas formais de plural nas estruturas predicativas, contrariando os resultados da análise individual da variável *idade*, já

que os informantes *de 26 a 35 anos*, nos resultados gerais da variável *idade*, eram os que demonstravam maior incidência de marcação de plural.

Diferentemente do cruzamento do fator *predicativos precedidos de predicativos/participios passivos com plural explícito* e a variável *idade*, que não se mostra muito revelador para a diferença de peso relativo em relação aos primeiros ciclos do Ensino Fundamental (tabela 21), observa-se que o cruzamento de *predicativos precedidos de predicativos com plural zero e ensino médio* distribuídos por faixa etária mostra-se qualitativamente relevante em termos estatísticos, contribuindo para o acréscimo de frequências no Ensino Médio em contraposição ao Fundamental. Vejamos essa disposição na tabela 25.

Tabela 25: Distribuição dos *predicativos precedidos de predicativos com plural zero* dos informantes do Ensino Médio em relação à variável *idade*

Ocorrência de predicativo preced. de pred. c/ plural zero x ensino médio							
de 16 a 25 anos		de 26 a 35 anos		de 36 a 55 anos		mais de 55 anos	
Apl.	%	Apl.	%	Apl.	%	Apl.	%
3/11	27,27	2/11	18,18	6/11	54,5	0/0	0

É possível considerar que esse acréscimo de *predicativos com plural zero* no Ensino Médio esteja mais relacionado à idade que à escolaridade. Na análise quantitativa geral, vimos que o cruzamento do fator *predicativo/participio passivo precedido de predicativo/participio passivo com plural* com o fator *Ensino Médio* aponta para um acréscimo de não-aplicação, no Ensino Médio e Superior. Esse resultado não era previsto, já que se estabelece uma correlação diretamente proporcional entre acréscimo de grau de escolaridade e de frequência na aplicação da regra.

O refinamento da análise quantitativa mostra, no entanto, um resultado revelador: a maior parte das ocorrências se concentra nos informantes *de 36 a 55 anos*, faixa que na análise da variável *idade* apresenta peso relativo negativo para a marcação de pluralidade (PR

0.37). O que se vê, portanto, é que é essa incidência de ocorrências nessa faixa etária que, afinal, contribui para a não utilização de marcas formais de plural no Ensino Médio.

Apesar de a variável *gênero* não ter sido considerada como estatisticamente significativa, nessa análise mais refinada, ela tem a função de esclarecer o acréscimo de incidência de predicativos com plural zero no Ensino Médio. Vejamos os resultados da tabela 26.

Tabela 26: Distribuição dos predicativos precedidos de predicativos com plural zero dos informantes do Ensino Médio em relação à variável *gênero*

Ocorrência de predicativo preced. de pred. c/ plural zero x ensino médio			
Masculino		Feminino	
Aplicação	%	Aplicação	%
7/11	63,6	4/11	36,6

Como se pode ver, a maior parte das ocorrências de não-aplicação da regra emerge do desempenho dos homens. Resultado aguardado, já que, de acordo com a literatura sociolinguística, são as mulheres mais sensíveis ao uso da flexão de plural e ao caráter normatizador da língua. Esse resultado revela-nos que gênero também tende a exercer uma influência negativa de marcação de pluralidade no cruzamento acima descrito.

A análise do cruzamento do fator *predicativo precedido de predicativo/particípio passivo com plural zero* com o fator *ensino superior* dispostos nas faixas etárias utilizadas nesta pesquisa, exibido na tabela 27, revela que o comportamento dos dados é distinto do cruzamento desse tipo de predicativo com o nível médio. Nesse caso, no nível superior, a maioria dos dados se concentra nos resultados de informantes mais velhos e de meia-idade.

Tabela 27: Distribuição dos predicativos precedidos de predicativos com plural zero dos informantes do Ensino Superior nas quatro faixas etárias do Banco de Dados Iboruna

Ocorrência de predicativo preced. de pred. c/ plural zero x ensino superior							
de 16 a 25 anos		de 26 a 35 anos		de 36 a 55 anos		mais de 55 anos	
Apl.	%	Apl.	%	Apl.	%	Apl.	%
1/10	10	0/0	0	3/10	30	6/10	60

Mesmo assim, as ocorrências se concentram numa faixa etária, cuja análise individual mostra um PR pouco acima de 0.50, o que significa uma tendência para a neutralidade. Outra variável que pode contribuir com o aumento de predicativos com plural zero no Ensino Superior é a variável *gênero do informante*, como se vê na tabela 28:

Tabela 28: Distribuição dos predicativos precedidos de predicativos com plural zero dos informantes do Ensino Superior em relação à variável gênero

Ocorrência de predicativo preced. De pred. c/ plural zero x ensino superior			
Masculino		Feminino	
Aplicação	%	Aplicação	%
6/10	60	4/10	40

Os dados mostram haver incidência majoritária nos informantes do sexo masculino. Apesar disso, não é possível fazer nenhuma generalização categórica definitiva com base num número tão reduzido de ocorrências. O que é possível fazer são conjecturas: (i) a tendência para a marcação de pluralidade nos predicativos é tão relevante para a variável *idade* quanto a variável *escolaridade* na análise quantitativa e principalmente na qualitativa; (ii) é possível confirmar a hipótese geral de que variáveis sociais exercem influência direta sobre o Princípio do Paralelismo Formal, nos termos de Labov (1994) e Scherre (1988).

Após a análise geral dos resultados, da discussão das variáveis e dos cruzamentos entre as variáveis independentes mais relevantes estatisticamente, abordaremos, na próxima seção, as implicações teóricas desses resultados, analisando mais especificamente as motivações formais e funcionais, foco central deste trabalho.

4. ALGUMAS GENERALIZAÇÕES E IMPLICAÇÕES TEÓRICAS

A análise de variáveis que cobrem a marcação de pluralidade nos predicativos e participios passivos serviu para avaliar o grau em que fatores propriamente internos, ou formais, interagem com os externos ou funcionais como verdadeiras motivações em competição (DU BOIS, 1995). Como a marcação de pluralidade nas estruturas predicativas representa nitidamente um processo de variação estável, a interação dessas motivações tem a função não de mostrar como elas podem favorecer a implementação de uma variante em detrimento da outra, mas de indicar como as variáveis dependentes se distribuem na comunidade como regras gramaticais da variedade dialetal considerada.

Poderíamos ainda nos juntar a Du Bois (1985) na indagação em termos de variação e mudança, de quanto do que ocorre na língua é fruto do trabalho social de seus usuários, a partir da informação cognitiva que eles retêm da estrutura da língua e da identidade social de que são dotados e quanto é derivado da organização estrutural do sistema linguístico, seja em termos de motivações funcionais ou formais.

A análise dessas questões se inicia na subseção 4.1, com a discussão dos efeitos das motivações externas. Em seguida, a subseção 4.2 trata do efeito das variáveis internas, principalmente focando a questão do princípio do Paralelismo Formal e da Saliência Fônica. Já a seção 4.3, finalmente, retoma as motivações em competição, trazendo as palavras finais desta pesquisa.

4.1 O efeito das motivações externas

Após o procedimento de análise operado na seção 3, cabe, agora, fazer um balanço final dos resultados alcançados para extrair deles algumas generalizações de poder explanatório. O primeiro tipo de motivações a ser discutida são as motivações externas, aqui

consideradas funcionais em larga escala, mas apenas quando cruzadas com as motivações internas. As três variáveis linguísticas estatisticamente significativas (*número de marcas precedentes no sujeito e no verbo, paralelismo formal no discurso e características formais do verbo*) mostram-se igualmente sensíveis a diferenças de escolaridade e idade e pouco sensíveis a diferenças de gênero em análise quantitativa. Entretanto, na análise qualitativa dos dados, a variável *gênero* se mostrou relevante na análise da disposição das ocorrências em diferentes faixas etárias.

As variáveis estão distribuídas em relação ao tempo aparente, representada por um padrão curvilíneo, indicando, desse modo, que a mudança em tempo real é pouco provável. A conclusão mais viável, nesse caso, é mais de uma tendência para a estabilidade que para a mudança. Bagno (2002) lembra que a implantação das teorias linguísticas nas instituições de Ensino Superior vem provocando significativas transformações na maneira de encarar o ensino das escolas de nível fundamental e médio. O resultado mais importante causado por essas transformações são as próprias diretrizes oficiais abrangidas pelos Parâmetros Curriculares Nacionais, doravante PCNs (MEC, 1998). Os PCNs trazem uma proposta de renovação do ensino de língua portuguesa no país, como se pode ver na citação abaixo.

A língua portuguesa, no Brasil, possui muitas variedades dialetais. Identificam-se geográfica e socialmente as pessoas pela forma como falam. Mas há muitos preconceitos decorrentes do valor social relativo que é atribuído aos diferentes modos de falar: é muito comum considerar as variedades linguísticas de menor prestígio como inferiores ou erradas. O problema do preconceito disseminado na sociedade em relação às falas dialetais deve ser enfrentado, na escola, como parte do objetivo educacional mais amplo de educação para o respeito das diferenças. Para isso, e também para poder ensinar Língua Portuguesa, a escola precisa livrar-se de alguns mitos: o de que existe uma única forma “certa” de falar – a que se parece com a escrita – e o de a escrita é o espelho da fala – e, sendo assim, seria preciso “consertar” a fala do aluno para evitar que ele escreva errado. Essas duas crenças produziram uma prática de mutilação cultural que, além de desvalorizar a fala do aluno, tratando sua comunidade como se fosse formada por incapazes, denota desconhecimento de que a escrita de uma língua não corresponde inteiramente a nenhum de seus dialetos, por mais prestígio que um deles tenha em um dado momento histórico. (PCN, séries iniciais do EF, p. 26 *apud* BAGNO, 2007, p.27)

Mesmo assim, o modo de se lidar com o ensino de fenômenos em variação ainda não se reflete, de modo integral, na prática pedagógica, visto que os professores ainda se consideram fortemente comprometidos com o ensino da tradição gramatical que, presa a formulações de base prescritiva, não se adapta às mudanças em progresso da língua. Bagno (2002) ressalta que a reação de muitos profissionais é a rejeição total aos novos parâmetros. Esse conservadorismo quase sempre se reflete na atribuição de estigmas e prestígios às variantes envolvidas, gerando julgamentos de valor que leva a discriminação dos usuários das formas não-padrão.

Conforme discutido, a marcação de pluralidade, seja no SN não-predicativo, seja nas estruturas predicativas, constitui fenômeno variável fortemente marcado por pressões normativas. A norma é um uso linguístico concreto, que corresponde à variedade social empregada pela classe de prestígio. É o que, segundo Castilho (1988), equivale à norma objetiva. No entanto, a norma pode também representar a atitude que o falante assume em face da norma objetiva, atitude essa que corresponde ao que a classe socialmente prestigiada “espera que as pessoas façam ou digam em determinadas situações” (RODRIGUES, 1968 *apud* CASTILHO, 1988, p.53). Portanto, a língua considerada correta pela classe de prestígio é aquela que está representada nas gramáticas pedagógicas, apesar de não ser a realidade da língua falada em uso na comunidade.

A distribuição dos resultados por grau de escolaridade revela que a distribuição dos índices cria um limite claramente demarcatório: de um lado, o ensino Superior e Médio e, de outros, os dois ciclos do Ensino Fundamental. Entretanto, há um crescimento gradual de marcas formais no predicativo, à medida que a escolaridade aumenta e que influencia, portanto, todos os níveis de escolaridade. Por essa razão, é possível afirmar que as regras gramaticais que regem o processo de variação são as mesmas para diferentes agrupamentos sociais de falantes.

A variável *idade* influencia a presença de marcas no predicativo, da mesma forma que a variável *escolaridade*. Por essa razão, diferentemente de outros trabalhos da sociolinguística, a variável independente *idade* se mostrou estatisticamente significativa na quantificação, sobrepujando em relevância a variável *gênero*. Conforme discutido, nas seções anteriores, as faixas de *16 a 25 anos* e de *26 a 35 anos* são as que mais se utilizam das marcas formais de pluralidade nos predicativos, devido ao alto grau de escolarização da região de São José do Rio Preto. São também as pessoas dessas faixas que ainda buscam auto-afirmação, em que a marcação de pluralidade é um indício de fala correta e prestigiada, gerando certo *status* linguístico, e até mesmo social, a quem pronuncie todas as marcas. Há queda no desempenho dos informantes da faixa seguinte, *de 36 a 55 anos*, o que pode representar uma fase de acomodação no mercado de trabalho ou até mesmo a proximidade da aposentadoria. Os informantes com mais de 55 anos apresentam um acréscimo de marcação, em relação à faixa anterior, que pode ser caracterizado pela própria pressão normativa que o falante sofreu em seu período de estudo.

4.2 O efeito das motivações internas

Em relação às motivações internas, a discussão se voltará para a atuação do princípio do Paralelismo Formal e para o princípio da Saliência Fônica. O primeiro é especificamente relacionado com a ocorrência ou não de marcas em posições precedentes ao predicativo em análise. Já o segundo diz respeito à diferença cognitiva derivada da diferença formal entre singular e plural no fenômeno variável investigado. No exame dessas motivações, interessa muito discutir a interação delas com outras motivações internas e motivações externas, de natureza social, que atuam, nesse caso, como motivações funcionais.

O princípio do Paralelismo Formal parece afetar as três variáveis internas, estatisticamente relevantes (com alguns fatores de maneira mais direta, outros nem tanto). O princípio é aqui analisado como marcas precedentes no nível oracional e marcas precedentes no nível do discurso. Segundo Camacho (2008)

A repetição de formas, considerada tanto na sucessão de construções passivas no discurso quanto na preservação da primeira posição estrutural para a alocação do mesmo referente em sentenças sucessivas, identifica-se com a tendência mecânica para a preservação de estruturas paralelas. Essa restrição contrapõe-se à distribuição de informação no discurso no estudo que Weiner e Labov (1977) dedicaram ao estudo das passivas não-agentivas do inglês, e ambas as restrições atuam claramente como motivações em competição com preponderância das restrições formais. (CAMACHO, 2008, p. 91-92)

O poder explanatório do princípio do Paralelismo Formal foi reiterado depois com o estatuto de verdadeiro universal variável postulado por Labov (1987, 1996) na consideração de casos já relatados na marcação de pluralidade no verbo e no SN do português brasileiro (SCHERRE, 1988) e do espanhol porto-riquenho (POPLACK, 1980). Labov (1996, p. 852) entende que todos os casos de repetição estrutural mostram uma tendência para que o falante continue repetindo o padrão fixado no início do SN de tal modo que um /s/ tende a produzir um /s/ e um zero tende a produzir um zero. O autor se baseia no trabalho de Poplack (1980) e se refere a essa tendência como perseverança ou concordância. Poplack (1980 *apud* LABOV, 1996) relaciona esta tendência concordante com um princípio de menor esforço no nível fonético antes proposto por Martinet (1961)²⁰.

O Paralelismo Formal, por muito tempo, foi absorvido como princípio teoricamente explanatório por Labov (1987; 1996) como forma de rejeição simultânea de princípios

²⁰ Vemos la tendencia a que el hablante continúe el patrón establecido al principio de la frase nominal: una /s/ tiende a producir una /s/, y un cero tiende a producir un cero. Podemos referirnos a esto como perseverancia o concordancia. Poplack relaciona esta tendencia concordante con un principio del menor esfuerzo en el nivel fonético articulado antes por Martinet (1961).

funcionalistas, avaliados como extremamente frágeis, como a necessidade de preservar informação. Segundo Labov (1996, p.840), há ampla evidência de que a variação morfológica e sintática é controlada por uma tendência da preservação de estruturas paralelas em orações sucessivas. Outras variantes resultam da avaliação social arbitrária de modos alternativos de dizer a mesma coisa, os subprodutos fortuitos do contato geográfico, sem função comunicativa.

Segundo Camacho (2008), Scherre (2001) se afasta da idéia simplista de Labov, defendendo a idéia de que o paralelismo tem motivação funcional, não por causa de sua função dentro do discurso, mas por causa de motivações externas a ele. Trata-se de um princípio de base cognitiva que possibilita ao ser humano fazer agrupamentos, formar blocos por similaridade formal e que encontra sua forma de manifestação maximizada quando atua em combinação com alguma outra função linguística, como a atribuição de prestígio social, a marcação de pluralidade dentre outras.

Camacho (2008) afirma, ainda, que, quando se evoca a participação humana, como faz Scherre (2001) ao postular um procedimento cognitivo na marcação variável de pluralidade, evoque também o princípio de que variação e mudança é o resultado do trabalho social dos usuários de uma língua a partir de informação cognitiva que eles retêm sobre a estrutura da linguagem e da identidade social específica de que são dotados, não o resultado de processos mecânicos derivados da organização interna, estrutural, do próprio sistema linguístico, que parece ganhar, assim, existência independente dos seus usuários.

4.2.1 O Princípio do Paralelismo Formal

Outro aspecto que cabe discutir mais detalhadamente são os resultados relativos ao Paralelismo Formal para a marcação de pluralidade nas estruturas predicativas. De maneira

geral, é possível afirmar que o Paralelismo Formal atua positivamente no condicionamento da marcação de plural no fenômeno estudado. Entretanto, ele não pode ser tratado apenas como uma simples repetição de estrutura precedente, um condicionamento fonético mecânico, nos termos de Labov (1996), na variedade estudada.

Conforme analisado na seção III, o princípio do Paralelismo Formal é totalmente influenciado por motivações externas, sociais, como as variáveis *escolaridade* e *idade*. Os resultados são muito claros: quanto maior a escolaridade, maior será o uso de marcas formais nos predicativos. Seria simplista demais afirmar que a presença de marcas nos predicativos seja apenas uma repetição mecânica.

Vale a pena ressaltar que os informantes da variedade estudada dispõem de alta incidência de escolarização, devido às boas condições econômicas da região de São José do Rio Preto. Dessa forma, os informantes, principalmente os mais escolarizados, sofrem pressão normativa muito forte das instituições de ensino, provocando um efetivo aumento no condicionamento de plural nos predicativos e participios passivos.

Os resultados discutidos, até o momento, são claros: informantes com uma identidade social específica, como o grau de escolaridade médio e superior inserido numa faixa etária de 26 a 35 anos, são os que lideram uma parte da aplicação do princípio do Paralelismo Formal – o que vai na direção do respeito às convenções normativas; caso em que marca leva à marca. Entretanto, informante com outro tipo de identidade social, como Ensino Fundamental e faixa etária de 36 a 55 anos, lideram outra ala da aplicação do princípio: a que vai em direção oposta às convenções normativas, caso em que ausência conduz a ausência de marcas.

O estudo da marcação de pluralidade em estruturas predicativos da variedade falada na região de São José do Rio Preto revela motivações formais e funcionais em competição. O Princípio de Economia (HAIMAN, 1983) e as Condições de Distintividade (KIPARSKY, 1972) aplicam-se aos dados do *cópus* Iboruna, mostrando-se mais decisivos na variável

número de marcas precedentes no sujeito e no verbo, primeira variável selecionada pelo *Goldvarb*. Apesar disso, não são capazes de explicar o processo de variação de pluralidade nos predicativos isoladamente. Além desses dois princípios funcionais, o que se nota é a atuação das variáveis sociais selecionadas, *escolaridade* e *idade*, como verdadeiras motivações funcionais quando cruzadas com as variáveis linguísticas. As variáveis sociais, nesta pesquisa, reduzem a influência no princípio do Paralelismo Formal. Por essa razão, podemos afirmar que a marcação de pluralidade nos predicativos, na comunidade de fala estudada, não pode ter explicação baseada apenas nas motivações formais ou apenas nas funcionais, pois as motivações competem entre si, nos termos de Du Bois (1985).

4.2.2 O princípio da Saliência Fônica

O princípio de Saliência Fônica era, em princípio, uma das motivações formais a ser investigada; no entanto, esse fator não se mostrou relevante estatisticamente na marcação de pluralidade nos predicativos. Mesmo assim, é possível observar que os plurais irregulares condicionam mais a presença de marcas formais que os regulares. Como a amostra dispõe de pouquíssimos casos com plural em - *ão*, um dos tipos mais salientes na CN, não foi possível chegar a nenhuma conclusão definitiva sobre a atuação desse fator.

O princípio Saliência Fônica, neste trabalho, pode mostrar apenas a atribuição de valores de estigma e prestígio social a variantes linguísticas, revelando também o papel das instituições, como a escola, na identificação da variedade padrão. Foi possível notar, ao longo da investigação, que a ausência de marca formal é socialmente estigmatizada e a atenção para esse atributo social ganha maior valor quanto maior for o grau de saliência. A variedade estudada, por um lado, mostra a realização de marcas explícitas com maior frequência entre os plurais irregulares, mais salientes; mas, por outro, percebe-se que os informantes em

relação aos plurais regulares, principalmente os paroxítonos, tendem a suprimir as marcas, em função de seu menor grau de saliência. Essa ausência de marcas formais, mesmo em predicativos menos salientes, pode acarretar o preconceito linguístico, já que a falta de CN é estigmatizada dentro da comunidade de fala.

O princípio de Saliência Fônica não pode ser considerado aqui como uma motivação em competição, como se suspeitou no início da pesquisa, já que não tem a mesma atuação relevante que o princípio de Paralelismo Formal. Desse modo, a Saliência Fônica é apenas um fator interno da língua que não se mostra estatisticamente significativo para a análise.

4.3 Motivações em Competição: Palavras Finais

Ao longo da pesquisa, tentamos desvendar se a marcação de pluralidade nos predicativos poderia ser explicada por motivações tipicamente formais, ou tipicamente funcionais, ou ainda, pela devida interação entre esses dois tipos de motivações, para ver se elas consistem em motivações em competição.

A análise dos resultados, discutida na seção III, mostra que os princípios funcionais, como o princípio de Economia (HAIMAN, 1983) e as Condições de Distintividade (KIPARSKY, 1972), atuam na marcação de pluralidade, presente principalmente na variável *número de marcas precedentes no sujeito e no verbo*, mas não têm força suficiente para governar todo o processo de CN na comunidade estudada. Em contrapartida, as motivações formais, gerenciada pelo Princípio do Paralelismo Formal, também exercem influência significativa na marcação explícita de pluralidade nos predicativos. Entretanto, o princípio do Paralelismo Formal, principal motivação formal, não é uma regra categórica dentro da comunidade de fala da região de São José do Rio Preto e também não atua tão decisivamente como mostrado em outros estudos sociolinguísticos.

O Princípio do Paralelismo Formal, na variedade estudada, sofre severa influência das variáveis sociais *escolaridade* e *idade*; desse modo, não é possível dar-lhe o poder explanatório que Labov (1996) lhe atribui, isto é, como um princípio explanatório de natureza totalmente interna, pois, na variedade riopretense, isso não se confirma. Por essa razão, a explicação mais plausível para a marcação de pluralidade nos predicativos é a de que há motivações em competição, nos termos de Du Bois (1985).

Como as motivações funcionais, Princípio de Economia (HAIMAN, 1983) e Condições de Distintividade (KIPARSKY, 1972), não governam com auto-suficiência a CN, as variáveis externas, *escolaridade* e *idade*, passam a ter um valor funcional, já que não pertencem à dimensão estrutural da língua, e são consideradas como motivações funcionais quando cruzadas com as variáveis internas. A essa constatação também chegaram outros estudos sociolinguísticos que consideram o Paralelismo Formal como princípio explanatório para a marcação de pluralidade. Há, todavia, uma marca decisiva no estudo aqui desenvolvido: o cruzamento entre fatores linguísticos e extralinguísticos foi seguido por análise qualitativa mais refinada, que possibilitou um enquadramento mais adequado da influência do Paralelismo Formal nos dados, antes visto como um condicionamento mecânico e repetitivo.

Outro fator que mostra a influência de idade é à disposição de alguns dados de variáveis internas nas quatro faixas etárias utilizadas na pesquisa. Essa distribuição das ocorrências revela que o uso de marcas é mais comum entre os informantes de 26 a 35 anos e a ausência de marcas está mais presente no grupo de 36 a 55 anos. A análise qualitativa nos mostra que a presença ou ausência de marcas tem relação com a faixa etária e o nível de escolaridade ao qual o informante pertence e não se trata apenas de uma característica interna da língua, a um condicionamento mecânico voltado a lei do menor esforço (LABOV, 1996).

É fato que em outras comunidades de fala, este princípio pode atuar como foi descrito por Labov (1996), já que as variedades do português são distintas; no entanto, na variedade riopretense, o que se vê é a idade e a escolaridade exercendo influência direta sobre o princípio do Paralelismo Formal; por essa razão, as motivações formais e funcionais estão em competição.

Os resultados encontrados vão ao encontro da reflexão de Du Bois (1985), visto que as motivações em competição podem ser ora funcionais ou discursivas, externas ao sistema linguístico, ora formais ou gramaticais, internas ao sistema linguístico. Para o autor, as motivações em competição podem ser ambas internas, ambas externas ou ainda uma interna e a outra externa, como no caso deste trabalho.

A marcação de pluralidade nos predicativos e participios passivos é o “bem limitado” deste trabalho, onde forças múltiplas, as motivações formais e funcionais competem entre si e não podem prevalecer sozinhas. Ao encerrar este trabalho com esta conclusão, esperamos contribuir para a discussão das motivações sociais e linguísticas da variação e para o debate sobre o grau e poder explanatório das condições formais e funcionais que podem governar a linguagem humana.

REFERÊNCIAS BIBLIOGRÁFICAS

ANTONINO, V. *A concordância nominal em predicativos do sujeito e em estruturas passivas no português popular do interior do estado da Bahia*. 123f. Dissertação de Mestrado. Universidade Federal da Bahia – Campus Universitário Ondina Salvador. 2007.

BAGNO, M., STUBBS, M., GAGNÉ, G. *Língua materna: Letramento, variação e ensino*. São Paulo: Parábola, 2002.

BAGNO, M. *A norma oculta – língua e poder na sociedade brasileira*. São Paulo: Parábola Editorial, 2003

_____. *Nada na língua é por acaso: por uma pedagogia da variação linguística*. São Paulo: Parábola Editorial, 2007

CALVET, L.J. *Saussure: pró e contra: para uma linguística social*. São Paulo: Cultrix, 1977.

CAMACHO, R. G. O Formal e o Funcional na Teoria Variacionista. In: RONCARATI, C.; ABRAÇADO, J. *Português brasileiro: contato linguístico, heterogeneidade e história*.(Org). Rio de Janeiro: Viveiros de Castro Editora, 2003. p. 55-65.

_____. *Motivações formais e funcionais da marcação de pluralidade no português falado na região de São José do Rio Preto*. Relatório encaminhado científico à FAPESP: não-publicado. São José do Rio Preto, 2008.

CASTILHO, A. T. de. Variação linguística. In; *Subsídios à proposta curricular de língua portuguesa para o 1º e 2º graus*. Coletânea de textos. v. 1. São Paulo: Secretaria da Educação. Coordenadoria de estudos e normas pedagógicas.v. 3, 1988.

CHAMBERS, J. K. *Sociolinguistic theory*.Linguistic variation and its social significance. Oxford/Cambridge: Blackwell, 1996.

CHOMSKY, N. *Aspectos da teoria da sintaxe*. Trad. de J.A. Meireles, E. Raposo. Coimbra: Armênio-Amado Editor, 1975.

DIAS, Juçá Fialho Vazzata. *A concordância de número nos predicativos e nos participios passivos na fala da região sul: um estudo variacionista*. Dissertação de Mestrado. UFSC, 1996.

DU BOIS, J. W. Competing motivations. In: HAIMAN, J. *Iconicity in syntax*. Amsterdam, John Benjamins, 1985. p. 342-365.

FARACO, C. A. *Norma culta brasileira desatando alguns nós*. São Paulo: Parábola Editorial, 2008

GARCÍA, E. C. Shifting variation. *Lingua*, v. 67, p. 189-224, 1985.

GONÇALVES, S. C. L. Projeto ALIP (Amostra Linguística do Interior Paulista): questões teóricas e metodológicas sobre a constituição de um banco de dados de língua falada. In: TAGNIN, E.O.; VALE, O. A. *Avanços da linguística de corpus no Brasil*. São Paulo: Humanitas, 2008, p.217-245.

GONÇALVES, S. C. L.; TENANI, L. E. Problemas teórico-metodológicos na elaboração de um sistema de transcrição de dados interacionais: o caso do Projeto ALIP (Amostra Linguística do Interior Paulista). *Gragoatá*. n. 25, p. 149-164, 2008.

GUIOTI, L. P. R-retroflexo na fala de São José do Rio Preto. *Estudos Linguísticos*, n. 31, São Paulo: Ed. Eletrônica, 2002.

HAIMAN, J. Iconic and economic motivation. *Language* v.59, p. 781-819, 1983.

HENGEVELD, K.; MACKENZIE, J. L. *Functional Discourse Grammar*. A typologically – based theory of language structure. Oxford: Oxford University Press, 2008.

HOCHBERG, J. G. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language*, LSA, 62(3): 609-621, set. 1986

KAY, P, McDANIEL, Ch.K. On the logic of variable rules. *Language in society*, v. 8, p. 151-87, 1979.

KIPARSKY, P. Explanation in phonology. In: PETERS, S. (ed.) *Goals of linguistic theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1972, p. 189-227.

LABOV, W. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972

_____. Where does the linguistic variable stop? A response to Beatriz Lavandera. *Sociolinguistic Working Paper*, v. 44, 1978.

_____. The overestimation of functionalism. In: Dirven, R.; Fried, V. (eds.) *Functionalism in linguistics*. Amsterdam/Philadelphia: John Benjamins, 1987, p.311-332.

_____. *Padrões sociolinguísticos*. São Paulo: Parábola Editorial, 2008

_____. La sobreestimación del funcionalismo. In: _____. *Principios del cambio lingüístico: factores internos*. Madrid: Editorial Gredos, 1996

LAVANDERA, B.R. Where does the sociolinguistic variable stop? *Language in Society*. v. 7, p. 171-82, 1978.

_____. *Variación y significado*. Buenos Aires: Librería Hachette, 1984.

_____. Where does the sociolinguistic variable stop? In: RAJENDRA, S. (org.). *Towards a critical sociolinguistics*. Amsterdam/Philadelphia: John Benjamins, 1996, p.17-30.

LEMLE, M. Heterogeneidade dialetal: um apelo à pesquisa. *Revista Tempo Brasileiro*, 53/54, p. 61-94, 1978.

LUCCHESI, D. A concordância nominal em estruturas passivas e de predicativo do sujeito em comunidades rurais isoladas afro-brasileiras no contexto da história da sociolinguística do Brasil. In: VOTRE, S; RONCARATI, C. (org.). *Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica*. Rio de Janeiro: 7Letras, 2008

NARO, A. J. Relevância da variável escolaridade. In: MOLLICA, C & BRAGA, M. L. (orgs). *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003.

PAIVA, M.C. (org.) *Amostras de fala do português falado no Rio de Janeiro*. Rio de Janeiro: UFRJ, 1999.

POPLACK, S. Delection and Disambiguation in Puerto Rican Spanish. *Language*, v.56, nº2, p. 371-385, 1980

Tagliamonte, S. A.; Robinson, J. S.; Lawrence, H. R. *GOLDVARB 2001: a multivariate analysis application for Windows*. 2001

ROMAINE, S. The status of variable rules in sociolinguistic theory. *Journal of Linguistics*. v. 17, p. 93-119, 1981

RUBIO, C. F. *A concordância verbal na língua falada na região noroeste do Estado de São Paulo*. 2008. 140f. Dissertação (Mestrado em Estudos Linguísticos). - Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto.

SALOMÃO, M. H. *Motivações Formais da Marcação de Pluralidade no SN Predicativo*. Relatório Final de Iniciação Científica submetido à FAPESP, 2007 (não-publicado)

SAUSSURE, F. Curso de linguística geral. São Paulo: Cultrix, 1997.

SCHERRE, M. M. P. *Reanálise da concordância nominal em português*. 1988. 555 f. Tese (Doutorado em Linguística). - Faculdade de Letras, Universidade Federal Rio de Janeiro, Rio de Janeiro.

_____. Sobre a Atuação do Princípio da Saliência Fônica na Concordância Nominal. In: TARALLO, F. *Fotografias Sociolinguísticas*. Campinas: Pontes: Editora da Universidade Estadual de Campinas, 1989.

_____. A concordância de número nos predicativos e nos participios passivos. *Organon*. Porto Alegre. v.5, n.17, 1991

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

_____. Sobre a Influência de Variáveis Sociais na Concordância Nominal. In: SCHERRE, M. M. P. & SILVA, G. M. O. (orgs) *Padrões sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro: Departamento de Linguística e Filologia, UFRJ, 1996.

_____. Concordância Nominal e Funcionalismo. In: *Alfa*. São Paulo, v.41, p.181-206, 1997.

_____. *Sobre a multiplicidade de variáveis independentes na sociolinguística variacionista*. Trabalho inédito apresentado na mesa-redonda 'O formal e o funcional na análise sociolinguística'. Marília: GEL, 2001.

SCHERRE, M. M. P.; NARO, A. J. Sobre as origens do português popular do Brasil. *DELTA*, v.9, n, 1, p. 01-14, 1993.

_____. A concordância de número no português do Brasil: um caso típico de variação inerente. In: HORA, D. da. (Org.). *Diversidade linguística no Brasil*. João Pessoa: Idéia, 1997, p. 93-114.

SILVA, G.M.O. Variáveis sociais e perfil do *corpus CENSO*. In: SILVA, G.M.O., SCHERRE, M.M.P. (orgs.) *Padrões sociolinguísticos*. Rio de Janeiro: Tempo Brasileiro, 1996, p. 51-81.

SILVA, G.M.O. Coleta de dados. In: MOLLICA, M.C., BRAGA, M.L. (Orgs.) *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003, p. 117-134.

TARALLO, F. *Tempos linguísticos: itinerário histórico da língua portuguesa*. São Paulo: Ática, 1990.

TERREL, T. Functional constraints on deletion of word-final /s/ in Cuban Spanish. *BLS*, v. 1, p. 431-437, 1975.

VOTRE, S., OLIVEIRA, M.R. *A Língua falada e escrita na cidade do Rio de Janeiro: materiais para seu estudo*. Rio de Janeiro: UFRJ, 1995.

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

História de São José do Rio Preto. Disponível em: <<http://www.riopreto.sp.gov.br/PortalGOV/do/conhecendoCidade?op=viewForm&coConteudo=11166&coSeqEstrutura=267&coEstruturaPai=12>>. Acesso em: 20/04/2009

BIBLIOGRAFIA CONSULTADA

CAMACHO, R. G. Uma reflexão crítica sobre a teoria sociolinguística. *D.E.L.T.A*, v. 26, n. 1, p.141-162, 2010.

GUY, G. R. Form and function in linguistic variation. In: GUY, G. et. al. *Towards a social science of language: papers in honor of William Labov*. Amsterdam: John Benjamins Publishing Company, 1996, p.221-252

GUY, G. R.; ZILLES, A. Sociolinguística Quantitativa - Instrumental de Análise. São Paulo: Parábola Editorial, 2007.

LEMLE, M. & NARO, A. J. Syntactic Diffusion. São Paulo, SBPC: *Ciência e Cultura* 29, 3, p. 259- 268, 1977

RAND, D.; SANKOFF, D. *Goldvarb version 2: a variable rule application for the Macintosh*. Montreal: University of Montreal, 1990.

SANKOFF, D. *VARBRULE2*. Mimeographed, Université de Montréal, 1975.

SCHERRE, M. M. P. Sobre a Influência de três Variáveis Relacionadas na Concordância Nominal em Português. In: SCHERRE, M. M P. & SILVA, G. M. de O e. (orgs) *Padrões sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro: Departamento de Linguística e Filologia, UFRJ, 1998

SCHERRE, M. M. P.; NARO, A. J. Mudança sem mudança: a concordância de número no português brasileiro. *Scripta*. v. 9, n. 18, p. 109-131, 2006.

VOTRE, S. J. Relevância da variável escolaridade. In: MOLLICA, M.C., BRAGA, M.L. (orgs.) *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003, p. 51-58.

GONÇALVES, S. C. L. *O português falado na região de São José do Rio Preto: constituição de um banco de dados anotado para o seu estudo*. Relatório Científico Final referente a auxílio à pesquisa concedido pela FAPESP (proc.: 03/08058-6)

Autorizo a reprodução xerográfica a fins de pesquisa.

São José do Rio Preto, 10 de setembro de 2010

MIRCIA HERMENEGILDO SALOMÃO