

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS CLÁSSICAS E VERNÁCULAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOLOGIA E LÍNGUA PORTUGUESA**

**CONCORDÂNCIA VERBAL NA LÍNGUA FALADA NAS
TRILHAS DAS BANDEIRAS PAULISTAS**

Deize Crespim Pereira

Dissertação apresentada ao Programa de Pós-Graduação em Filologia e Língua Portuguesa do Departamento de Letras Clássicas e Vernáculos da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para obtenção do título de Mestre em Letras.

Orientadora: Profa. Dra. Angela Cecília de Souza Rodrigues

**São Paulo
2004**

**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS CLÁSSICAS E VERNÁCULAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOLOGIA E LÍNGUA PORTUGUESA**

**CONCORDÂNCIA VERBAL NA LÍNGUA FALADA NAS
TRILHAS DAS BANDEIRAS PAULISTAS**

Deize Crespim Pereira

**São Paulo
2004**

Dedicado a Rui Evangelista.

AGRADECIMENTOS

A Profa. Angela Cecília de Souza Rodrigues, pela disponibilidade, dedicação e incentivo.

A Fapesp, que proporcionou os recursos financeiros necessários para a realização desta pesquisa.

A todas as pessoas que contribuíram de forma direta ou indireta para a concretização deste trabalho.

RESUMO

Esta dissertação apresenta um estudo de natureza sincrônica e quantitativa da variação entre aplicação x não-aplicação da regra de concordância verbal de 3ª e 1ª pessoas do plural. Os dados são oriundos da variedade do português popular falada por pessoas idosas residentes na zona rural dos estados de São Paulo e Minas Gerais, na área correspondente às trilhas das bandeiras paulistas. A variação da concordância verbal de 3ª pessoa do plural constitui um fenômeno comum no português do Brasil e tem sido amplamente documentada pela literatura lingüística, a qual evidencia que formas como *Eles fala x Eles falam* podem ser encontradas na fala de pessoas de diversas regiões brasileiras e de diferentes graus de escolaridade. São poucos os estudos que tratam da variação da regra de concordância com sujeitos de 1ª pessoa do plural (*Nós fala x Nós falamos*), manifestada em variedades populares do português brasileiro. Adotando os pressupostos teórico-metodológicos da Sociolingüística e da Lingüística Funcional, nosso objetivo é descrever, analisar e explicar os padrões de concordância encontrados, buscando identificar os fatores lingüísticos (semânticos, gramaticais, discursivos) e sociais que condicionam a realização ou a não-realização da regra formal de concordância verbal.

PALAVRAS-CHAVE: concordância verbal, variação lingüística, português popular, Sociolingüística, Lingüística Funcional.

ABSTRACT

This research presents a synchronic and quantitative study of the variation between application x non-application of the subject-verb agreement rule with the 3rd and 1st plural persons. The data were collected from Popular Portuguese spoken by elderly who live in the rural areas of São Paulo and Minas Gerais, in the regions that coincide with the historical route of the “*Bandeirantes*” (members of colonial exploratory expeditions). The variation of subject-verb agreement with the 3rd plural person is an ordinary phenomenon in Brazilian Portuguese, and it has been extensively documented by linguistic literature, which gives evidence that forms as *Eles fala x Eles falam* can be found in the speech of people of various Brazilian regions and of different education degrees. Few are the studies that deal with the variation of subject-verb agreement with the 1st plural person (*Nós fala x Nós falamos*), displayed in popular varieties of Brazilian Portuguese. Using the theories of Sociolinguistics and Functional Linguistics, our goal is to describe, analyze and explain the patterns of agreement found, searching for linguistic (semantic, grammatical, discursive) and social factors that condition the application or the non-application of the formal subject-verb agreement rule.

KEY WORDS: subject-verb agreement, linguistic variation, Popular Portuguese, Sociolinguistics, Functional Linguistics.

SUMÁRIO

Introdução.....	10
Capítulo 1: Pressupostos teórico-metodológicos.....	12
1.1 A Sociolingüística.....	15
1.2 A Lingüística Funcional.....	19
1.3 Uma análise funcional das categorias sujeito e verbo e da concordância.....	28
Capítulo 2: A concordância verbal na literatura lingüística.....	33
2.1 Fatores lingüísticos.....	34
2.2 Fatores sociais.....	38
2.3 As hipóteses que explicam a origem da variação da concordância verbal.....	42
Capítulo 3: O <i>corpus</i> sob análise.....	47
3.1 O Projeto Filologia Bandeirante.....	47
3.2 Perfil social dos informantes.....	48
3.3 O português popular.....	48
3.4 A coleta de dados.....	49
3.5 Normas de transcrição.....	51
Capítulo 4: Análise dos dados.....	52
4.1 Contextos de variação da concordância verbal no português popular e critérios para seleção de ocorrências.....	52
4.1.1 Critérios de inclusão.....	52
4.1.2 Critérios de exclusão.....	55
4.2 A realização da regra de concordância verbal.....	60
4.2.1 Concordância de 3ª pessoa do plural: fatores lingüísticos.....	61
4.2.1.1 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	63
4.2.1.2 Presença/ausência do sujeito pronominal.....	67
4.2.1.3 Paralelismo discursivo.....	71
4.2.1.4 Traço semântico do sujeito.....	75
4.2.1.5 Paralelismo oracional.....	76
4.2.1.6 Papel semântico do sujeito.....	80
4.2.2 Concordância de 1ª pessoa do plural: fatores lingüísticos.....	85
4.2.2.1 Acentuação da forma verbal.....	86
4.2.2.2 Paralelismo discursivo.....	87
4.2.2.3 Status informacional do sujeito.....	90
4.2.2.4 Distância do referente do sujeito zero.....	98
4.2.3 Fatores sociais.....	100
4.2.3.1 Escolaridade.....	101
4.2.3.2 Sexo.....	102
4.2.3.3 Procedência.....	103
Considerações finais.....	105
Referências bibliográficas.....	113

ÍNDICE DE QUADROS

1. Paradigma Formal x Paradigma Funcional (Dik,1978, In: Castilho, no prelo).....	13
2. Gramática Formal x Gramática Funcional (Halliday, 1994).....	14
3. Metafunções (Halliday, 1994).....	21
4. Fatores lingüísticos utilizados na análise da concordância de 3ª pessoa do plural (P3).....	61
5. Fatores lingüísticos (P3) selecionados e descartados pelos programas estatísticos do Varbrul.....	62
6. Fatores lingüísticos utilizados na análise da concordância de 1ª pessoa do plural (P1).....	85
7. Fatores lingüísticos (P1) selecionados e descartados pelos programas estatísticos do Varbrul.....	86
8. Fatores sociais utilizados na análise da concordância de P3 e P1.....	100

ÍNDICE DE TABELAS

1. Frequência e probabilidade de aplicação da regra de concordância verbal (P3) conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	64
2. Frequência de não-aplicação da regra de concordância verbal (P3) conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural, no <i>corpus</i> sob análise e entre os falantes de 3ª faixa etária cujos dados compõem parte do <i>corpus</i> do português popular de São Paulo (Rodrigues, 1987:231).....	65
3. Frequência de aplicação da regra de concordância verbal (P3) com sujeitos pospostos ao verbo conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	66
4. Frequência e probabilidade de aplicação da regra de concordância verbal (P3) de acordo com a presença/ausência do sujeito pronominal.....	68
5. Frequência de aplicação da regra de concordância verbal (P3) com sujeito não-pronominal e pronominal explícito conforme a posição e distância do sujeito em relação ao verbo.....	69
6. Frequência de aplicação da regra de concordância verbal (P3) segundo a presença/ausência do sujeito pronominal e o grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	70
7. Frequência e probabilidade de aplicação da regra de concordância verbal (P3) segundo o paralelismo discursivo.....	73
8. Frequência de aplicação da regra de concordância verbal (P3) conforme o paralelismo discursivo e a presença/ausência do sujeito pronominal.....	73
9. Frequência e probabilidade de aplicação da regra de concordância verbal (P3) de acordo com o traço semântico do sujeito [+humano].....	75

10. Frequência de aplicação da regra de concordância verbal (P3) conforme o traço semântico do sujeito [+humano] e a presença/ausência do sujeito pronominal.....76
11. Frequência de aplicação da regra de concordância verbal (P3) segundo o paralelismo oracional (1ª categorização).....78
12. Frequência e probabilidade de aplicação da regra de concordância verbal (P3) segundo o paralelismo oracional (2ª categorização).....78
13. Frequência de aplicação da regra de concordância verbal (P3) conforme o paralelismo oracional e a presença/ausência do sujeito pronominal.....79
14. Frequência e probabilidade de aplicação da regra de concordância verbal (P3) de acordo com o papel semântico do sujeito.....82
15. Frequência e probabilidade de aplicação da regra de concordância verbal (P1) de acordo com a acentuação da forma verbal.....86
16. Frequência e probabilidade de aplicação da regra de concordância verbal (P1) segundo o paralelismo discursivo.....88
17. Frequência de aplicação da regra de concordância verbal (P1) conforme o paralelismo discursivo e a acentuação da forma verbal.....89
18. Frequência e probabilidade de aplicação da regra de concordância verbal (P1) conforme o status informacional do sujeito.....96
19. Frequência de aplicação da regra de concordância verbal (P1) conforme o status informacional do sujeito e a presença/ausência do sujeito pronominal.....96
20. Frequência e probabilidade de aplicação da regra de concordância verbal (P1) de acordo com a distância do referente do sujeito zero.....99
21. Frequência e probabilidade de aplicação da regra de concordância verbal (P3 e P1) segundo o grau de escolaridade do informante.....101
22. Frequência e probabilidade de aplicação da regra de concordância verbal (P3) conforme o sexo do informante.....102
23. Frequência de aplicação da regra de concordância verbal (P3) de acordo com o sexo e o grau de escolaridade do informante.....102
24. Frequência e probabilidade de aplicação da regra de concordância verbal (P3) segundo a procedência do informante.....103
25. Frequência geral de não-aplicação da regra de concordância verbal (P3 e P1) no *corpus* sob análise e entre os falantes de terceira faixa etária cujos dados integram o *corpus* do português popular de São Paulo (Rodrigues, 1987:218).....103

INTRODUÇÃO

O objetivo deste trabalho é descrever os padrões de concordância verbal de 3ª e 1ª pessoas do plural, encontrados na variedade do português popular falada por pessoas idosas residentes na zona rural dos estados de São Paulo e Minas Gerais, na área correspondente às trilhas das bandeiras paulistas. Trata-se de um estudo de natureza sincrônica e quantitativa que busca analisar e explicar a variação entre aplicação x não-aplicação da regra formal de concordância verbal, tomando por base os pressupostos teórico-metodológicos da Sociolinguística e da Linguística Funcional. A pesquisa inclui ainda a comparação dos resultados obtidos com os encontrados por Rodrigues (1987), no português popular de São Paulo.

As seguintes hipóteses orientam o presente estudo:

- (i) os falantes da variedade popular, registrada nas trilhas das bandeiras, tenderiam a não aplicar sistematicamente a regra padrão de concordância verbal, estabelecida pela gramática normativa;
- (ii) a variação entre aplicação/não-aplicação da regra de concordância verbal não seria aleatória, mas condicionada tanto por fatores lingüísticos, internos ao sistema, como por fatores extralingüísticos, de natureza social;
- (iii) partindo do pressuposto de que o modelo de realização lingüística se encontra mais disponível em áreas urbanas, o *corpus* da presente pesquisa, composto por inquéritos de falantes que residem na zona rural, apresentaria uma freqüência maior de formas de não-concordância do que o *corpus* do português popular de São Paulo (Rodrigues, 1987), constituído de inquéritos de informantes que habitam uma comunidade urbana;
- (iv) casos de não-concordância ocorreriam com a mesma freqüência tanto na 3ª, como na 1ª pessoa do plural.

O trabalho apresenta-se dividido em quatro capítulos. No primeiro, discorremos sobre a metodologia sociolinguística utilizada por Labov (1991, 1994) para o estudo de fenômenos linguísticos variáveis, e sobre a proposta de análise funcionalista de Halliday (1994, 1987, 1974), as quais serviram de base para o estudo da concordância verbal no português popular.

O capítulo 2 compreende uma resenha de trabalhos sobre a concordância no português do Brasil, ressaltando os aspectos considerados importantes para o desenvolvimento da presente pesquisa: os fatores linguísticos e extralinguísticos que condicionam a realização x não-realização da concordância verbal e as hipóteses formuladas para explicar a origem da variação na aplicação da regra.

No capítulo 3, fazemos uma apresentação do *corpus* sob análise, bem como das características sociais dos informantes cujos inquéritos são estudados.

O capítulo 4 corresponde à parte nuclear da dissertação e abrange a descrição, a análise e a interpretação dos dados. Tomando por base os cálculos fornecidos pelo pacote de programas computacionais Varbrul, discorremos, neste capítulo, sobre os fatores linguísticos e sociais que condicionam a variação na aplicação da regra de concordância verbal.

Em Considerações Finais, retomamos brevemente os resultados mais relevantes da pesquisa e fazemos algumas sugestões para estudos futuros.

Esperamos que a realização deste trabalho contribua para uma melhor caracterização e compreensão das estruturas de concordância verbal no português do Brasil. Conforme atestam Rodrigues & Campos (2002:12), “*uma visão completa, ou mais abrangente da questão de concordância verbal em português só poderá ser obtida se levarmos em conta diferentes tipos de textos realizados no plano da oralidade e da escrita, representativos de múltiplas variedades do português em diferentes momentos de sua história*”.

Acreditamos, sobretudo, que é a partir do “*exame dos usos reais que se podem descobrir as normas linguísticas*” (Rodrigues, 1987:82), sendo a tarefa do linguista precisamente esta: descrever e explicar o funcionamento dessas regras normais no discurso de uma comunidade de falantes.

A ausência de marcas formais de concordância é considerada um erro do ponto de vista da norma padrão, que se mantém rigidamente vinculada à língua escrita. Não obstante, formas verbais não-marcadas se mostram freqüentes na fala popular e são encontradas até mesmo entre falantes escolarizados do português do Brasil. Estudos como este se mostram importantes e necessários, pois comprovam a existência de uma norma brasileira, que se configura no uso efetivo que os falantes fazem da língua portuguesa no Brasil.

CAPÍTULO 1:

PRESSUPOSTOS TEÓRICO-METODOLÓGICOS

O suporte teórico-metodológico desta pesquisa é fornecido pela Sociolingüística e pela Lingüística Funcional. Adotamos principalmente os postulados da Sociolingüística Quantitativa de William Labov e da Gramática Funcional Sistêmica de Michael A. K. Halliday. Estas duas perspectivas de análise se combinam perfeitamente, uma vez que ambas: (i) adotam como objeto de estudo o uso efetivo da linguagem; (ii) vêem a língua falada como um *corpus* privilegiado para os estudos lingüísticos; (iii) postulam que a forma da língua é determinada pelo seu uso – concepção que implica uma interpretação da língua como um sistema variável, flexível e dinâmico; (iv) procuram depreender as relações entre o uso lingüístico e o contexto social.

Estes pressupostos integram o que podemos chamar de uma análise lingüística funcional, em contraposição à análise formal. Conforme Halliday (1994), o principal contraste entre as teorias lingüísticas do século XX se encontra justamente na oposição entre as Gramáticas Formais, com suas raízes na Lógica e na Filosofia, incluindo-se aí o Estruturalismo e o Gerativismo, e as Gramáticas Funcionais, com suas raízes na Retórica e na Etnografia¹. Em outra oportunidade, Halliday (1987:56) denomina estas duas perspectivas para o estudo da linguagem de, respectivamente: (i) “*intra-organism perspective*”: língua como conhecimento (“*we see language as what goes on in the head*”); (ii) “*inter-organism perspective*”: língua como interação (“*what goes on between people*”).

Vistas ora como alternativas, ora como complementares, ambas constituem, com efeito, as duas principais linhas de análise da lingüística contemporânea. Segundo Halliday, suas origens residem parte no pensamento ocidental e parte na própria natureza da linguagem: “*which is equally at home in humanities, social science, natural science, medicine and engineering, but appears very different according to where one starts*” (1994:xxviii).

As principais distinções entre uma e outra teoria são resumidas nos quadros a seguir, retirados de Dik (1978, In: Castilho, no prelo) e de Halliday (1994):

¹ Conforme notam diversos autores (cf. Pezatti, 1992; Scherre, 1988; Lavandera, 1984; Castilho, no prelo), a abordagem sociolingüística certamente se inclui na área funcionalista, na medida em que estuda a língua em uso em seu contexto social, considerando as relações entre forma e função da linguagem.

PARADIGMA FORMAL	PARADIGMA FUNCIONAL
A língua é um conjunto de sentenças.	A língua é um instrumento de interação social.
A função primária da língua é a expressão dos pensamentos.	A função primária da língua é a comunicação.
O correlato psicológico da língua é a competência: a capacidade de produzir, interpretar e julgar sentenças.	O correlato psicológico da língua é a competência comunicativa: a habilidade de conduzir a interação social por meio da língua.
O estudo da competência tem uma prioridade lógica e metodológica sobre o estudo do desempenho.	O estudo do sistema lingüístico deve ter lugar no interior do sistema de usos lingüísticos.
As sentenças de uma língua devem ser descritas independentemente do contexto em que ocorreram.	A descrição dos elementos lingüísticos de uso de uma língua deve proporcionar pontos de contato com o contexto em que ocorreram.
A aquisição da língua é inata. Os inputs são restritos e não estruturados. A teoria do estímulo é pobre.	A criança descobre o sistema que subjaz à língua e ao uso lingüístico ajudada por inputs de dados lingüísticos extensos e altamente estruturados, presentes em contextos naturais.
Os universais lingüísticos são propriedades inatas do organismo biológico e psicológico dos homens.	Os universais lingüísticos são especificações inerentes às finalidades da comunicação, à constituição dos usuários da língua e aos contextos em que a língua é usada.
A Sintaxe é autônoma em relação à Semântica. A Sintaxe e a Semântica são autônomas com relação à Pragmática, e as prioridades vão da Sintaxe à Pragmática via Semântica.	A Pragmática é a moldura dentro da qual a Semântica e a Sintaxe devem ser estudadas. A Semântica é dependente da Pragmática, e as prioridades vão da Pragmática para a Sintaxe via Semântica.

Quadro 1: Paradigma Formal x Paradigma Funcional (Dik, 1978:4, In: Castilho, no prelo)

GRAMÁTICA FORMAL	GRAMÁTICA FUNCIONAL
Orientação sintagmática (“ <i>chain grammar</i> ”).	Orientação paradigmática (“ <i>choice grammar</i> ”).
A língua é interpretada como um conjunto de estruturas entre as quais relações podem ser estabelecidas.	A língua é interpretada como uma rede de relações, sendo as estruturas realizações, manifestações dessas relações.
Ênfase nos universais lingüísticos.	Ênfase na variação entre as diferentes línguas.
A Sintaxe é o fundamento da língua.	A Semântica é o fundamento da língua.
A Gramática é arbitrária.	A Gramática é natural.
Organização em torno da sentença.	Organização em torno do texto ou discurso.

Quadro 2: Gramática Formal x Gramática Funcional (Halliday, 1994: xxviii)

A escolha da Sociolingüística e da Lingüística Funcional como norteadoras da presente pesquisa está relacionada: (i) ao *corpus* sob análise: este registra o português popular falado, variedade de língua que é definida com base tanto em parâmetros sociolingüísticos, como funcionais; (ii) ao fenômeno lingüístico analisado: a metodologia sociolingüística fornece os instrumentos necessários para dar conta da variação na aplicação da regra de concordância do verbo com o sujeito, e a Lingüística Funcional, por seu turno, proporciona uma retaguarda teórica para explicar os resultados encontrados. Acreditamos que as questões postas pelo discurso são certamente pertinentes, inclusive para a análise de fenômenos gramaticais como a concordância; (iii) a uma inclinação pessoal: conforme nota Possenti (1979), a visão do lingüista, além de não ser neutra, é necessariamente parcial e influenciada pelo quadro geral de sua época.

Faremos a seguir uma exposição dos pressupostos teóricos e metodológicos da Sociolingüística Variacionista Laboviana, bem como da Gramática Funcional Sistêmica de Halliday. Na medida em que prosseguirmos em nossa explanação, outros autores também serão mencionados cujas obras são relevantes para a compreensão das duas linhas de pesquisa em questão.

1.1 A SOCIOLINGÜÍSTICA

Na introdução à tradução espanhola de *Sociolinguistic Patterns* (Labov, 1991), Herreras (1983:12) nota que se tivesse que escolher um único aspecto para definir o estilo teórico e a estratégia de investigação de Labov, este seria o do “*realismo empírico*”, isto é, a fidelidade à experiência mesma da língua, em sua circulação e atuação em um dado contexto social. De fato, os estudos realizados por Labov têm como principal característica a análise de discursos concretos, objetivando descrever a gramática efetiva de uma determinada comunidade lingüística, bem como depreender as relações entre padrões lingüísticos e sociais. Esta lingüística “*racional*” e “*socialmente realista*” (Labov, 1991: xiii, xv, xvi) recebeu o nome de “Sociolingüística”, o que para o autor equivale a uma “*utilização enganosa de um termo estranhamente redundante*” (1991:183), posto que a linguagem constitui uma forma de comportamento social, um instrumento de comunicação utilizado por seres reais em um dado contexto situado no tempo e na história.

A Teoria Variacionista Laboviana ou Sociolingüística Quantitativa interpreta a variação e a heterogeneidade como características inerentes ao sistema lingüístico. Labov (1991) nota que grande parte de nossas teorias sobre a linguagem concerne ao isolamento de unidades funcionais invariantes, como fonemas, morfemas, etc. Contudo, o estudo empírico da língua em uso nos mostra que os elementos da estrutura lingüística não estão confinados a essas unidades, abrangendo também elementos que se encontram em variação sistemática, refletindo mudanças temporais e processos extralingüísticos.

Para incorporar a variação na descrição e na teoria lingüísticas, Labov (1991) introduz alguns conceitos teórico-metodológicos extremamente relevantes para nossa pesquisa. Segundo o autor, uma determinada gramática abrange não só *regras categóricas* (i.e. regras lingüísticas que sempre se aplicam), mas também *regras variáveis* – conceito que é utilizado para substituir a noção de regra opcional, na medida em que não pressupõe variação livre, mas sim sistemática. As formas lingüísticas em variação em uma determinada comunidade de fala são denominadas *variantes lingüísticas*. Estas são definidas como formas alternativas de se dizer a mesma coisa, em um mesmo contexto. Embora sejam idênticas em seu valor referencial, as variantes podem opor-se quanto ao seu significado social e/ou estilístico. Ao conjunto de variantes dá-se o nome de *variável lingüística*. Assim, a concordância verbal no português do Brasil constitui precisamente uma *regra variável*, ou uma *variável lingüística* que abrange duas *variantes*: a aplicação e a não-aplicação da regra.

Na perspectiva sociolingüística, a variação é sistemática, não-aleatória, na medida em que é condicionada tanto por fatores lingüísticos, de ordem estrutural, internos ao sistema, como por fatores extralingüísticos, ou seja, de natureza social. Estes abrangem: (i) a história, a cultura e o sistema de organização política, social, econômica e geográfica de uma comunidade; (ii) características individuais dos falantes, como sexo, idade, escolaridade, classe social, ocupação, etc; (iii) aspectos da situação imediata, ou seja, do contexto em que um discurso é produzido. Os fatores condicionantes lingüísticos e sociais são também denominados *variáveis independentes*, em contraposição à *variável dependente*, que constitui o fenômeno lingüístico estudado.

O estudo da variação lingüística proposto por Labov é necessariamente de natureza quantitativa. O pacote de programas computacionais Varbrul (“Variable Rules” de D. Sankoff) é utilizado para o tratamento estatístico dos dados, efetuando cálculos de frequência e probabilidade de aplicação de uma determinada regra gramatical, avaliando o peso relativo de cada fator condicionante e realizando o cruzamento dos fatores. Somente deste modo, torna-se o lingüista apto a identificar os fatores internos e externos que estão a reger a variação lingüística, bem como estabelecer a correlação entre comportamento lingüístico e estrutura social.

A análise da variação é pertinente, na medida em que esta é interpretada como uma condição indispensável para entender a mudança lingüística. Para Labov, os processos de mudança que se verificam em uma comunidade de fala se atualizam na variação observada em cada momento nos padrões de comportamento lingüístico dos falantes, sendo que se a mudança implica a variação, a variação não implica necessariamente a mudança.

O processo de mudança lingüística envolve três estágios: origem, propagação e conclusão: *“In the ‘origin’ of a change, it is one of innumerable variations confined to the use of a few people. In the ‘propagation’ of the change, it is adopted by such large number of speakers that it stands in contrast to the older form along a broad front of social interaction. In the ‘completion’ of the change, it attains regularity by the elimination of competing variants”* (Labov, 1991:123).

Para determinar se uma variável está envolvida em um processo de mudança, Labov analisa sua distribuição em *tempo real* (dimensão diacrônica) e em *tempo aparente* (dimensão sincrônica – distribuição da variável nas várias faixas etárias da população). O autor acredita que obteremos uma melhor compreensão do mecanismo da mudança lingüística e das mudanças completadas no passado, se estudarmos detalhadamente aquelas que estão em progresso no presente. Labov aproxima a análise sincrônica da diacrônica, apoiando-se no *Princípio da Uniformidade*, que estabelece que os mesmos mecanismos que produziram grandes mudanças no passado podem ser observados operando sobre as mudanças que se dão atualmente.

O estudo empírico da mudança lingüística em progresso compõe-se de cinco problemas distintos: (i) *restrição* (“*constraints*”) – verificar quais são as restrições universais sobre a mudança; (ii) *transição* (“*transition*”) – descrever os caminhos pelos quais a língua muda, ou seja, os estágios intermediários da mudança; (iii) *encaixamento* (“*embedding*”) – correlacionar as variáveis lingüísticas entre si, com outros elementos do sistema lingüístico, e com aspectos do contexto social; (iv) *avaliação* (“*evaluation*”) – investigar como os membros da comunidade de fala avaliam a mudança e quais informações expressivas as variantes veiculam; (v) *implementação* (“*actuation*”) – analisar os motivos pelos quais uma mudança ocorre num momento e num lugar particulares, ou seja, descobrir quais os eventos lingüísticos ou sociais que são os gatilhos da mudança (Labov, 1991:283).

O autor classifica os elementos lingüísticos envolvidos no processo de mudança, conforme o julgamento social a que são submetidos em: (i) *indicadores* – apresentam diferenciação social, mas não estilística e não têm efeito sobre o julgamento social dos falantes; (ii) *marcadores* – apresentam variação social e estilística e se refletem de forma mais ou menos consciente no julgamento dos falantes; (iii) *estereótipos* – são formas socialmente estigmatizadas pela sociedade.

Labov nota que mudanças lingüísticas em seu estágio inicial dificilmente recebem uma avaliação social ou um reconhecimento, e nem todas as mudanças se tornam objeto de atenção consciente, mesmo se estiverem em um estágio avançado. O autor observa ainda que mudanças gramaticais dificilmente passam despercebidas. De qualquer forma, um significado social somente pode ser atribuído a uma forma lingüística quando há variação. Conforme o processo de mudança avança, o fenômeno lingüístico é passível de ser associado aos valores sociais atribuídos ao grupo no qual a mudança se originou: se este é um grupo de prestígio, a nova forma pode ser incorporada na variedade de língua dominante, se o grupo é excluído, por outro lado, a forma lingüística pode ser estigmatizada, corrigida ou até extinta². Contudo, Labov observa que a correção não tem o caráter regular da mudança, pois focaliza em determinados sons ou palavras proeminentes.

Quanto aos estereótipos, Labov nota que pode ou não haver uma relação entre estes e o uso lingüístico. Poucas pessoas têm consciência de que utilizam formas estigmatizadas, mas em toda comunidade existem falantes que são mais conscientes dos modelos de prestígio e cujo comportamento lingüístico é mais influenciado pela norma padrão. O autor estabelece ainda que este reconhecimento está relacionado à escolarização, sendo que somente no primeiro ano ginasial o

² Labov postula ainda que podemos estabelecer se um determinado elemento lingüístico está envolvido em um processo de mudança, observando o seu padrão de distribuição social: se há uma *estratificação brusca* (i.e. grande variação no padrão de uso, conforme o grupo social considerado), temos uma variável com um significado social estável, não envolvida em um processo de mudança, mas se há uma *estratificação gradual* (i.e. sem descontinuidade na frequência de uso entre os grupos sociais), a variável está envolvida em um processo de mudança.

falante começa a adquirir as normas de avaliação, tornando-se sensível ao significado social de sua própria forma de falar assim como daqueles que o rodeiam.

É importante ressaltar também que as noções de *prestígio* e de *estigma* são definidas em função das pessoas que utilizam a linguagem e da situação. Assim, uma mesma variante lingüística pode ter valor positivo para uns, e negativo para outros. Isto acontece, por exemplo, com o falante rural: em seu local de origem, determinadas características de sua fala podem ser tomadas como uma marca de identidade, apresentando um valor positivo, mas na medida em que este falante migra para os centros urbanos, estas mesmas características podem ser estigmatizadas. Como consequência da mudança de ambiente, há uma rápida transformação das formas lingüísticas rurais mais salientes.

A *comunidade lingüística* é vista como um grupo de pessoas que compartilham um conjunto de normas e atitudes com respeito à linguagem, e não como um grupo de pessoas que falam da mesma maneira. Ainda que seus membros partilhem uma *variedade de língua*³ e avaliem as formas lingüísticas de maneira uniforme, há uma enorme diferenciação entre eles quando consideramos sua performance objetiva.

Labov nota que, embora os estudos tradicionais de dialetos regionais postulem que o isolamento leva à diversidade lingüística, enquanto a mistura de populações leva à uniformidade, nos seus próprios estudos realizados em centros metropolitanos, ao invés de uma diferenciação horizontal (geográfica), há uma diferenciação vertical (social) que não pressupõe isolamento do estrato lingüístico considerado: grupos que mantêm relações estreitas podem participar de rápidas mudanças lingüísticas que levam a um grau ainda maior de diversidade.

Em suma, o autor postula que a chave para uma melhor compreensão da mudança lingüística, bem como da própria linguagem, é a possibilidade de descrever a diferenciação ordenada de uma dada língua servindo uma comunidade real. Acreditamos que todos estes conceitos teórico-metodológicos fornecidos pela Sociolingüística são essenciais para uma pesquisa que adota como objeto de estudo um fenômeno de variação observado a partir do uso lingüístico em seu contexto social.

³ Rodrigues (1987:77) define variedade lingüística como “*um conjunto de itens lingüísticos com distribuição social semelhante*”.

1.2 A LINGÜÍSTICA FUNCIONAL

Assim como a Sociolingüística, a Lingüística Funcional é projetada para dar conta de como a língua é utilizada, ou seja, de como “*as estruturas semânticas e gramaticais funcionam no próprio ato de comunicação*” (Danes in TLP1, In: Halliday, 1974:44). Todo discurso, isto é, tudo que é falado ou escrito, toda forma de interação lingüística efetiva, se desdobra em algum contexto de uso e, segundo Halliday, é justamente o uso da língua que molda o sistema lingüístico: “*no doubt language has evolved in this way because of the ways in which it is used*” (1987:187). A língua se desenvolveu para satisfazer certas necessidades na vida dos homens e a maneira como ela é organizada é funcional com respeito às finalidades que cumpre. O autor postula que uma gramática funcional é essencialmente uma gramática “natural”, no sentido de que todas as estruturas podem ser explicadas com referência ao modo como a língua é usada.

O objetivo da Lingüística Funcional é, pois, analisar o uso lingüístico, mas para interpretá-lo, deve relacioná-lo ao sistema: “*in order to make sense of the text, what the speaker actually says, we have to interpret it against the background of what he ‘can say’*” (Halliday, 1987:40). O autor, no entanto, minimiza a dicotomia competência x desempenho, na medida em que parte do concreto para o abstrato, do uso efetivo para o sistema ideal, interpretando o texto como uma realização do sistema.

Halliday (1994) denomina sua linha de análise como uma “*Gramática Funcional Sistêmica*”, que segue a tradição funcional européia, sendo baseada principalmente em Firth (Escola de Londres) e Hjelmslev (Escola de Praga). A Gramática Funcional Sistêmica prioriza as relações paradigmáticas, interpretando a língua não como um inventário de estruturas, mas como uma rede de sistemas, isto é, um conjunto de opções interrelacionadas para construir sentido. Estas opções são abstratas e a estrutura é o meio pelo qual elas são realizadas. Em suas próprias palavras: “*language as resource – resource for meaning, with meaning defined in terms of function*” (1987:17).

Há dois aspectos fundamentais na interpretação funcional do sistema lingüístico: sua organização: (1) em estratos ou níveis, (2) em um pequeno número de componentes funcionais.

1. Na Gramática Funcional de Halliday, os termos utilizados para os níveis da língua – “*the stages in the coding process from meaning to expression*” (1994:xiv) – são: Semântica, Gramática e Fonologia. O estrato central da língua é a Gramática, também denominada Léxico-gramática. Esta é expressa na forma de som ou escrita, daí a existência dos níveis Fonologia e Grafologia, que servem como meios alternativos de expressão. Utilizando a metáfora de um espaço vertical, Halliday postula que a Fonologia e a Grafologia são os estratos “abaixo” da Gramática. Esta, por sua vez,

realiza padrões de um outro nível “mais alto”: o da Semântica, também referida pelo termo Semântica do Discurso, para explicitar que o que se investiga é como as unidades gramaticais são construídas no discurso.

O autor prefere não empregar o termo “Sintaxe”, uma vez que esta é vista somente como uma parte da Gramática, que consiste, além disso, de Léxico e Morfologia. Há ainda uma outra razão para não utilizá-lo: a designação sugere uma direção particular: a língua é interpretada como um sistema de formas às quais significados são atribuídos. Na história da lingüística ocidental, esta foi a direção adotada: estudava-se primeiramente as formas das palavras (Morfologia), em seguida, a formação das sentenças (Sintaxe), e só depois o significado. A Gramática Funcional sugere, porém, a direção inversa: a língua é interpretada como um sistema de significações acompanhado de formas as quais as realizam. Sob esta perspectiva, a forma da língua é vista como um meio para um fim e não como um fim em si mesma. Como os significados são expressos, como se codificam gramaticalmente, constituem, pois, questões centrais da análise funcionalista.

Segundo Halliday, a relação entre Semântica (“*meaning*”) e Gramática (“*wording*”) não é arbitrária⁴: “*the form of the grammar relates naturally to the meanings that are being encoded*”; “*the general kinds of grammatical pattern that have evolved in language, and the specific manifestations of each kind, bear a natural relation to the meanings they have evolved to express*” (1994:xvii, xviii). A língua adulta construiu estruturas semânticas que nos permitem pensar sobre nossa experiência e o sistema de significações engendrou estruturas léxico-gramaticais correspondentes, como por exemplo, verbos e substantivos para encaixar com a análise da experiência em processos e participantes. É assim que as crianças são capazes de aprender a gramática: elas podem estabelecer uma ligação entre as categorias gramaticais e a sua realidade exterior e interior. A Gramática Funcional é projetada com o objetivo de trazer à luz esta relação natural: é um estudo das formas lingüísticas, mas que as interpreta com referência ao que significam.

2. Conforme Halliday (1994), os componentes fundamentais de significado na língua são funcionais. Todas as línguas são organizadas em torno de dois principais tipos de significação: a) Ideacional (reflexivo), b) Interpessoal (ativo). Estes componentes, denominados “metafunções”⁵, são as manifestações no sistema lingüístico de duas finalidades gerais que subjazem todos os usos da língua: a) entender o ambiente, b) atuar sobre os outros, neste ambiente. Combinado com estes

⁴ O autor admite uma relação arbitrária entre Gramática e Fonologia.

⁵ Halliday (1987) distingue “função” e “uso” da linguagem. Segundo o autor, os usos lingüísticos são inumeráveis. O fato de a língua servir a uma variedade imensa de propósitos se deve justamente à sua organização em um pequeno número de componentes funcionais generalizados. O termo “metafunção” refere-se, portanto, às funções da língua incorporadas no sistema lingüístico.

dois, há um terceiro componente metafuncional: o Textual, que organiza a expressão como mensagem.

É importante ressaltar que cada estrato, cada componente metafuncional, é descrito como um sistema de potencialidades, um conjunto de opções, de alternativas de que o falante se vale, para construir significado. De acordo com Halliday, os falantes de uma língua podem, assim, “brincar” com o sistema, realizando uma série de escolhas correlacionadas que têm relevância semântica.

Halliday (1994) analisa a oração como uma combinação de três estruturas distintas, derivadas dos componentes funcionais (metafunções), cada uma expressando um tipo de significado: “*oração como representação*”, “*oração como troca*”, “*oração como mensagem*”. É a estrutura como um todo, a configuração total das funções, que realiza o sentido da oração.

Metafunção	Definição (tipo de sentido)	Status na sentença
1.Experiencial	Construindo um modelo da experiência	Oração como representação
2. Interpessoal	Desempenhando relações sociais	Oração como troca
3. Textual	Criando relevância para o contexto	Oração como mensagem
4. Lógica	Construindo relações lógicas	----- ⁶

Quadro 3: Metafunções (Halliday, 1994:36)

Três estruturas expressam estes três tipos de escolha semântica realizadas pelo falante:

- (1) estruturas de transitividade expressam o sentido representacional: o que a mensagem concerne, tipicamente algum processo associado a participantes e a circunstâncias; Halliday (1994) classifica os processos verbais em: material (“*doing*”), mental (“*sensing*”), relacional (“*being*”), comportamental (“*behaving*”), existencial (“*existing*”) e verbal (“*saying*”); conforme o tipo de processo verbal, os participantes assumem funções diversas, como por exemplo: Ator, Meta, etc.;
- (2) estruturas de modo expressam o sentido interacional: o que a mensagem realiza como uma troca verbal entre falante x ouvinte, os papéis adotados pelos interlocutores e suas atitudes; em sua análise, Halliday divide a oração em: Modo – constituído de sujeito e operador verbal expressando tempo ou modalidade – e Resíduo – consistindo de verbo, complemento e adjunto;
- (3) estruturas temática e informacional expressam a organização da mensagem: como a oração se relaciona com o discurso e com o contexto em que está sendo produzida.

⁶ As funções Experiencial e Lógica compõem juntas a função Ideacional. A função Experiencial se refere ao sentido como organização da experiência, e a função Lógica, às relações entre orações, que se manifestam na forma de subordinação, ou de independência.

Estes três níveis de análise se referem, portanto, às funções semânticas, gramaticais e discursivas exercidas pelos elementos lingüísticos. Vale ressaltar que tanto a forma da expressão lingüística, como o seu conteúdo semântico, podem variar de acordo com a atribuição de diferentes funções aos seus constituintes, em cada um dos três níveis. Nos parágrafos subseqüentes, nós nos deteremos na função Textual, por acreditarmos que esta constitui um instrumento de análise que pode ajudar a explicar os padrões de concordância sujeito-verbo encontrados no *corpus*.

Segundo Halliday (1994), a função que dá à oração um caráter de mensagem, o status de um evento comunicativo, é a função Textual. Esta é realizada através dos seguintes recursos:

- (1) estrutura Temática (Tema x Rema)
- (2) estrutura Informacional e Foco (Dado x Novo)
- (3) coesão (referência, elipse, conjunção e organização lexical)

1. No inglês, como em muitas outras línguas, a oração é organizada como mensagem, assinalando um status especial a uma parte dela: um elemento é enunciado como Tema. Seguindo a terminologia dos lingüistas da Escola de Praga, Halliday (1994) define *Tema* como o elemento que serve como ponto de partida da mensagem, aquilo sobre o que a oração concerne. O restante da mensagem, a parte em que o Tema é desenvolvido, é denominado *Rema*.

Na língua inglesa a estrutura temática está relacionada à ordem das palavras: o Tema é indicado por ocupar a primeira posição na sentença. O autor ressaltava que a primeira posição não corresponde à definição de Tema, mas apenas à maneira como ele se realiza em inglês. Assim, na língua japonesa, por exemplo, o Tema é indicado pela partícula *-wa*, que indica que tudo que a precede integra o Tema da oração. A fórmula recorrentemente utilizada por Halliday para identificar o Tema é: “*I’ll tell you about...*”.

Nas orações declarativas do inglês, o Tema natural, não-marcado, corresponde ao sujeito da oração; nas interrogativas, ao pronome interrogativo (*who, when, etc.*), ou ao elemento finito do predicado mais o sujeito (*can’t you, can you, is it, etc.*); nas exclamativas, ao pronome exclamativo; nas imperativas, ao verbo. As imperativas diferem das restantes, uma vez que o verbo é regularmente o tema não-marcado da oração. Em outros tipos de construções, há a possibilidade de colocar o verbo em primeira posição precisamente para lhe dar status temático altamente marcado (ex.: *forget it I never shall*) (1994:47). Segundo Halliday, o Tema se estende desde o primeiro elemento da oração até o primeiro elemento que tenha alguma função na transitividade (participantes, ou circunstâncias, ou processos), sendo este denominado “*Topical Theme*”.

Danes (1974) nota que a seleção e a distribuição do Tema dificilmente são fortuitas, não-motivadas, sem conexão estrutural com o texto. Halliday (1994) afirma que a escolha do Tema

constitui o fator mais importante para o desenvolvimento do texto como um todo e tende a diferir conforme o gênero textual. Assim, numa narrativa, por exemplo, é provável que o mesmo participante exerça a função de Tema por um trecho do discurso, enquanto num diálogo, há geralmente alternância de Temas, especialmente entre os pronomes de 1ª e 2ª pessoas, representando falante e ouvinte.

2. Halliday delimita a unidade de informação baseando-se em aspectos prosódicos. Uma unidade de informação corresponde a um grupo tonal⁷, podendo ser mais ou menos do que uma oração gramatical. No caso não-marcado há a equivalência: uma unidade de informação (estrutura informacional): um grupo tonal (estrutura prosódica): uma oração (estrutura gramatical).

A unidade de informação serve para estruturar os componentes do discurso de acordo com o status que o falante dá a cada elemento. Uma parte é o *Novo* – o que o ouvinte está sendo convidado a interpretar como novo, inesperado, imprevisível, ou importante. A outra parte corresponde à *informação dada* – o que está sendo apresentado como já conhecido pelo ouvinte.

A maneira como a estrutura informacional é realizada é natural (não-arbitrária) em dois sentidos: (i) O Dado tipicamente precede o Novo; (ii) o Novo é assinalado por proeminência tônica. Segundo Halliday (1994:296), a localização do elementoônico (pé, palavra ou sílaba) identifica o *Foco informacional* (“*the culmination of what is New*”).

O Dado é geralmente um elemento anafórico⁸, referindo-se a algo que já foi apresentado no contexto verbal ou situacional. Há, no entanto, instâncias em que ocorre a elipse do elemento dado. Por isto, Halliday postula que a unidade de informação consiste de um elemento novo obrigatório, mais um elemento dado opcional.

O elemento que está sendo tratado como Dado pode ser recuperável porque já foi mencionado, mas esta não é a única possibilidade: pode ser algo que está presente na situação, ou algo que não está presente, mas que o falante apresenta como Dado, por alguma razão (retórica, estilística). Da mesma forma, o Novo pode ser algo que não foi mencionado, mas pode ser também algo inesperado ou importante, previamente mencionado ou não. Assim, uma forma de novidade é a ênfase assinalada pela entonação e/ou pela estrutura, como nos seguintes exemplos, em que os

⁷ Baseando-se em Cagliari (1978, apud Ilari, 1992:19), que analisou estas noções no português, Ilari define *pé* como “*unidade rítmica que se inicia com uma sílaba forte*” e *grupo tonal* como “*uma seqüência de pés organizada em torno de uma sílaba tônica*”. No exemplo citado a seguir, barras duplas delimitam o grupo tonal e as simples assinalam os pés: //*Todos /foram à/ praia/*// . A partir do grupo tonal, há três opções entoacionais que têm relevância semântica: (i) “*a de localizar em diferentes pontos do texto falado os limites dos sucessivos grupos tonais*”; (ii) “*a de localizar a sílaba tônica em diferentes pontos do grupo tonal*”; (iii) “*a de associar à sílaba tônica diferentes configurações melódicas*” (Ilari, 1992:19). As duas primeiras opções são, segundo Ilari, “*os expoentes fônicos do recurso textual que Halliday chama de Informação*” (1992:20).

⁸ Há um número de elementos na língua que são inerentemente Dados, no sentido de que não são interpretáveis a não ser por menção anterior ou referência a algum aspecto da situação. Este é o caso dos elementos anafóricos e dos dêiticos. Tipicamente estes elementos não carregam o Foco informacional e se isto acontece, eles expressam contraste.

pronomes pessoais teoricamente Dados são apresentados como Novos: *You can go if you like, I'm not going; It's you who were to blame* (Halliday, 1994:298,301).

Os principais critérios que Halliday utiliza para determinar o que é Novo são portanto: (i) fonológicos (proeminência tônica); (ii) contextuais (contexto precedente verbal ou não-verbal).

Halliday postula que há uma relação semântica entre a estrutura informacional e a temática: *“other things being equal, a speaker will choose the Theme from within what is Given and locate the focus, the climax of the New, somewhere within the Rheme”* (1994:299). Apesar de correlacionadas, estas estruturas não são idênticas: Tema corresponde ao que o falante escolhe como ponto de partida, e Dado, ao que o ouvinte já conhece ou a algo a que ele tem acesso. Halliday ilustra a distinção, dizendo que o Tema é *“speaker-oriented”*, enquanto o Dado é *“listener-oriented”*. As duas estruturas são é claro selecionadas pelo falante, mas este leva em conta o conhecimento partilhado com o ouvinte. Há ainda uma outra diferença fundamental entre estas funções: Halliday analisa a estrutura temática na oração, dividindo os seus componentes em Tema e Rema. A estrutura informacional, por seu lado, não é um sistema da oração, seu domínio é o da unidade informacional que tipicamente, mas não necessariamente, corresponde a uma oração.

As estruturas temática e informacional constituem juntas recursos para estruturar a oração como mensagem, dando-lhe um status particular em relação ao discurso que a cerca e ao contexto situacional. Tema e Foco Informacional são os aspectos básicos do que os lingüistas da Escola de Praga denominavam *“Functional Sentence Perspective”* (FSP). As distinções entre Dado x Novo e Tema x Rema remontam a V. Mathesius (1939, In: Danes, 1974). Uma outra distinção, no âmbito da FSP, foi introduzida por Firbas (1964, In: Danes, 1974): a de *dinamismo comunicativo* – extensão com que um determinado elemento da sentença contribui para o desenvolvimento da comunicação. Segundo Firbas, o Tema é constituído pelos elementos da sentença que carregam o menor grau de dinamismo comunicativo. Assim como Halliday (1994), Firbas dissocia Tema e Dado. Danes (1974) nota que o que justifica a distinção é a existência de casos em que o Tema não veicula informação que pode ser recuperada pelo contexto verbal ou situacional, embora estes constituam a minoria e sejam percebidos como marcados. Mathesius (1939, In: Danes, 1974) afirma igualmente que o ponto de partida da expressão, definido como informação dada, geralmente corresponde ao Tema.

Ilari (1992) analisa a expressão entoacional da articulação Tema-Rema nas orações do português e ressalta a dificuldade de tratar a Tematização e a Informação como subsistemas independentes. Para identificar o Tema e o Rema, Ilari aplica o *“teste da pergunta natural”* – toda oração pode ser vista como uma resposta a uma pergunta virtual, o conteúdo compartilhado entre a oração afirmativa e a pergunta virtual corresponde ao Tema, e o não compartilhado, ao Rema. O

autor estabelece ainda que a expressão do Rema está relacionada à proeminência entoacional, ou seja, ao Foco. Assim, em sua análise, ele considera a possibilidade de o Rema ocupar a primeira posição na sentença, embora sua posição natural, não-marcada, seja a final, conforme ilustrado nos exemplos a seguir (Ilari, 1992:42):

(Quem chamou o Paulo?) (pergunta virtual)

Maria chamou o Paulo (Rema identificado pela entonação)

O Paulo, chamou-o ***Maria*** (Rema indicado pela posição final e pela entonação)

Neste último exemplo, vemos que há um segmento separado por pausa no início da oração (*O Paulo*) que é retomado por um pronome átono. Ilari denomina este segmento deslocado de *Tópico* ou *Antitópico*, conforme ele se encontre no início ou no final da sentença. Embora o termo *Tópico* seja utilizado por alguns autores como sinônimo de Tema⁹, acreditamos que devemos manter distintas as duas noções para dar conta destes casos em que um segmento da oração é deslocado de sua posição natural. Koch (2000), que também analisa o português, ocupa-se justamente dessas construções segmentadas, nas quais um elemento do enunciado é destacado e colocado em posição inicial ou final. Em seu estudo, porém, o *Tópico* e o *Antitópico* são designados como “*Temas marcados*”. Estes exercem um papel importante na construção da coerência de um texto, bem como no processamento pragmático-cognitivo do sentido. Segundo a autora, “*o falante seleciona um elemento (...) que deseja ativar ou reativar na memória do interlocutor e sobre o qual seu enunciado deverá lançar nova luz*” (2000:138). Estas noções são úteis para nossa pesquisa, aplicando-se aos casos em que o elemento em questão corresponde ao referente do verbo.

3. A unidade fundamental que Halliday (1994) considera é a oração, abrangendo também o período. Contudo, o autor não se restringe a isso, tratando também das relações denominadas “*não-estruturais*”, ou seja, aquelas que são determinadas pela semântica do discurso e que permitem a interpretação de um item em dependência de outro que integra o mesmo texto. Estas relações que não se limitam à sentença, podendo em um texto ligar elementos de qualquer extensão a qualquer distância, são englobadas sob o título geral de *coesão*. Halliday divide a coesão nos seguintes itens: referência, elipse, conjunção e organização lexical. Nós nos deteremos nas duas primeiras.

⁹ Halliday (1994) utiliza o termo *Tópico* para designar o Tema que é simultaneamente o elemento Dado. Para Dik (1989, In: Neves, 2001), o *Tópico* constitui a entidade acerca da qual uma informação é oferecida no discurso.

A questão da referência é extremamente relevante para a presente pesquisa. Conforme nota Halliday, as relações anafóricas criam coesão, ligando duas passagens do texto em uma unidade coerente. O autor estabelece que a textualidade é obtida parte pela estrutura, parte pela coesão. Por exemplo, se um pronome e seu referente se encontram no mesmo período, há textualidade pela relação estrutural entre as orações e a coesão meramente adiciona textualidade. Se, por outro lado, não há relação estrutural (i.e. o pronome e seu referente estão distantes um do outro, em dois períodos diferentes), a coesão se torna o elemento de ligação exclusivo, fundamental para a criação da textualidade. O autor postula ainda que se um referente se mantém na estrutura narrativa do discurso, não importa se ele é mencionado por um nome lexical, por um pronome, ou não mencionado (elíptico), ele irá produzir coerência: “*Whatever requires the listener or reader to store and retrieve what has gone before has this effect*” (Halliday, 1994:312).

A elipse constitui da mesma forma um meio de se estabelecer coesão. A elipse do sujeito é de particular interesse para a nossa pesquisa. No inglês, os contextos de sujeito elíptico são reduzidos, uma vez que esta língua praticamente não dispõe de flexão número-pessoal para identificar o referente do verbo e, além disso, necessita do sujeito explícito para indicar o modo (i.e. a ordem do sujeito e do elemento finito do predicado distingue uma oração declarativa de uma interrogativa). Halliday (1994:94) nota, no entanto, que o modo também pode ser expresso pela entonação, e o sujeito do verbo pode ser interpretado com referência tanto ao contexto verbal (“elipse anafórica”), como ao situacional (“elipse exofórica”). Há então a possibilidade de omissão do sujeito no inglês, como ilustram os exemplos: *Seen Fred?* (“Have you seen Fred?”) (elipse exofórica); *No must be away.* (“He must be away.”) (elipse anafórica).

Além disso, Halliday (1994:94) postula que há um sujeito não-marcado para cada tipo de oração. Assim, se uma oração com sujeito elíptico pode ser interpretada como uma oferta ou uma afirmação, o ouvinte identificará o sujeito como sendo de 1ª pessoa (exs.: *Carry your bag?*; *Met Fred on the way here.*); e se a oração corresponde a uma pergunta ou a uma ordem, o sujeito será identificado como 2ª pessoa (exs.: *Seen Fred?*; *Play us a tune.*).

Para a questão: “*Which subject is to be understood if none is present?*”, o autor tem, portanto, uma resposta funcional: o ouvinte preencherá o sujeito: “*there’s rarely any misunderstanding since the listener operates on the basic principle of all linguistic interaction – the principle that what the speaker says makes sense in the context in which he is saying it*” (1994:95)¹⁰.

¹⁰ Halliday (1994) estabelece uma diferença importante entre a referência e a elipse: enquanto esta constitui uma relação no nível léxico-gramatical e tem alcance restrito, aquela corresponde a uma relação semântica e pode abranger longas passagens de um texto.

Dik (1997) estabelece que quando uma entidade é introduzida no discurso de um modo definido e específico, ela estabelece um referente discursivo, isto é, adquire o status de existência no mundo criado pelo discurso, podendo ser referida por meios anafóricos que criam coerência. Segundo o autor, Tópicos novos não são introduzidos apenas para figurarem em uma única oração, mas para estarem disponíveis, como Tópicos dados, ao longo de uma unidade discursiva¹¹, ou até mesmo ao longo de todo o discurso. De acordo com a importância da entidade no discurso, haverá um maior ou menor grau de continuidade tópica. No exemplo citado pelo autor: *Once upon a time there was an ugly duckling* (1997:422), sabemos que não só a próxima sentença, mas toda a história diz respeito ao pato em questão.

Dik (1997:411-412) ressalta ainda que a informação pragmática compartilhada entre falante e ouvinte exerce um papel fundamental na produção e na interpretação do discurso. O autor postula que o falante tem uma intenção comunicativa que é parcialmente codificada no conteúdo da expressão lingüística, sendo o grau de explicitação determinado pelo conhecimento partilhado entre os interlocutores. Este engloba não só o conhecimento lingüístico e não-lingüístico de que os interlocutores dispunham antes do evento comunicativo (“*long-term knowledge*”), como também o conhecimento derivado do próprio evento comunicativo e da situação em que este toma lugar (“*short-term knowledge*”). Falante e ouvinte constroem um modelo mental dinâmico (“*Discourse Model*”¹²), constantemente atualizado, do que está sendo transmitido no discurso e que vai auxiliar em sua interpretação. “*This is why later parts of the discourse may rely upon contextual information which has been provided in the preceding context. Following context, however, also plays a crucial role: each step in the discourse typically project structure onto the discourse that is yet to come; each move creates an expectation pattern about the further steps that are yet to be taken*” (Dik, 1997:413).

Vemos que a Lingüística Funcional toma o discurso essencialmente como um processo cuja organização é semântica, mais do que formal. Halliday (1987) postula que um texto não se compõe de sentenças, mas é realizado através delas. Dik, por sua vez, nota que “*natural language users do not speak in isolated sentences or clauses, but combine these into longer and more complex stretches for which we may use the general cover term ‘discourse’*”(1997:409).

¹¹ Segundo Dik (1997), o evento discursivo pode ser subdividido nos seguintes tipos de unidades: (i) de um ponto de vista interacional: atos de fala, seqüência de atos de fala, turnos, seqüência de turnos; (ii) de um ponto de vista do seu conteúdo: proposições, predicacões e episódios discursivos.

¹² Dik (1997:413) define “*Discourse Model*” como: “*a kind of book-keeping of what the discourse has so far achieved (...) and what is probably yet to come*”.

1.3 UMA ANÁLISE FUNCIONAL DAS CATEGORIAS SUJEITO E VERBO E DA CONCORDÂNCIA

É possível analisar o sujeito, o verbo, e a relação que se estabelece entre ambos por meio da concordância, nas três funções propostas por Halliday (1994): (i) Interpessoal, (ii) Experiencial, (iii) Textual. O argumento externo do verbo corresponde ao: (i) Sujeito; (ii) Participante: Ator, etc.; (iii) Tema x Rema; Dado x Novo. O verbo é o constituinte que funciona como: (i) Predicador na estrutura modal; (ii) Processo na estrutura transitiva: material, mental, etc.; (iii) Tema x Rema e Dado x Novo nas estruturas temática e informacional. Analisando a concordância sob a perspectiva funcional, podemos estabelecer que a flexão número-pessoal constitui: (i) de um ponto de vista gramatical: a marca formal que explicita a relação sintática entre o sujeito e o verbo; (ii) de um ponto de vista semântico: o elemento que identifica o referente do verbo; (iii) de um ponto de vista discursivo: informação nova ou dada.

Em sua análise da articulação Tema-Rema na língua portuguesa, Ilari (1992:10) critica as fórmulas que são utilizadas pela tradição escolar como definições intercambiáveis de sujeito: “*quem faz a ação*”; “*o termo com o qual concorda o predicado*”; “*o assunto da oração*”. O autor nota que a não-coincidência entre Agente, sujeito e Tema, que torna a oração marcada, justifica a distinção entre os três níveis de análise. A definição de sujeito dada por Dik (1981, In: Pezatti, 1992) e por Halliday (1994) é semântica. Dik o define como a entidade tomada como ponto de partida, o centro natural de atenção a partir do qual um estado de coisa é apresentado, e Halliday, como o elemento apresentado pelo falante como responsável pela validade do que está sendo enunciado. Chafe (1979), por sua vez, observa que, nos casos menos marcados, há uma correspondência entre sujeito e informação dada. Castilho (no prelo) nota igualmente que há geralmente coincidência entre sujeito, informação dada e Tema da oração, enquanto o verbo normalmente representa o Rema.

Na quinta parte de Benveniste (1976), intitulada “O homem na língua”, o autor faz uma reflexão sobre a estrutura das relações de pessoa no verbo. Conforme nota Benveniste, “*o verbo é, com o pronome, a única espécie de palavra submetida à categoria da pessoa*” (1976:247). É comum classificar as formas de conjugação do verbo conforme a sua referência à pessoa, distinguindo-se três no singular, no plural, e eventualmente no dual. O autor discorre sobre uma questão importante: “*poderá existir um verbo sem distinção de pessoa?*” (1976:248). Cita então algumas línguas cujas formas verbais são indiferentemente marcadas quanto à pessoa e ao número (coreano, línguas paleo-siberianas), notando, no entanto, que estas dispõem de pronomes pessoais que cumprem a função de distinguir as pessoas. Benveniste deduz então que “*não parece que se conheça uma língua dotada de um verbo no qual as distinções de pessoa não se marquem de uma*

ou de outra maneira nas formas verbais. Podemos concluir, pois, que a categoria da pessoa pertence realmente às noções fundamentais e necessárias do verbo” (1976:250)¹³.

Se o falante do português popular tende a não concordar o verbo com o sujeito sintático da oração, conforme estabelece uma de nossas hipóteses, uma análise funcional prediz que a noção de sujeito será explicitada mediante outras estratégias.

Nos capítulos 19 e 20 de *Principles of Linguistic Change: Internal Factors*, Labov (1994) vai justamente promover uma reflexão sobre o peso dos fatores funcionais na explicação da variação e da mudança lingüísticas. Os próprios títulos dos dois capítulos, respectivamente, “The Overestimation of Functionalism” e “The Maintenance of Meaning”, refletem a posição de Labov, que pode ser resumida através de duas proposições: (i) a escolha de uma ou outra variante lingüística não é determinada pela necessidade de preservar a informação, mas sim influenciada por fatores de natureza mecânica e estrutural; (ii) quando uma mudança lingüística se completa, o sistema realiza reajustes internos para preservar a informação. O autor postula basicamente que os efeitos funcionais são primordialmente as conseqüências das escolhas lingüísticas e não as suas causas.

Labov resume a orientação funcionalista através da seguinte proposição: “*The function of language is for the speaker (or writer) to communicate meaning to the listener (or reader)*” (1994:548). É esta interpretação da língua essencialmente como um instrumento de comunicação que vai determinar a maneira como os funcionalistas vêem a variação e a mudança lingüísticas.

No que se refere estritamente à variação lingüística, Labov (1994) nota que a hipótese funcionalista estabelece que quando os falantes se comunicam, eles levam em consideração a informação de que seus interlocutores dispõem e, se houver a possibilidade de escolha entre duas formas alternativas, eles favorecem aquela que transmite o sentido da maneira mais eficiente e efetiva. Labov, no entanto, questiona esta formulação, argumentando que não podemos captar a intenção do falante, quando este profere uma sentença e, mesmo se conhecêssemos suas motivações, isto não implicaria que elas teriam determinado a escolha de uma ou outra variante lingüística.

No que diz respeito à variação que envolve mudança lingüística, especialmente a de natureza fonética que constitui o tema central de Labov (1994), os funcionalistas postulam que esta é

¹³ Benveniste (1976:252) destaca ainda que a 1ª e a 2ª pessoas não estão no mesmo plano da 3ª: “a 3ª pessoa é, em virtude de sua própria estrutura, a forma não-pessoal da flexão verbal”. A 3ª pessoa, vista como a não-pessoa, justifica nossa utilização do termo “forma verbal não-marcada” para designar a forma de 3ª pessoa do singular com sujeito no plural de 1ª ou 3ª pessoas (exs.: *Nós fala; Eles fala*). Além disso, relaciona-se com a questão da indeterminação no português, que elege precisamente a forma de 3ª pessoa do singular ou do plural, para explicitar a noção de sujeito indeterminado.

diretamente afetada pela necessidade de preservar o sentido, conforme ilustrado nos trechos de Kiparsky e de Labov abaixo:

“[T]here is a tendency for semantically relevant information to be retained in surface structure...It characteristically originates as a blocking of rules in environments in which their free application would wipe out morphological distinctions on the surface.” (Kiparsky, 1982, In: Labov, 1994:554)

“Whenever a single segment is variably deleted, it will be deleted less often if it is a separate morpheme.” (Labov, 1971, In: Labov, 1994:554)

Contudo, revisando sua posição anterior, Labov (1994) vai questionar esta perspectiva, notando que diversas mudanças fonéticas apagam flexões, levando à perda de distinções e a um aumento concomitante da homonímia. O autor lança então a seguinte questão: *“se a comunicação da informação não determina a forma da variação e da mudança lingüísticas, o que a determina?”* (1994:550). Uma possível resposta está na posição dos neogramáticos: o mecanismo da mudança lingüística opera de um modo mecânico, sem considerar o sentido ou as necessidades comunicativas da sociedade. Labov vai sustentar uma posição semelhante: enumerando diversos estudos de variação sincrônica no âmbito fonológico e morfológico, o autor mostra que esta opera independentemente da necessidade de transmitir informação. Entre estes, são citados dois estudos sobre a concordância verbal de 3ª pessoa do plural: um no espanhol (Poplack, 1980, In: Labov 1994:557-561), e outro no português (Scherre & Naro, 1991, In: Labov, 1994:566-568).

Poplack constata dois efeitos contrafuncionais: (i) a consoante /n/, marca verbal de 3ª pessoa do plural, é mais freqüentemente omitida nos casos em que constitui a única marca de plural, ou seja, em verbos regulares, como em *mandan*, do que em verbos irregulares, como em *son* – observação que se coaduna com o princípio da saliência fônica¹⁴ levantado no estudo de Naro & Lemle (1977) e que constitui, segundo Labov, uma prova de que a ação de fatores funcionais pode ser anulada por fatores fonéticos; (ii) o /n/ dos verbos é omitido com maior freqüência em estruturas de sujeito posposto do que nas de sujeito anteposto. Para Labov, um argumento funcional prediria justamente o oposto, estabelecendo que se o sujeito ocorre antes do verbo e carrega marca de plural, a marca de plural no verbo é redundante, mas se o sujeito é posposto, a marca de plural do verbo torna-se necessária.

¹⁴ O princípio da saliência fônica estabelece que quanto mais proeminente a marca flexional do verbo, maior a tendência de o falante retê-la.

Scherre & Naro, por sua vez, verificam que o princípio do paralelismo formal (i.e. tendência de marcas levarem a marcas e zeros levarem a zeros) atua de maneira decisiva na aplicação da concordância verbal: se um verbo não-marcado precede a ocorrência, a probabilidade de realização da concordância diminui e, se esta é precedida por um verbo marcado, a chance de realização da regra aumenta. Os autores analisam também a relação entre a realização da concordância e a presença/ausência de marca de plural no último elemento do sintagma nominal (SN) correspondente ao sujeito, comprovando o mesmo efeito do paralelismo: se a marca está presente, a concordância verbal tende a ser aplicada, e se não o está, a regra é menos frequentemente realizada. Segundo Labov, este estudo não deixa dúvida de que a restrição dominante na variação da marcação dos verbos do português não é funcional, já que foi verificada uma tendência não de preservar a informação semântica, mas de usar as marcas quando são redundantes e omiti-las quando são necessárias.

Labov conclui que, no âmbito da variação fonológica e morfológica, a hipótese funcionalista postula uma tendência de os falantes escolherem uma ou outra variante, de modo a preservar a informação. Mas os resultados citados mostram que os principais fatores que determinam estas escolhas são mecânicos: condicionamento fonológico ou simples repetição da estrutura precedente. Todos os fenômenos lingüísticos examinados pelo autor são interpretados como casos de variação estável. O tratamento da mudança lingüística, por outro lado, evidencia que Labov não nega totalmente os argumentos funcionais, como ilustram os trechos abaixo:

“When language changes, its information carrying capacity is often threatened; but in the long run, most languages do preserve their means of conveying information, more or less, by one route or another. Though speakers may not behave wisely and thoughtfully as they choose one variant or the other, somehow the system does react.” (1994:568)

“It is important to note that in the course of language evolution, change does go to completion, and variable rules have become invariant. When this happens, there is inevitably some other structural change to compensate for the loss of information involved.” (1991:223)

“Principle of Structural Compensation”: *“When the rate of deletion of a meaningful feature of a language increases, the frequency of features that redundantly carry this meaning will increase”* (1994:604)¹⁵.

¹⁵ Para explicar como se dá este reajuste interno que serve para preservar o sentido, Labov (1994) levanta a hipótese do *“Probability Matching”*: uma espécie de mecanismo de aprendizagem que indica que os falantes são capazes de perceber as formas lingüísticas mais frequentes, empregando-as em sua fala.

Labov cita então alguns estudos de mudança fonológica, morfológica e sintática que comprovam tais formulações. Entre estes, dois trabalhos sobre a concordância verbal no espanhol (Hochberg, 1986,a,b, Cameron, 1992, In: Labov, 1994) que evidenciam que a supressão de /s/, marca verbal da 2ª pessoa do singular, é compensada pelo uso do pronome pessoal *tu*, de modo a não comprometer a informação.

Todas estas considerações são essenciais para nossa análise, uma vez que sugerem uma série de conceitos, idéias e observações que permanecem válidas como hipóteses de trabalho. No capítulo seguinte, apresentamos alguns estudos que tratam da variação na aplicação da regra de concordância verbal no português do Brasil.

CAPÍTULO 2:

A CONCORDÂNCIA VERBAL NA LITERATURA LINGÜÍSTICA

Há na literatura lingüística um grande número de estudos que se ocupam da análise da concordância verbal (CV) no português do Brasil (PB). Tais obras indicam que a variação entre aplicação x não-aplicação da regra de concordância constitui um fenômeno recorrente na fala de pessoas de diferentes regiões brasileiras, sendo encontrada: em variedades populares rurais (cf. Assis, 1988; Baxter & Lucchesi, 1993; Mello, 1996; Pereira & Rodrigues, 2003), em variedades populares urbanas (cf. Lemle & Naro, 1977; Naro & Lemle, 1977; Naro, 1981; Guy, 1981; Bortoni-Ricardo, 1985; Rodrigues, 1987, 1989, 2000) e no português culto (cf. Camacho, 1993; Rodrigues & Campos, 2002)¹⁶.

A maior parte destes estudos incide sobre a regra de CV de 3ª pessoa do plural. Bortoni-Ricardo (1985) e Rodrigues (1987, 1989, 2000) incluem, além disso, a análise da CV de 1ª pessoa do plural. Baxter & Lucchesi (1993) descrevem somente a de 1ª pessoa do singular. Por fim, Assis (1988) e Mello (1996) contemplam todas as pessoas do paradigma.

É importante notar também que grande parte destes lingüistas utiliza a metodologia quantitativa, buscando identificar os fatores lingüísticos e sociais que favorecem ou desfavorecem a aplicação da regra de CV. Alguns procuram ainda explicar a origem da variação, baseando-se em pressupostos da Sociolingüística, Gerativismo, Crioulística, bem como na história social do PB.

Faremos uma apresentação de alguns destes estudos, ressaltando os aspectos que de algum modo contribuíram para o desenvolvimento da presente pesquisa. Nos itens 2.1 e 2.2 consideramos, respectivamente, os fatores lingüísticos e os fatores sociais que condicionam a realização/não-realização da regra de CV, tanto na 1ª (P1), quanto na 3ª pessoa do plural (P3). O item 2.3 versa sobre as hipóteses que foram formuladas para explicar a origem da variação de concordância no PB.

¹⁶ Há ainda os estudos de Scherre & Naro (1993) e de Naro & Scherre (1999) cujos *corpora* incluem dados de falantes populares e de falantes escolarizados. O *corpus* utilizado por Bortoni-Ricardo (1985) também contém dados de um pequeno número de informantes escolarizados.

2.1 FATORES LINGÜÍSTICOS

Há um consenso, entre lingüistas, no que se refere à relação entre a saliência fônica da oposição entre as formas verbais do singular e do plural e a CV de 3ª pessoa. Embora existam diferenças entre as categorizações utilizadas para medir esta saliência, todos os estudos que consideram este fator (cf. Lemle & Naro, 1977; Naro & Lemle, 1977; Naro, 1981; Guy, 1981; Bortoni-Ricardo, 1985; Rodrigues, 1987; Naro & Scherre, 1999; Pereira & Rodrigues, 2003) concluem que ele exerce uma influência decisiva sobre a concordância: quanto maior for a diferença entre as formas verbais do singular e do plural, maior será a probabilidade de aplicação da regra e, por outro lado, quanto menor essa diferença, menor a chance de realização da concordância.

É interessante destacar o estudo de Naro (1981), que analisa a saliência fônica através de dois fatores distintos: (i) grau de diferenciação material entre as formas verbais do singular e do plural; (ii) intensidade dos segmentos fonéticos que realizam a oposição: não-acentuados (menos salientes) x acentuados (mais salientes).

Guy (1981) nota que das duas dimensões da saliência fônica propostas por Naro (1981), somente a primeira constitui um fator condicionante da CV. Quanto à separação entre segmentos acentuados e não-acentuados, Guy acredita que esta seja um reflexo da distinção entre: (i) verbos cuja forma no plural pode ser convertida para o singular através do processo fonológico de desnasalização (exs.: *come/comem; fala/falam; faz/fazem*); (ii) verbos em que, mesmo se houver a possibilidade de ocorrer desnasalização, esta não é suficiente para apagar a marca de plural (exs.: *dá/dão; comeu/comeram; falou/falaram; é/são; disse/disseram*). Em seu estudo, o autor propõe que há uma intersecção entre duas regras variáveis distintas: uma regra puramente fonológica que suprime a nasalidade das vogais átonas finais, atuando não somente em verbos, mas em outras classes de palavras (exs.: *homem>homi; ontem>onti*), e uma regra de natureza morfossintática referente à CV de 3ª pessoa do plural.

No que se refere à CV de 1ª pessoa do plural, também há um fator condicionante de natureza fonética: a acentuação da forma verbal. Bortoni-Ricardo (1985) e Rodrigues (1987) comprovam que formas paroxítonas (ex.: *falamos*) tendem a favorecer a aplicação da regra de concordância, e formas proparoxítonas (ex.: *falávamos*) a desfavorecem. Rodrigues constata que, quando a forma verbal esperada é proparoxítona, a não-concordância é praticamente categórica no português popular.

Um outro fator amplamente reconhecido como condicionante da CV em P3 é a posição e a distância do sujeito em relação ao verbo (cf. Lemle & Naro, 1977; Naro & Lemle, 1977; Naro, 1981; Guy, 1981; Bortoni-Ricardo, 1985; Rodrigues, 1987; Camacho, 1993; Naro & Scherre, 1999;

Rodrigues & Campos, 2002; Pereira & Rodrigues, 2003). Tais estudos atestam que o sujeito imediatamente anteposto ao verbo tende a favorecer a concordância, e o sujeito antecedendo o verbo, mas dele separado por elementos intervenientes, a desfavorece. O sujeito posposto corresponde à categoria menos saliente, uma vez que o elemento determinante da concordância segue o elemento determinado, tornando menos óbvia a falta de CV¹⁷.

Em conexão com o fator posição do sujeito, Rodrigues & Campos (2002) consideram também sua função discursiva. Nas orações que apresentam a ordem sujeito-verbo, o sujeito exerce a função de Tema, e o verbo, a de Rema. Já nas orações com posposição do sujeito, este assume o papel discursivo de Rema, e o verbo, o de Tema. Neste segundo contexto, a aplicação da regra de CV é desfavorecida, uma vez que o SN posposto geralmente não é percebido como sujeito pelo falante.

Partindo de uma perspectiva funcional, que considera o sujeito como “*ponto de partida*” ou “*ponto de vista*”, Pezatti (1992:143) propõe que “*a ordem VS (verbo-sujeito) configura sentenças absolutamente destituídas de sujeito*”, ou seja, sentenças nas quais não é possível detectar a relação sujeito/predicado e tópico/comentário, constituídas apenas de predicado ou comentário. Baseando-se em Keenan (1976, In: Pezatti, 1992), que sugere que o status do sujeito é gradual, a autora postula que o português elege três traços para configurar o sujeito prototípico: um traço semântico (ser ativo) e dois traços sintáticos: estar na posição pré-verbal e ser controlador da CV. Contudo, o argumento das estruturas VS não dispõe destes traços: ele é “*paciente ou neutro (inativo), ocupa a posição pós-verbal e nem sempre controla a concordância, mesmo na variedade culta*” (1992:143). Os exemplos por ela arrolados comprovam tal hipótese: *existe leis aí; falta elementos técnicos; bateu seis horas; saiu as brigas em família*. Pezatti conclui que estes argumentos não são intuitivamente considerados sujeitos pelo falante justamente por não possuírem as características prototípicas de sujeito.

Camacho (1993) chega a conclusões semelhantes. O autor postula que a CV é obrigatória ao aplicar-se para trás, isto é, quando o sujeito está anteposto ao verbo, exercendo a função de Tópico, e optativa ao aplicar-se para frente, quando o SN controlador da concordância está posposto ao verbo. O autor (1993:635) constata também que grande parte das ocorrências de não-concordância corresponde a sentenças com um verbo de natureza existencial, como nos exemplos a seguir: *inclusive parece que existe...leis; a não ser quando era as estudantes que iam para a escola; o fato de em um segundo morrer duas mil pe ahn duzentas mil pessoas apavora*.

¹⁷ Bortoni-Ricardo (1985) também analisa a atuação do fator posição do sujeito na variação da CV de 1ª pessoa do plural, comprovando o mesmo efeito constatado na 3ª pessoa. Dos autores acima citados, somente Guy (1981) não confirma a hipótese de que o sujeito anteposto distante do verbo desfavorece a CV de P3.

Segundo Camacho, os verbos que são usados nessas construções são apresentacionais, do tipo existencial (exs.: *ser, existir*), intransitivos (exs.: *viver, vir, ficar*) e verbos de ação e locativos, todos eles exercendo a função pragmática de introduzir um argumento novo no discurso. O autor observa que, de uma perspectiva formal, o argumento novo exerce a função de sujeito do verbo, e de um ponto de vista funcional, estas construções se caracterizam por duas marcas: “*usar-se uma sintaxe do tipo verbo-em-primeiro*” e “*retirar do sujeito lógico o controle da concordância gramatical dos verbos existenciais apresentacionais, neutralizando-as completa ou parcialmente, como parece ser o caso do português*” (1993:636).

A presença/ausência do sujeito também é um fator comumente apontado para explicar a variação da CV (cf. Lemle & Naro, 1977; Naro & Lemle, 1977; Naro, 1981; Guy, 1981; Bortoni-Ricardo, 1985; Rodrigues, 1987, 1989; Naro & Scherre, 1999, Rodrigues & Campos, 2002; Pereira & Rodrigues, 2003). A hipótese estabelecida e geralmente confirmada é a de que o sujeito oculto favorece a aplicação da regra e o sujeito explícito condiciona o uso de formas verbais não-marcadas. Este fato é interpretado sob um ângulo funcional: conforme destaca Rodrigues (1987, 1989), se o sujeito não se encontra na frase, a flexão não é redundante e tende a ser utilizada, pois estabelece a relação entre o verbo e o sujeito extra-sentencial; por outro lado, a perda da informação causada pelo apagamento da desinência número-pessoal dos verbos é compensada pelo uso do pronome lexical.

Bortoni-Ricardo (1985) e Rodrigues (1987, 1989) comprovam ainda que, embora o sujeito oculto seja um fator condicionante da aplicação da regra de CV tanto na 1ª quanto na 3ª pessoa, a ausência do sujeito pronominal *nós* tem um peso maior do que a de *eles/elas* na realização da concordância.

Para explicar esta tendência, Rodrigues (1987, 1989) introduz o conceito de “saliência de pessoa”, pressupondo que, para o falante popular, a noção de pessoa é mais saliente do que a de número gramatical. A autora propõe que a presença física do pronome *nós* estabelece a diferença entre 1ª e 3ª pessoas, conseqüentemente o seu índice de retenção é mais elevado do que o de *eles/elas*, e a sua realização tende a levar ao apagamento das marcas de flexão número-pessoal dos verbos, porque esta informação já está contida no pronome e é interpretada como redundante pelo falante. A ausência do pronome *nós*, por outro lado, favorece decisivamente o uso de uma forma flexionada de 1ª pessoal do plural.

Já os pronomes *eles/elas* determinam apenas a oposição de número em relação a *ele/ela* e, como a noção de plural é menos saliente, os falantes se inclinam a não utilizar nem o pronome explícito, nem a forma verbal flexionada, desde que o contexto esclareça o referente. Exemplo:

Doc. *E as crianças cê acha que acostumaram aqui ficam bem?/ Inf. Costumô porque porque nasceu aqui né? e tá sendo criada aqui então costumô (Rodrigues,1989:551).*

Rodrigues (1989:552-3) destaca ainda que “*assim como no sintagma nominal a marca de pluralidade pode restringir-se, no vernáculo popular, ao determinante, ou primeiro elemento da frase nominal, numa sucessão de frases de mesmo sujeito plural, este pode ser referido na 1ª delas e apagado nas subseqüentes, mesmo quando associado a formas verbais do singular*”. Exemplo: *Nós não tinha medo de sucuri, não tinha medo de onça, não pensava em nada, só pensava em Deus* (Rodrigues,1989:551).

A autora analisa os casos em que a concordância não-padrão ocorre quando o pronome lexical não é explícito. A ausência de marca por completo teoricamente comprometeria o conteúdo informacional da frase e caracterizaria a ação de forças de natureza contrafuncional. Contudo, em todas as ocorrências examinadas, a ausência simultânea do pronome sujeito e das marcas de pluralidade no verbo não constitui um impedimento para a veiculação da informação, uma vez que o sujeito está fisicamente presente nas orações que antecedem essas formas. Isto leva a autora a introduzir um outro fator condicionante da aplicação/não-aplicação da regra de CV no português popular: a categoria [+ - referido] e o princípio da não-repetição: se o sujeito já foi mencionado, as formas verbais que o sucedem não apresentam flexão número-pessoal; o ouvinte já sabe que se trata de formas de plural, por conseguinte o falante elimina a redundância, não repetindo a informação gramatical velha de número e pessoa, mas mantendo apenas o dado novo de natureza lexical e modo-temporal.

Rodrigues & Campos (2002) levantam um outro aspecto importante em relação ao sujeito oculto: a localização de seu referente (conferir também: Naro & Scherre, 1999; Pereira & Rodrigues, 2003). As autoras verificam que o índice de aplicação da CV de 3ª pessoa do plural é maior: (i) nos contextos em que o referente está distante da forma verbal com sujeito zero do que nos casos de referente próximo, (ii) nos contextos em que o antecedente do verbo figura no próprio turno do informante do que naqueles em que está na fala do interlocutor. Ambos os achados são encarados sob uma perspectiva discursiva: o primeiro está ligado à própria construção da coerência no discurso, e o segundo é explicado pelo distanciamento entre a fala dos interlocutores e pelo fato de “*cada turno constituir uma espécie de ilha sob o ponto de vista da progressão tópica*” (2002:6).

Guy (1981) e Pereira & Rodrigues (2003) também examinam a relação entre presença/ausência do sujeito de 3ª pessoa do plural e CV. Contudo, a hipótese de que o sujeito oculto favorece a concordância não é confirmada pelos dados destes estudos.

Guy (1981) estabelece diversas categorias para o sujeito não-explícito: (i) mencionado no discurso, (ii) na pergunta do documentador, (iii) não-especificado ou não-visível (indeterminado). O

autor comprova que, enquanto os dois últimos casos favorecem decisivamente o uso de formas verbais marcadas, o primeiro não se mostra tão significativo. Para Guy, o pressuposto de que o sujeito oculto condiciona a concordância em P3 é um artefato que provém do tratamento indiscriminado de contextos distintos.

Pereira & Rodrigues (2003) igualmente destacam que sujeitos determinados e indeterminados não podem ser tratados em conjunto, uma vez que os últimos impõem uma taxa maior de plural¹⁸. Em sua análise, as autoras descartam os casos de casa vazia de sujeito indeterminado, considerando apenas aqueles em que há a presença do pronome pessoal *eles* que, mesmo não remetendo a um SN no plural expresso no contexto precedente, constitui um elemento determinante da concordância. Este estudo verifica que, embora o falante popular tenda a não aplicar a CV, ele emprega com maior frequência formas verbais no plural para expressar a noção de sujeito indeterminado. Pereira & Rodrigues observam ainda que o princípio da saliência fônica atua muito mais fortemente em estruturas de sujeito indeterminado do que nas de sujeito determinado. Nas ocorrências de sujeito oculto, a CV somente é favorecida nos contextos em que o referente se encontra distante da forma verbal com sujeito zero.

Há ainda um último fator lingüístico relevante que foi levantado no estudo de Scherre & Naro (1993) para explicar a variação da CV de 3ª pessoa do plural: o princípio do paralelismo formal – tendência de marcas levarem a marcas, e zeros levarem a zeros. Este fator atua tanto no nível intra-sentencial (a presença de marca de plural no último elemento do SN sujeito favorece a CV, e a ausência de marca a desfavorece), como no nível extra-sentencial (o padrão de marcação ou não-marcação de pluralidade dos verbos tende a ser mantido numa seqüência de ocorrências). Guy (1981) também verifica que havendo marca de plural em todos os elementos do SN, a CV tende a ser aplicada.

2.2 FATORES SOCIAIS

Vários foram os fatores sociais considerados para explicar a variação da CV: (i) idade do informante (cf. Naro, 1981; Guy, 1981; Bortoni-Ricardo, 1985; Rodrigues, 1987, 2000); (ii) sexo (cf. Naro, 1981; Guy, 1981; Bortoni-Ricardo, 1985; Rodrigues, 1987, 2000); (iii) escolaridade (cf. Rodrigues, 1987, 2000); (iv) procedência (cf. Naro 1981; Rodrigues, 1987, 2000); (v) orientação

¹⁸ A maior parte dos estudos sobre a CV de P3 não considera este critério semântico. Lemle & Naro (1977) analisam este aspecto através do fator tipo de sujeito (definido x indefinido) e também constatam que sujeitos impessoais tendem a favorecer a marca de plural. Já Naro (1981) e Guy (1981) acabam excluindo as ocorrências de casa vazia de sujeito indeterminado em suas análises.

cultural (cf. Naro, 1981); (vi) exposição à mídia (cf. Bortoni-Ricardo, 1985); (vii) rede social (cf. Bortoni-Ricardo, 1985); (viii) situação social em que o discurso é produzido (cf. Lemle & Naro, 1977; Naro & Lemle, 1977; Guy, 1981); (ix) significado social das variantes (cf. Bortoni-Ricardo, 1981). Em nossa exposição, discorreremos apenas sobre os mais significativos para nossa pesquisa.

De todos os fatores mencionados, o mais importante é certamente a idade, uma vez que um estudo da distribuição dos padrões de variação nas diversas faixas etárias de uma determinada comunidade permite estabelecer se a CV se encontra em processo de mudança, ou se constitui uma variação estável. Naro (1981) constata que os informantes mais velhos são mais propensos a aplicar a CV de 3ª pessoa do plural do que os jovens, evidenciando um processo lento de eliminação da regra. Guy (1981) contesta esta hipótese: examinando o mesmo *corpus*, o autor não acha uma diferença significativa nos padrões de CV de P3 conforme a idade do informante e conclui que não há evidências de uma mudança em progresso. Rodrigues (1987, 2000) verifica que a probabilidade de não-aplicação da CV, tanto na 1ª quanto na 3ª pessoa, é maior entre os adultos jovens do que entre os adultos de meia idade e os mais velhos. Este resultado reforça a hipótese de uma mudança em progresso, mas pode também ser explicado meramente em termos das pressões sociais exercidas sobre os indivíduos: os adultos jovens, sob a influência de seus pares, tenderiam a manter os hábitos lingüísticos da fase de aquisição de seu vernáculo, já sobre os adultos mais velhos pesam as pressões da sociedade mais ampla e por isso estes tenderiam a utilizar formas verbais de maior aceitação social. Bortoni-Ricardo (1985) constata a tendência oposta: os falantes jovens apresentam uma taxa muito maior de concordância em P1 e em P3 do que os adultos, fato que a autora atribui à influência da educação formal.

O sexo é freqüentemente apontado como um outro fator que tem influência na realização da CV. Rodrigues (1987, 2000) verifica que, na 3ª pessoa do plural, este fator não é significativo: a não-concordância está igualmente disseminada entre falantes de ambos os sexos. Já na 1ª pessoa do plural, há uma diferença relevante: os homens mostram uma maior freqüência de concordância em sua fala do que as mulheres. Bortoni-Ricardo (1985) igualmente comprova que as mulheres empregam mais freqüentemente formas verbais não-padrão do que os homens, diferença esta que se manifesta nas duas pessoas, ainda que seja mais significativa em P1 do que em P3. Para explicar esta diversidade nos padrões de concordância conforme o sexo, Rodrigues (1987, 2000) considera o papel e a rede social do indivíduo: enquanto as mulheres da classe social desfavorecida dispõem de menos opções de trabalho, sendo em geral donas de casa, com um contato normalmente restrito às pessoas da comunidade na qual residem, os homens têm mais possibilidades profissionais e um

maior grau de interação com falantes de outros grupos sociais, e conseqüentemente tornam-se mais conscientes do significado social da ausência de CV¹⁹.

Rodrigues (1987, 2000) considera a influência da escolaridade nos padrões de realização de CV. Seu estudo comprova que este fator é extremamente significativo para a concordância de 1ª pessoa do plural: os falantes que passaram pela escola, ainda que tenham permanecido nela pouco tempo, apresentam uma taxa muito maior de aplicação da concordância do que os falantes de escolaridade nula. Com relação à 3ª pessoa do plural, porém, há uma tendência generalizada de não-aplicação da regra que independe do nível de escolaridade do falante.

A autora também analisa o fator procedência dos informantes, verificando que os falantes nascidos na capital paulistana tendem a aplicar com maior freqüência a regra de CV do que os falantes procedentes de outras regiões (Noroeste de São Paulo, Norte do Paraná, Norte de Minas Gerais, Sul da Bahia, e região Nordeste). Este fator se mostra relevante tanto na 3ª, quanto na 1ª pessoa, mas tem peso maior em P1.

O último fator extralingüístico de que trataremos concerne o significado social da realização/não-realização da CV. Ainda que não estudemos a avaliação social da concordância em nossa pesquisa, consideramos relevante mencioná-la aqui, na medida em que pode sugerir pistas para explicar os padrões de concordância de 1ª e 3ª pessoas encontrados na fala dos informantes do meio rural e do meio urbano cujos dados compõem o *corpus*. De certa forma, todos os fatores sociais sobre os quais discorreremos – idade, sexo, escolaridade, procedência – entrecruzam-se com o significado social que as variantes aplicação e não-aplicação da CV possam assumir em um determinado contexto.

Os estudos de Rodrigues (1987, 2000) e de Bortoni-Ricardo (1985) são extremamente significativos para nossa pesquisa, uma vez que mostram que os índices de CV diferem substancialmente de acordo com a pessoa considerada: as freqüências de aplicação da regra constatadas por estas duas autoras são de, respectivamente, 54% e 56% na 1ª pessoa do plural, e 29% e 35% na 3ª pessoa do plural. Para Rodrigues (1987, 2000), o uso de uma forma verbal não-marcada tem, nos grandes centros urbanos, um significado social diverso, conforme o sujeito seja de 1ª ou de 3ª pessoa do plural. De fato, a ausência de CV em P3 parece não ter assumido um significado social negativo generalizado, sendo encontrada até mesmo na fala de informantes cultos, nos contextos em que é menos perceptível (cf. Camacho, 1993; Rodrigues & Campos, 2002). A não-concordância com um sujeito de 1ª pessoa do plural, por outro lado, é saliente não só de um

¹⁹ Naro (1981) e Guy (1981) não acham uma diferença significativa nos padrões de CV conforme o sexo do falante. O mesmo acontece com o fator procedência analisado por Naro (1981), fato que o autor atribui à irregularidade da distribuição dos informantes do *corpus*.

ponto de vista lingüístico, mas também de um ponto de vista social. Segundo Rodrigues (1987, 2000), a não-realização da CV em P1 é estigmatizada, identifica o falante como analfabeto e/ou procedente da zona rural e o encaixa no grupo de mais baixa posição na escala social: “*se, em tese, o falante popular tende a eliminar as marcas de flexão verbal, tal tendência é refreada, principalmente em P4 (1ª pessoa do plural), pelo modelo de realização dos grandes centros urbanos, que sugerem o emprego de formas verbais padrão, marca característica da língua popular urbana em contraste com a variedade rural*” (1987:215)²⁰.

Bortoni-Ricardo (1985) concebe o PB como um continuum: no extremo esquerdo, encontram-se os vernáculos rurais isolados; no extremo direito, o padrão urbano das classes altas; e no meio, o “rurbano” – conjunto de variedades não-padrão faladas pela classe baixa urbana, analfabeto ou semi-escolarizada, na sua maior parte proveniente da zona rural, ou pela população que vive em áreas rurais modernizadas. Conforme a autora, a estratificação de um determinado dado lingüístico pode ser: (i) “gradual” se ele ocorre nos dois extremos do continuum; ou (ii) “descontínua” se o elemento lingüístico é submetido a estigmatização, definindo uma distinção profunda entre fala urbana e rural. A ausência de concordância verbo-sujeito é em alguns casos gradual, mas nos contextos em que é mais perceptível e menos freqüente é descontínua.

Bortoni-Ricardo (1981) constata que até mesmo a não-realização da CV de 3ª pessoa do plural pode veicular um significado social negativo principalmente nos contextos em que é mais saliente. A autora realiza um exame de reações subjetivas à falta de CV em P3 entre dois grupos de falantes: (i) universitários e (ii) alunos do curso supletivo noturno de 1º grau, constatando que a estigmatização está presente no primeiro grupo, mas não no segundo. Bortoni-Ricardo conclui então que a avaliação negativa da não-concordância é, em parte, decorrente do ensino sistemático da variedade normativa na escola. Isto explica por que os indivíduos de alto nível de escolaridade tendem a reagir negativamente em relação ao uso de formas verbais não-marcadas, embora as empreguem eventualmente em sua fala. Já “*entre as classes menos favorecidas, que se conservaram durante muitas décadas em zonas rurais ou em regiões interioranas onde havia poucas oportunidades culturais, a concordância não-padrão, perfeitamente instituída (...) ainda não é saliente*” (1981:94). A autora ressalta ainda que, mesmo se este falante chegar a perceber o significado social negativo da ausência de CV, através do convívio com os falantes urbanos ou de um período de escolarização, ele tenderá a utilizar com menos freqüência as formas verbais padrão, uma vez que as assimilou tardiamente.

²⁰ Mello (1996) também nota que o vernáculo urbano, apesar de apresentar formas não-marcadas ou fonologicamente reduzidas, mostra-se mais próximo da língua padrão do que o vernáculo rural. Assim, a autora postula ser improvável um falante urbano dizer *nós é feio*, preferindo a forma *nós como feio*, em que o verbo simplesmente perde o -s final.

2.3 AS HIPÓTESES QUE EXPLICAM A ORIGEM DA VARIAÇÃO DA CONCORDÂNCIA VERBAL

Neste item, tratamos das seguintes questões: qual a origem da variação da CV no português popular? Esta variação implica mudança lingüística? Caso a resposta seja afirmativa, qual seria a direção desta mudança? Há, por um lado, os estudos que sugerem que as regras de CV estão sendo ou foram adquiridas pelos falantes populares e, por outro lado, os que postulam que elas estão sendo eliminadas de sua gramática. Há os que acreditam que a variação na aplicação da concordância tem uma origem européia e os que a atribuem ao contato lingüístico ocorrido no Brasil colonial, decorrente da massiva presença de falantes não-nativos do português.

Assis (1988), Baxter & Lucchesi (1993) e Mello (1996) analisam a realização da regra de CV em variedades do português popular registradas em comunidades rurais brasileiras. Estes autores notam que a simplificação do sistema flexional de número-pessoa é mais intensa em variedades rurais, predominando o uso de apenas duas formas verbais – a da 1ª pessoa do singular e uma forma não-marcada coincidente com a da 3ª pessoa do singular para todas as outras pessoas do discurso. Assis (1988:67) cita alguns exemplos que comprovam a simplificação do paradigma verbal, que se reduz à oposição de duas pessoas (falante exclusivo x falante não-exclusivo), sem apresentar distinções quanto ao número: *eu alembro de todos; nós pranta é mamona; eles tira o quarto ano.*

Esta redução das flexões número-pessoais dos verbos, pode, entretanto, ser ainda mais drástica em comunidades rurais isoladas de origem étnica africana, como apontam os trabalhos de Baxter & Lucchesi (1993) e de Mello (1996). Ambos os estudos destacam a possibilidade de utilização de uma única forma verbal não-marcada correspondente a da 3ª pessoa do singular para todas as pessoas do discurso. Exemplos: *eu sabe fazê isso* (Mello,1996:125); *eu fala* (Baxter & Lucchesi, 1993:61).

Mello (1996) nota que esta tendência é ainda mais forte entre a população mais velha da comunidade de Helvécia (Bahia), uma vez que os falantes jovens são mais expostos à educação formal e a outras variedades do português vernacular brasileiro. Segundo a autora, isto constituiria uma evidência de que os estágios mais antigos do português popular foram mais reestruturados, ou seja, predominava apenas uma forma verbal para todas as pessoas e só num segundo momento, os falantes adquiriram marcas de flexão da 1ª pessoa do singular.

Um fato interessante apontado pela autora é que crianças que adquirem o PB como primeira língua também fazem uso dessa simplificação morfológica, utilizando a forma de 3ª pessoa do singular para qualquer sujeito, independentemente da pessoa ou número. A alta frequência de uso

desta forma na fala do adulto levaria a sua generalização para todas as pessoas, sendo que a forma de 1ª pessoa do singular só seria aprendida pela criança num estágio avançado de aquisição da língua.

Este dado é relevante porque ilustraria, segundo Mello, como deve ter sido incompleta a aprendizagem do português como segunda língua pelos escravos no Brasil que, devido ao pouquíssimo contato com falantes nativos do português, não teriam adquirido as marcas de flexão verbal. As formas monoverbais poderiam ter sido cristalizadas em seu falar e teriam permanecido na fala de seus descendentes, antes que tivessem um maior contato com uma comunidade mais ampla de falantes do português. Esta seria uma das razões para a presença de um paradigma verbal reduzido em variedades específicas do PB.

A outra explicação apontada para o uso de uma forma verbal para todas as pessoas é a existência de um pidgin ou crioulo no Brasil colonial. Mello propõe que nas plantações, por exemplo, é provável que o código de comunicação utilizado para interação não só entre escravos, mas entre escravos e brancos, era um pidgin baseado no português. Esse pidgin teria uma só forma verbal baseada no infinitivo ou na 3ª pessoa do singular. Para reforçar tal suposição, Mello (1996:128-129) cita exemplos retirados de crioulos de base portuguesa da África, em que também ocorre a utilização de uma forma neutra baseada na 3ª pessoa do singular para todas as pessoas: *Tempu ki nôsu nasêu* (“No tempo em que nós nascemos”); *Ezi matava zenti* (“Eles matavam a gente”); *Yô skiseu* (“Eu esqueci”).

Essas formas monoverbais encontradas em algumas variedades rurais do PB e em crioulos de base portuguesa são utilizadas como evidências de que o português vernacular brasileiro teria passado por um processo de reestruturação parcial e simplificação, ocasionado pelos fenômenos de contato lingüístico, pidginização e criouliização, e aquisição imperfeita da segunda língua por povos de origem não-européia. Uma posterior descrioulização teria ocorrido, devido ao contato com variedades mais padronizadas, mas o paradigma verbal reduzido teria resistido em comunidades isoladas do Brasil.

Este ponto de vista é compartilhado por Guy (1981) e por Baxter & Lucchesi (1993:61), que argumentam da mesma forma em favor de um processo prévio de descrioulização, que levou a uma mudança de um sistema sem regra de CV para um sistema em que esta regra se aplica de forma variável (exs.: *eu fala/falo*; *nós fala/falamo*; *vocês, eles fala/falam*). Para Guy, o princípio da saliência fônica governando a variação seria exatamente o que se esperaria encontrar em uma variedade lingüística descrioulizada.

É interessante notar que esta visão de que a mudança na CV se dá em direção à aquisição da regra também está presente num plano sincrônico. Bortoni-Ricardo (1985) postula que os falantes

provenientes de regiões rurais que se transferem para regiões urbanas, sendo expostos à língua padrão, tendem a recuperar as marcas de CV e as mais salientes são as primeiras a serem adquiridas.

A perspectiva oposta, isto é, a de que as regras de CV estão sendo perdidas, é evidenciada nos estudos de Lemle & Naro (1977), Naro & Lemle (1977) e Naro (1981). Estes autores concluem que a regra de CV de 3ª pessoa do plural encontra-se em pleno processo de mudança e caminha para a extinção na variedade popular²¹. O surgimento desta mudança “natural” seria essencialmente de natureza fonológica, decorrente da desnasalização de vogais átonas finais; já a implementação da mudança no sistema dar-se-ia de acordo com o princípio da saliência, estendendo-se mais fortemente aos contextos em que a falta de concordância é menos perceptível.

Cabe notar que a hipótese de que a origem da variação da CV no PB está ligada a fatores de natureza fonética tem sido levantada por vários lingüistas²². Naro & Scherre (1993,1999), por exemplo, argumentam igualmente que a variação na CV tem um componente puramente fonológico: quando o plural *comem* se reduz ao singular *come*, a única diferença é a perda da nasalidade da vogal não-acentuada final. Tal regra atuou no latim clássico, em textos medievais portugueses e existe no PB, operando também sobre outras classes de palavras. Por outro lado, quando *comeram* é substituído por *comeu*, já há uma substituição de desinência. Esta redução morfológica da concordância é, segundo eles, “*um desenvolvimento mais tardio, criado a partir da generalização da redução fonológica*” (1993:442).

Admitindo, portanto, que a mudança lingüística que envolve a CV tenha se iniciado na fonologia, Naro & Scherre (1993,1999) defendem a hipótese de que o fenômeno de concordância variável no PB provém de Portugal, sendo resultado da deriva – tendências inerentes às línguas românicas, e indo-européias de maneira geral, em direção à uniformização dos padrões morfológicos, com a permanência apenas das formas “irregulares” mais salientes.

Naro & Scherre (1999) de fato constatam variação na aplicação da regra de CV de P3 em oito textos pré-clássicos²³ e em variedades populares do português europeu falado moderno. Comparando os textos medievais portugueses com dados da fala de informantes do Rio de Janeiro,

²¹ Rodrigues (1987) também observa que, enquanto a CV em P1 constitui uma variação estável, os padrões de CV em P3 evidenciam uma tendência de eliminação da regra.

²² Bortoni-Ricardo (1981) igualmente acredita que a redução flexional no português tenha sido implementada pela desnasalização.

²³ A média de instâncias de não-concordância nos textos pré-clássicos é, porém, muito baixa: cerca de 1%. Mattos e Silva (1991:70-71), que estuda o português arcaico, também comprova que a CV apresentava-se já facultativa em alguns contextos, sendo condicionada pela posição e distância do sujeito em relação ao verbo, tipo de sujeito (simples ou composto) e interpretação singular ou plural do sujeito. Os exemplos por ela utilizados demonstram que a variação não se restringia aos contextos menos salientes: *o sabor do luxurioso e o prazer he vermen e fedor; e aos braados yeo o bispo e todos aqueles.*

comprovam que os mesmos fatores lingüísticos que condicionavam a aplicação da regra de CV com sujeitos de 3ª pessoa do plural no passado se mostram igualmente atuantes no presente. São eles: saliência fônica, posição do sujeito, e traço semântico do sujeito – o traço [+humano] tende a favorecer a aplicação da regra, e o traço [-humano] a desfavorece. Para os autores, houve apenas uma mudança na freqüência de uso de formas verbais não-marcadas e no peso relativo dos fatores condicionantes.

Naro & Scherre (1993,1999) acreditam que as condições de pidginização endêmica e aquisição da língua portuguesa pelos povos de origem não-européia somente teriam acelerado tendências prefiguradas no sistema lingüístico do português. Esta visão é oposta a de Mello (1996), que postula que a deriva não é suficiente para explicar a simplificação das formas verbais do português vernacular brasileiro. As evidências apresentadas por Mello (1996:247) contra um simples condicionamento fonológico são os verbos irregulares, que comprovam uma mudança morfológica, já que a forma de 3ª pessoa do singular é utilizada como forma não-marcada para todas as pessoas, com exceção da 1ª, em muitas variedades do PB (exs.: *Nós foi; Eles é*). Para Mello, como vimos, foram sobretudo os processos de contato e mudança lingüística, devidos à adoção do português por milhões de indígenas²⁴ e africanos, que levaram à simplificação da morfologia verbal do PB.

Guy (1981) da mesma forma destaca que, em grande parte dos casos, a variação da marcação de pluralidade dos verbos de 3ª pessoa se deve à regra fonológica de desnasalização, reconhecendo que esta constitui uma tendência antiga na língua portuguesa. Mas a variação da CV é vista como uma inovação: para o autor, não há evidências históricas de que as regras de concordância verbal tenham sido variáveis ou condicionadas por fatores de natureza morfológica e/ou sintática. Já em línguas crioulas, a redução da complexidade morfológica, a perda de flexões e o nivelamento de paradigmas constituem fenômenos comuns. Daí a variação na concordância ser tomada como prova de que houve criouliização no PB. Estes traços do tipo crioulo teriam permanecido no português popular, ainda que atenuados pelo processo de descriouliização.

Cabe, por fim, mencionar alguns estudos que partem de uma perspectiva gerativista para explicar a redução das flexões número-pessoais do PB. Duarte (1993), Roberts (1993) e Galves (1993) propõem que a simplificação da morfologia verbal não se deve a uma mudança fonológica, mas à reorganização do sistema pronominal. Com a substituição de *tu/vós* por *você/vocês* e a

²⁴ No contato lingüístico ocorrido no Brasil colonial, os lingüistas geralmente consideram importante apenas a influência dos africanos, uma vez que os índios utilizavam uma língua geral de base tupi para comunicar-se com os brancos e com outras tribos indígenas, não tendo necessidade de utilizar um pidgin ou crioulo baseado no português. Silva Neto (1963, In: Mello, 1996:245) cita, no entanto, um exemplo retirado de um livro jesuítico retratando o português falado pelos índios que mostra que também entre eles a CV poderia ter sido simplificada: *Nós não coiesse esso gente*.

covariação entre os pronomes *a gente* e *nós*, o paradigma verbal passa a ter somente quatro formas possíveis, resultantes da combinação entre um traço sintático (pessoa) e um traço de número (plural):

- (i) [+ pessoa] [- plural], ex.: *canto*;
- (ii) [+ pessoa] [+ plural], ex.: *cantamos*;
- (iii) [- pessoa] [+ plural], ex.: *cantam*;
- (iv) [- pessoa] [- plural], ex.: *canta* (forma totalmente não-marcada).

Galves (1993) considera esta concordância “enfraquecida” tanto de um ponto de vista morfológico (não contém 2ª pessoa), como de um ponto de vista semântico (há a possibilidade de interpretação da 3ª pessoa do singular como indeterminada). Sob o ângulo gerativista, este empobrecimento da morfologia verbal tem como consequência uma mudança paramétrica no PB: a progressiva perda do sujeito nulo, uma vez que a CV não seria mais capaz de licenciá-lo e de permitir a recuperação de seu referente²⁵.

Todos os estudos aqui apresentados são fundamentais para o desenvolvimento da presente pesquisa, fornecendo subsídios não só para a análise das ocorrências do *corpus*, como também para a explicação dos padrões de concordância encontrados no português popular.

²⁵ Em seu estudo sobre a gramaticalização, Castilho (1997:44) faz especulações sobre este fenômeno, afirmando que “no Português do Brasil, a confluência de fenômenos tais como a simplificação da morfologia verbal e as alterações do quadro dos pronomes pessoais anuncia a cliticização destes ao verbo, e o possível surgimento de prefixos número-pessoais”.

CAPÍTULO 3: O *CORPUS* SOB ANÁLISE

3.1 O PROJETO FILOGIA BANDEIRANTE

Para estudar os padrões de realização da regra de concordância verbal, utilizamos parte do *corpus* de língua falada do Projeto Filologia Bandeirante. Iniciado em 1998, sob a coordenação do Prof. Dr. Heitor Megale, tal Projeto está realizando uma incursão em núcleos rurais de quatro estados brasileiros, localizados na rota dos bandeirantes²⁶ (os caminhos percorridos à exaustão em busca do ouro das minas, nos séculos XVII e XVIII), com o objetivo de pesquisar dados lingüísticos que possam favorecer “*a busca da ancianidade do português do Brasil*” (Megale, 1998, In: Oliveira, 2000:223). Pesquisadores das Universidades Federais de Minas Gerais, de Goiás, do Mato Grosso e da Universidade de São Paulo têm trabalhado a fim de constituir um *corpus* de língua falada – a partir da gravação de inquéritos junto a falantes idosos radicados em localidades fundadas por bandeirantes, ou que surgiram em conseqüência de sua passagem – e de língua escrita – através da busca de documentos produzidos no período das bandeiras, ou que tenham algum vínculo com os bandeirantes.

Tal *corpus* é extremamente significativo para a história sociolingüística do português do Brasil. Na região do ciclo do ouro, o movimento das bandeiras adquiriu um caráter colonizador. Segundo Megale (comunicação pessoal, In: Cohen, 2000:212), “*não há rota de colonização mais importante (do que as bandeiras) dentro do país, que tenha tido uma movimentação demográfica de maior repercussão*”. Alguns autores acreditam que as bandeiras tiveram um papel decisivo na penetração e posterior fixação da língua portuguesa nos sertões brasileiros²⁷. Abreu (1883, 1930, In: Mendes, 2000:66), por exemplo, afirma que “*foi no lombo de burro e pata do boi (...) que os sertões foram progressivamente conquistados e integrados na nova cultura que os descobridores trouxeram ao Brasil. E dessa cultura, que partia do litoral, fazia parte integrante a língua portuguesa*”.

²⁶ O termo *bandeirante* se refere a um conjunto heterogêneo de pessoas que participavam de expedições ao sertão: portugueses, mamelucos, índios, negros e estrangeiros. O movimento das bandeiras é eminentemente paulista, com maior atuação dos moradores da Vila de São Paulo.

²⁷ Mattos e Silva (2001:286) observa que historiadores acreditavam que a língua geral (de base tupi) seria, até o século XVIII, a língua familiar da área cultural que mais tarde seria denominada “Brasil caipira”.

Documentando a realização oral de pessoas da zona rural que adquiriram a língua há 60 anos ou mais, o Projeto Filologia Bandeirante espera colher traços da língua portuguesa antiga que tenham permanecido ao longo dos três séculos que nos separam do início do processo de colonização e povoamento dessas áreas (Cohen *et al.*, 1997:79). Tem, portanto, “*o audacioso objetivo de buscar as raízes do português do Brasil falado nas trilhas das bandeiras, interpretando-o como uma modalidade conservadora, que reflete o falar seiscentista-setecentista trazido pelos colonizadores e implantado no Brasil pelos sertanistas (Castilho, 1997)*” (Oliveira, 2000: 223). O estudo do padrão de concordância verbal nessa variedade específica é relevante, pois reúne elementos lingüísticos que, se comparados a dados do português arcaico, podem colaborar na verificação da hipótese da ancianidade do português do Brasil.

3.2 PERFIL SOCIAL DOS INFORMANTES

O *corpus* que constitui o objeto principal de análise é composto de 15 inquéritos de informantes idosos (em média 78 anos) de ambos os sexos, analfabetos ou semi-escolarizados, nascidos e criados na zona rural dos estados de São Paulo e de Minas Gerais, na área correspondente às trilhas das bandeiras paulistas. Em sua grande maioria, são pessoas já aposentadas que trabalharam somente em atividades típicas do campo, ligadas à lavoura e à criação de gado.

Os dados obtidos na fala destes 15 informantes são comparados com os encontrados por Rodrigues (1987), em seu estudo sobre a concordância verbal no português popular de São Paulo. O *corpus* estudado por Rodrigues (1987) compreende 40 inquéritos de falantes de três faixas etárias – (i) 20 a 35 anos, (ii) 36 a 50 anos, (iii) mais de 50 anos – e de ambos os sexos, residentes na zona urbana: a favela do Carombé (zona norte da cidade de São Paulo). São adultos de baixa ou nula escolaridade, em sua maior parte, provenientes da zona rural, não só do estado de São Paulo, como de outras regiões do Brasil.

3.3 O PORTUGUÊS POPULAR

Os informantes cujas características são descritas no item anterior são autênticos falantes do português popular. Esta noção de variedade lingüística popular é construída com base em parâmetros sócio-culturais. Conforme a definição de Rodrigues & Campos (2002:1), “*português*

popular brasileiro, ou não-padrão, ou vernáculo brasileiro é aquela variedade de Português do Brasil utilizada por brasileiros do mundo rural ou do mundo urbano, analfabetos ou de baixo nível de letramento, que não têm na língua escrita possível modelo de realização oral”.

O grau de escolaridade constitui, portanto, um parâmetro decisivo para a caracterização do estrato sociolingüístico popular: “*por não terem acesso à escola, que, em tese, constitui instrumento de divulgação de um padrão culto de realização lingüística, tendem a manter os hábitos adquiridos na fase de aquisição e definição de seu vernáculo. Se já nos primeiros anos de vida aprenderam a dominar os mecanismos básicos do português oral, a sociedade não lhes permite, ou deles não exige, no decorrer da vida, alcançar outros estágios no processo de aquisição ou aprendizagem da língua materna, ou certas habilidades de expressão que instituições sociais típicas, como a escola, se encarregam de difundir.*” (Rodrigues & Ferreira Netto, 2000:172).

Segundo Rodrigues (1987), o português popular constitui uma variedade lingüística eminentemente de cunho oral, não-codificada e não-normalizada, que tende a modificar-se com maior vitalidade, posto que abrange variação em todos os níveis. Embora seja uma variedade de menor prestígio, que pode servir de índice de classificação social, ela cumpre perfeitamente as necessidades comunicativas de seus usuários.

3.4 A COLETA DE DADOS

As entrevistas com os informantes foram realizadas, em 1998, pelos pesquisadores do Projeto Filologia Bandeirante e gravadas em áudio. Embora este seja o método de coleta de dados mais utilizado na Sociolingüística, há alguns inconvenientes. Segundo Labov, o *corpus* ideal para o estudo lingüístico é a fala utilizada em situações cotidianas: “*language as it is used in everyday life by members of the social order, that vehicle of communication in which they argue with their wives, joke with their friends, and deceive their enemies.*” (1991:xiii)²⁸. Este constitui o que Labov (1991:86) denomina *estilo casual*, ou seja, o discurso utilizado em situações informais, quando a atenção do falante não está diretamente voltada para a linguagem²⁹.

O contexto da entrevista, todavia, condiciona um *estilo formal*. Este é o estilo de fala que normalmente ocorre quando o falante está respondendo questões que são formalmente reconhecidas como parte da entrevista (Labov, 1991). No geral, a entrevista constitui um discurso público,

²⁸ Halliday compartilha dessa visão, argumentando que: “*It is in spontaneous, operational speech that the grammatical system of a language is most fully exploited*” (1994: xxiv).

²⁹ Um outro termo utilizado por Labov (1991:208) é *vernáculo* – o estilo em que a mínima atenção é dada ao monitoramento da própria fala.

monitorado e controlado em resposta à presença de um observador externo. Qualquer tipo de observação sistemática de um falante determina um contexto formal em que mais do que uma atenção mínima é prestada a sua linguagem. Ainda assim, a única maneira de obter informações suficientes sobre a fala de uma pessoa é através da entrevista individual gravada, isto é, através da forma mais óbvia de observação sistemática. Daí surge o que Labov (1991:61) denomina de *Paradoxo do Observador*: o nosso objetivo consiste em observar como as pessoas utilizam a linguagem quando não estão sendo observadas. A solução para o Paradoxo do Observador consiste na utilização de diversas técnicas com o objetivo de suscitar o estilo casual na situação de entrevista, sendo este definido como *discurso espontâneo*: “*Spontaneous speech refers to a pattern used in excited, emotionally charged speech when the constraints of a formal situation are overridden (...) Spontaneous speech is defined here as the counterpart of casual speech which does occur in formal contexts, not in response to the formal situation, but in spite of it*” (Labov, 1991:86).

Embora houvesse grandes diferenças sociais entre informante e documentador, as entrevistas do Projeto Filologia Bandeirante foram realizadas com o cuidado de atenuar a formalidade da situação. Para isto, os documentadores adotaram uma série de procedimentos válidos: (i) não disseram ao informante que se tratava de uma pesquisa sobre a língua – já que isto levaria a um aumento do grau de formalidade –, mas sim de uma pesquisa sobre aspectos históricos e culturais da região por onde passaram os bandeirantes; (ii) as entrevistas foram realizadas na própria residência dos informantes, com a presença muitas vezes de seus parentes e vizinhos; (iii) os documentadores não seguiram um roteiro fixo de perguntas, deixando o informante à vontade para falar do que desejasse, aproximando assim a entrevista de uma conversação natural³⁰.

O êxito na obtenção do discurso espontâneo é evidenciado nas longas histórias contadas pelos informantes³¹. Os próprios assuntos sobre os quais discorrem os levam muitas vezes a um alto grau de descontração. Estes giram em torno de: narrativas sobre sua experiência de vida; atividades de lazer, como caça e pesca, além de festas religiosas; lendas e histórias da região; infância (tema recorrente que evidencia grande nostalgia); vida na roça (trabalho na lavoura e na criação de gado; extrema pobreza); família; morte (tema constante nas entrevistas); modernização: estradas, transporte, luz elétrica, acesso a serviços médicos, água encanada, etc.

³⁰ Procedimentos semelhantes foram adotados por Rodrigues (1987). Todos os informantes sabiam que a entrevista estava sendo gravada. Os inquéritos do Projeto Filologia Bandeirante têm em média 1 hora de duração.

³¹ Labov (1991:91) nota que as digressões costumam suscitar a utilização do discurso espontâneo, sendo comuns entre falantes idosos: “*Some older speakers, in particular, pay little attention to the questions as they are asked. They may have certain favorite points of view which they want to express, and they have a great deal of experience in making a rapid transition from the topic to the subject that is closest to their hearts.*”

3.5 NORMAS DE TRANSCRIÇÃO

A partir das gravações em áudio, as entrevistas foram transcritas conforme as normas estabelecidas em Rodrigues & Ferreira Netto (2000) e Ferreira Netto & Rodrigues (2001). Trata-se de uma transcrição convencional ou transcrição ortográfica modificada: leva-se em conta a produção real, utilizando em parte as normas convencionais da ortografia portuguesa. São registradas as realizações fônicas características da fala do informante, tentando respeitar ao máximo a realização oral.

Os morfemas de flexão número-pessoal só são representados quando efetivamente realizados, com possíveis alterações fônicas:

- (i) a desinência número pessoal de 1ª pessoa do plural é grafada *-mos* ou *-mo*, conforme sua realização (ex.: *falamos/falamo*);
- (ii) nas formas verbais de 3ª pessoa do plural, o ditongo nasal decrescente /ãw/ é representado por *am* (exs.: *falam, comeram*); quando se reduz a vogal desnasalada /u/, é representado por *u* (exs.: *falu, comeru*).

Ao final de cada exemplo citado neste estudo, identificaremos entre parênteses o número do inquérito do qual foi retirado e da página em que se encontra nas transcrições.

CAPÍTULO 4: ANÁLISE DOS DADOS

4.1 CONTEXTOS DE VARIAÇÃO DA CONCORDÂNCIA VERBAL NO PORTUGUÊS POPULAR E CRITÉRIOS PARA SELEÇÃO DE OCORRÊNCIAS

4.1.1 CRITÉRIOS DE INCLUSÃO

No exame inicial dos dados, pudemos constatar que existem diferentes contextos que envolvem variação da concordância verbal no português popular. A presente pesquisa contempla o estudo das seguintes estruturas³²:

I – SUJEITOS DE 3ª PESSOA DO PLURAL REPRESENTADOS POR:

a) nome substantivo no singular com um ou mais determinantes no plural, exs.:

- (1) *us pai deli eram* muito pobre (I.10,p.142)³³
- (2) quando *nasce seti homi* ni uma família (I.9,p.131)
- (3) *us dois filhu meu* já *morreu* (I.13,p.192)
- (4) *ficô umas locomotiva* funcionando lá (I.6,p.74)

b) nome substantivo no plural, ex.:

- (5) *os pais* num *dava* tempu não (I.7,p.101)

c) numeral no plural, ex.:

- (6) porque si *pegaru* nu caminhu *us dois* (I.12,p.183)

³² Foram consideradas todas as ocorrências encontradas no *corpus*, com exceção das apontadas em Critérios de Exclusão.

³³ As ocorrências com o verbo *ser* que foram consideradas compreendem uma das seguintes estruturas: (i) sujeito no plural e predicativo no plural (ainda que o plural do predicativo não seja realizado no nível superficial da oração), como em (1); (ii) expressões equativas, em que o verbo *ser* denota a equivalência entre dois substantivos no plural, exs.: *as duas qui vem pra cima é só as duas cuitadinha* (I.2, p.25), *os índio é os negru memu* (I.12,p.178); (iii) voz passiva, ex.: *o dia que eles foram sepultado* (I.1,p.5); (iv) *ser* transitivo circunstancial, exs.: *meus pais eram daqui* (I.1,p.10); *ái eu perguntei a ela...oh: :: pra que que era aqui/aqueles ressalto* (I.6,p.85).

d) dois ou mais nomes substantivos, ou nome substantivo e pronome *ele/ela* (sujeito composto), exs.:

(7) *o nirso e a cida mora em taubaté (I.2,p.23)*

(8) *ela i a fia dela ia pra lá (I.14,p.214)*

e) pronomes pessoais *eles/elas* (explícito ou oculto) com referência determinada, exs.:

(9) *o meus neto vai lá na seti vorta (...) então...é lá qui *elis* estuda (I.2,p.25)*

(10-11) *us vizinhu tamém tinha tamém quatru homi tamém a: :: a: :: só qui *elis* eram bem de vida né?... () mai trabalhavam muito (I.10,p.142)*

f) pronome pessoal *eles* com referência indeterminada³⁴, ou casa vazia do sujeito indeterminado numa seqüência de orações em que figura o pronome *eles*, exs.:

(12) **eles* perguntava ele quanto: :: valia o boi ne/ne/nessas exposição (I.6,p.86)*

(13-14) *carregaru...seis boi de carro i: ::...i duas (manguera)...*eles* foi embora lá em jacaraí (I.2,p.22)*

(15-16) *então *eles*...exporta o leite cru do do do jeito que tira do da: :: do curral...exporta (I.6,p.88)*

g) pronome de tratamento *vocês* (explícito ou oculto) – ainda que *vocês* se refira do ponto de vista semântico e discursivo à 2ª pessoa do plural, sob uma perspectiva gramatical tal pronome exige verbos com marca formal de 3ª pessoa do plural, o que justifica a sua inclusão no estudo da concordância de 3ª pessoa; exs.:

(17-18) *ó *ocêis* já cumeu a carni *cêis* têm qui ruê u ossu tamém já cumeu a carni têm qui ruê u ossu tamém né? (I.14,p.209)*

(19) *eu se eu fosse estudado *cês* tavu perdido comigo (I.5,p.67)³⁵*

h) pronome relativo *que* retomando SN de 3ª pessoa do plural, exs.:

(20) *então *aquelis* *peessoar* qui mexe com criação né? (I.11,p.154)*

³⁴ Tomando por base o estudo de Micheletti (2000), consideramos o pronome *eles* indeterminado, quando não é possível recuperar o seu referente no contexto lingüístico, ou situacional.

³⁵ Foram consideradas quaisquer marcas de plural nos verbos, sejam elas padrão, ou não-padrão como em (19).

- (21) *essis um memu qui saía pregá missões pas pas igrejas né? (I.10,p.148)³⁶*
 (22-23) *tudo é pra rezá mandá celebrá missa pros otro...pros que já morreu pros que tá vivo (I.1,p.9)*

i) outros pronomes, exs.:

- (24) *num tô falandu que você u/ usa o remédio mais muitos usa né? (I.3,p.44)*
 (25) *naqui u governu nem sabia quantus existia nu mundu (I.12,p.173)*

II – SUJEITOS DE 1ª PESSOA DO PLURAL REPRESENTADOS POR:

a) pronome pessoal *nós* (explícito ou oculto), exs.:

- (26-27) *ói nói vamo pra lá se se nu tivé sór quenti nu tivé chovendu nói nu vai na casa di ninguém di parenti (I.10, p.139)*
 (28-31) *nói era muito pobrezinho nu tinha nada...é...tinha nada só usava uma prantinha (I.2,p.19)*

b) pronome pessoal *eu* mais um nome lexical (sujeito composto)³⁷, exs.:

- (32) *trabaiava aqui *eu i us dois fio meu* (I.13,p.192)*
 (33) *eu cu joaquim meu irmão nós vivia remendanu eli (...) *i eu cu joaquim vivia* remendanu eli (I.12,p.180)³⁸*

c) pronome relativo *que* retomando SN de 1ª pessoa do plural, ex.:

- (34) *são tiago hoje é uma cidade boa pra nóis que somo aqui...desse lugarzinho pobre né? (I.1,p.15)*

³⁶ Foram interpretados como sujeitos plurais todos aqueles representados por SNs que comportam marca formal de plural, ainda que seu núcleo apresente um nome morfologicamente singular, como em (20-21).

³⁷ Não foram encontradas ocorrências de sujeito composto de 1ª pessoa do plural com a utilização de dois pronomes (ex.: *ela e eu*), tampouco de 3ª pessoa do plural (ex.: *ele e ela*).

³⁸ Só foram considerados os casos claros de sujeitos compostos ligados pela partícula *com*, isto é, quando há o pronome *nós* no trecho do discurso comprovando que se trata realmente de um sujeito de 1ª pessoa do plural, como ilustra o exemplo (33).

4.1.2 CRITÉRIOS DE EXCLUSÃO

Não se incluem entre as ocorrências selecionadas para a análise os contextos (1-4) abaixo:

- 1) Frases truncadas (anacolutos), ex.:

(35) quato cinco trazi/ (I.4,p.54)

- 2) Alguns verbos no Presente do Indicativo cujas formas do singular e do plural são homófonas: *tem/têm, vem/vêm, contém/contêm*; ex.:

(36) sempre eles vêm aqui (I.2,p.24)

- 3) Respostas em que se repete a forma verbal da pergunta feita pelo documentador, exs.:

(37) Doc. os lugares de perto *manda* os bloco tamém? / Inf. manda (I.8,p.107)

(38) Doc.2 aí *começaro* a falá / Inf.1 aí aí aí começaro a falá né? (I.6,p.79)

- 4) Verbo no plural com casa vazia do sujeito indeterminado (sem referente *eles* no trecho do discurso) – contexto em que a marca de plural do verbo seria condicionada não pelo critério sintático (concordância com o sujeito), mas pelo critério semântico (noção de indeterminação do sujeito); ex.:

(39) eu nu vi mai um dia mataru aí pra cima (I.2,p.21)

Também não foram consideradas as estruturas com verbos impessoais, exemplificadas em (5-10):

- 5) Orações com o verbo *ter* com valor existencial (nas 39 ocorrências encontradas no *corpus*, o uso de verbos não-flexionados é categórico), ex.:

(40) tinha diversas coisa ali dentu daqueli par(qui) ali (I.9,p.117)

- 6) Orações com o verbo *fazer* com o sentido de “completar, atingir” (nas 38 ocorrências encontradas, a concordância padrão é categórica), exs.:

(41) quandu feiz treis meses que eli morreu (I.3,p.36)

(42) tá fazendu uns quarenta anu (I.14,p.218)

7) Orações com o verbo *dar* com o sentido de “haver”³⁹, exs.:

(43) desses terremotu qui dá qui mata (quantia de) pessoa (I.10,p.142)

(44) essas inchenti qui dá (I.10,p.142)

(45) quandu dá uns truvão muito forti (I.13,p.195)

(46) as ota banana pra nói qui num dá mai (I.4,p.49)

8) Verbo *ser* em expressões com o sentido de afirmação e marcador conversacional, exs.:

(47) Doc. sexta feira santa essas coisas era tudo na cidade? / Inf. é (I.4,p.52)

(48) os os negru aqueli tempu era judiadu nu é? (I.12,p.175)

9) Orações impessoais com o verbo *ser* (casos em que a concordância se dá com *isso*, *aquilo*), exs.:

(49) seis sete carro...(uma porção de/) tud/ tudo com oitu boi tudu de oitu boi...de oitu di deiz...é era bunitu demais (I.7,p.98)

(50) mais sabe qui aquele tempu eh os pai era brabu fazia issu mais era bom purqui os filhu tudu obedecia (I.10,p.141)

10) Orações com o verbo *ser*, nas designações de tempo, distância, quantidade, valor – casos em que as gramáticas prescrevem a concordância com o predicativo no plural (de um total de 61 ocorrências, apenas 6 apresentam o verbo no plural), exs.:

(51) é seis dia né? (I.13,p.204)

(52) é novi hora (I.15,p.244)

(53) daqui a são joão (...) são noventa quilômetros (I.6,p.74)

(54) é onze bisneto (I.1,p.8)

(55) era mil réis pur meis (I.10,p.145)

Além dos casos enumerados em (10) acima, não foram incluídos alguns outros contextos em que a concordância poderia ser estabelecida com o predicativo:

³⁹ Cf. exemplos de Borba (1991:366): É que aqui dá muito assaltante; Em terra que eu não vou, feijão dá na raiz.

11) Orações com sujeito singular e predicativo no plural – casos em que as gramáticas prescrevem a concordância com o predicativo no plural (há 12 ocorrências no *corpus*, apenas 2 com verbo no plural), exs.:

(56) u ordenado já era dois i quinhentu (I.13,p.192)

(57-59) família minha é sãu pocus sãu: ::... doi rapai i uma minina (I.9,p.119)

12) Orações com sujeito no plural e predicativo no singular (casos em que a concordância pode se dar com o predicativo no singular), exs.:

(60) e: :: os netos só contando...num é muito não (I.1,p.7)

(61-62) as coisa é um é fogu (I.12,p.181)

13) Expressões equativas denotando equivalência entre um substantivo no plural e outro no singular (contexto em que a concordância pode se dar com o substantivo no singular), ex.:

(63) os parenti da minha mãe era pir... o marido é juventino (de souza) dos reis (I.2,p.24)

Os contextos (14-18) ilustram instâncias desconsideradas por não apresentarem, ou praticamente não apresentarem, variação entre aplicação x não-aplicação da concordância:

14) Orações infinitivas (nas 54 ocorrências encontradas, foi constatado o uso categórico de verbos não-flexionados), exs.:

(64) o resto dei tudo pa eis comê (I.4,p.48)

(65) vam/ vamo levá quarqué coisa pa nós tomá café (I.6,p.80)

15) Sujeito representado pelo pronome de tratamento *a gente* (as 137 ocorrências encontradas evidenciam concordância categórica), ex.:

(66) *a genti* sofreu na roça pa criá os filho né? (I.1,p.4)

16) Expressão cristalizada *vamos* com casa vazia do sujeito (há 19 ocorrências no *corpus* que mostram o uso categórico do verbo no plural), ex.:

(67) mai vamo lá pra dentru eu vô fazê um café pa nós (I.11,p.162)

17) Verbo *ser* em estruturas clivadas, pseudo-clivadas, e na expressão cristalizada *é que* (há 26 ocorrências de tais estruturas no *corpus*, e apenas 1 com verbo flexionado), exs.:

- (68) aqueleli conventu ali foi elis qui fizeru (I.10,p.151)
 (69-70) foi foi nós que fizemo (I.2,p.34)
 (71) é us pai qui mandava nas fia i nus fiio (I.12,p.183)
 (72) então quem vai: :: ganhá o: :: ICM essa coisa tudo sob o leite daqui...são
 essas companhia de fora (I.6,p.88)
 (73) nóis é que tinha que fazê o (nosso) separado (I.3,p.40)
 (74) agora o vossos filho o vossos netu i bisnetu é qui vão tê o que contá
 (I.10,p.141)

18) Verbo *ser* em orações com valor existencial (as 12 ocorrências encontradas evidenciam que a não-concordância é categórica), exs.:

- (75) era pinta sirva canáriu da terra (...) i sabiá (I.9,p.125)
 (76-77) naquele tempo nu tinha instalação nem nada era era aquelas fossa na horta
 né? (I.11,p.158)

Restringimo-nos à análise da concordância do verbo com sujeito formalmente plural, razão pela qual foram excluídos os contextos (19-23) enumerados abaixo, ainda que envolvessem variação na aplicação da regra:

19) Orações com sujeito singular e verbo no plural (encontramos 4 casos na 3ª pessoa, e 1 na 1ª pessoa), exs.:

- (78) decerto *ele* enguiçaram alguma coisinha (I.4,p.46)
 (79) *eu* memu carreguemu (I.15,p.242)

20) Orações em que o sujeito não tem marca formal de plural, embora provavelmente seja plural de um ponto de vista semântico⁴⁰, exs.:

- (80) essis missionáriu qui vinha re/rezá...qui elis ficava muito tempu no na
 igre/aondi elis iam...eli ficava muito tempu na igreja (I.10,p.147)

⁴⁰ Não contamos o número de ocorrências em que isto acontece, mas sabemos que não se trata de um fenômeno incomum: Guy (1981) também verifica a existência de estruturas deste tipo.

21) Sujeito representado por substantivo coletivo no singular (de um total de 173 ocorrências, apenas 7 apresentam o verbo no plural), exs.:

(81) *o pessoal viviam de lavora (I.4,p.47)*

(82) *o povo depois pagava né? (I.1,p.6)*

22) Sujeito representado pelo pronome indefinido *tudo* retomando substantivo coletivo no singular (encontramos 4 ocorrências de verbo no plural e 13 de verbo no singular), exs.:

(83) *esse povo do rio abaixo tudo são parenti meu (I.2,p.24)*

(84) *o pessoal só trabaiava coitado né? era tudo pobre né? *tudo mundo* (I.4,p.47)*

23) Sujeito representado pelo pronome indefinido *tudo* remetendo a um SN de 3ª pessoa do plural, ou casa vazia do sujeito retomando *tudo* (de um total de 66 ocorrências, apenas 6 figuram com verbo no plural), exs.:

(85) *começô *as criança* morrê...morreu tudo no mesmo dia (I. 1, p.4)*

(86) *tudu já foi imbora...*us capelão* qui morava aqui (I.15,p.241)*

(87-88) *meus pai foi tudo enterrado aqui ó avô pai *manhi* foi tudo enterrado aqui (I.8,p.105)*

(89-90) *é gente pobre é é rico é branco é preto de quarqué forma...tudo gosta muito de mim num gosta? (I.1,p.15-16)*

(91-92) *us netu meu aí tão tudo forgadu porque durmi até oitu hora (I.13,p.192)*

(93) *us *otru tudo* lá começaru a gozá di mim (I.10,p.141)*

4.2 A REALIZAÇÃO DA REGRA DE CONCORDÂNCIA VERBAL

Seguindo a metodologia da Sociolinguística Quantitativa, estabelecemos uma série de fatores de natureza lingüística e social que condicionariam a realização/não-realização da regra de concordância verbal. Grande parte deles é baseada em estudos sobre a concordância verbal no português do Brasil. Alguns fatores lingüísticos foram formulados a partir da leitura de textos que discorrem sobre a Lingüística Funcional (Halliday, 1994; Dik, 1997; Chafe, 1979). Em nossa análise lingüística, procuramos contemplar todos os níveis: semântico, gramatical e discursivo.

Após a codificação de todas as ocorrências conforme os fatores lingüísticos e sociais estabelecidos, os dados foram submetidos ao conjunto de programas computacionais Varbrul, o qual nos forneceu:

- (1) a frequência geral de aplicação/não-aplicação⁴¹ da regra de concordância verbal, bem como as frequências associadas aos fatores estabelecidos⁴² (programa Makecell);
- (2) a seleção por ordem de importância dos fatores estatisticamente pertinentes (programa Step Up);
- (3) os fatores descartados como numericamente irrelevantes (programa Step Down);
- (4) a probabilidade de aplicação da regra conforme os fatores lingüísticos e sociais relevantes (programa Ivarb)⁴³ – cálculo que permite uma visão mais apurada dos dados, uma vez que leva em conta a relação entre o conjunto de fatores presentes em um determinado contexto. Os valores de probabilidade vão de 0.0 a 1.0. Um fator com probabilidade: (i) maior que 0.5 favorece a aplicação da concordância; (ii) menor que 0.5 inibe a sua realização; (iii) igual a 0.5 não interfere na realização da regra;
- (5) o cruzamento dos fatores (programa Crosstab).

Nos itens subseqüentes – respectivamente, 4.2.1 e 4.2.2 – trataremos dos fatores lingüísticos utilizados na análise da concordância de 3ª pessoa do plural e de 1ª pessoa do plural. A última parte deste capítulo (item 4.2.3) versa sobre os fatores de natureza social.

⁴¹ Nas tabelas relativas aos resultados da análise, usamos os valores de frequência de *aplicação* da concordância. Quando comparamos nossos resultados com os encontrados por Rodrigues (1987), porém, fazemos uso dos valores de *não-aplicação*, já que esta autora adotou este critério em seu estudo.

⁴² Para analisar os índices de frequência de cada categoria que compõe os fatores, tomamos como parâmetro a taxa geral de aplicação da regra de concordância encontrada no *corpus* (24% na 3ª pessoa do plural, e 33% na 1ª pessoa do plural).

⁴³ Este programa trabalha com níveis diversos de análise, realizando comparações progressivas entre os fatores. Seguindo a instrução de Scherre (1993), utilizamos os valores de probabilidade retirados do nível de análise que contém conjuntamente todos os fatores selecionados, tanto os lingüísticos, como os sociais. Justamente pelo fato de o programa Ivarb considerar a relação entre os diversos fatores, nem sempre há compatibilidade entre os números de probabilidade e os índices de frequência de concordância. Manteremos a nomenclatura de Labov, utilizando “probabilidade”, e não “peso relativo”.

4.2.1 CONCORDÂNCIA DE 3ª PESSOA DO PLURAL: FATORES LINGÜÍSTICOS

Fatores lingüísticos	Categorias
1) Posição e distância do sujeito em relação ao verbo	A – anteposto contíguo (SV) D – anteposto distante (S ... V) P – posposto contíguo (VS) Z – posposto distante (V... S)
2) Grau de saliência fônica da oposição entre as formas verbais do singular e do plural	R - (<i>fala/falam, come/comem</i>) V - (<i>faz/fazem, quer/querem</i>) L - (<i>dá/dão, está/estão, vai/vão</i>) E - (<i>falou/falaram, vendeu/venderam, partiu/partiram</i>) F - (<i>fez/fizeram, trouxe/trouxeram</i>) W - (<i>é/são</i>)
3) Presença/ausência do sujeito pronominal (<i>eles, elas, vocês</i>)	n - não-pronominal p - pronominal explícito z - pronominal não-explícito
4) Distância do referente do sujeito zero	a - referente anterior próximo na fala do informante b - referente anterior distante na fala do informante c - referente na fala do interlocutor e - referente mencionado posteriormente à ocorrência f - referente no contexto situacional
5) Função discursiva do sujeito	T - Tema R - Rema
6) Status informacional do sujeito	D - Dado N - Novo
7) Traço semântico do sujeito	A - [+humano] O - [-humano]
8) Função sintático-semântica do sujeito	D - determinado I - indeterminado
9) Paralelismo discursivo (entre verbos)	p - verbo precedido de verbo com marca de plural z - verbo precedido de verbo sem marca formal de plural u - verbo isolado ou primeiro de uma série
10) Paralelismo oracional (entre SN sujeito e verbo)	1 - presença da marca formal de plural -s no último elemento do SN sujeito 2 - zero no último elemento do SN sujeito
11) Definitude do sujeito	d - definido (identificável, específico, ex.: <i>as criança</i>) i - indefinido (não-identificável, ex.: <i>uns fornecedor</i>)
12) Valor semântico do verbo	A - Ação P - Processo 2 - Ação-Processo E - Estado C - Copulativo
13) Papel semântico do sujeito	1 - Agente 2 - Beneficiário 3 - Causativo 4 - Experimentador 5 - Inativo 6 - Paciente
14) Traço semântico do sujeito	1 - [+animado] 2 - [-animado]

Quadro 4: fatores lingüísticos utilizados na análise da concordância de 3ª pessoa do plural (P3)

O quadro da página anterior contém uma enumeração concisa de todos os fatores lingüísticos utilizados na análise da concordância de 3ª pessoa do plural.

No quadro a seguir, mostramos os fatores lingüísticos selecionados e descartados pelos programas do Varbrul conforme a relevância estatística (os fatores selecionados são apresentados por ordem de importância)⁴⁴. Nos itens subseqüentes, trataremos apenas daqueles que se mostraram significativos, com algumas observações referentes aos fatores não-selecionados, apenas quando estas forem relevantes.

Fatores selecionados	Fatores descartados
2) Grau de saliência fônica da oposição entre as formas verbais do singular e do plural 3) Presença/ausência do sujeito pronominal (<i>eles, elas, vocês</i>) 9) Paralelismo discursivo 7) Traço semântico do sujeito [+/-humano] 10) Paralelismo oracional 13) Papel semântico do sujeito	1) Posição e distância do sujeito em relação ao verbo 4) Distância do referente do sujeito zero 5) Função discursiva do sujeito 6) Status informacional do sujeito 8) Função sintático-semântica do sujeito 11) Definitude do sujeito 12) Valor semântico do verbo

Quadro 5: fatores lingüísticos (P3) selecionados e descartados pelos programas estatísticos do Varbrul

⁴⁴ Apenas o fator (14) traço semântico do sujeito [+/-animado] não pôde ser submetido a este programa de seleção de fatores, uma vez que houve não-concordância categórica com sujeito [-animado] (ver adiante comentários no item 4.2.1.4).

4.2.1.1 GRAU DE SALIÊNCIA FÔNICA DA OPOSIÇÃO ENTRE AS FORMAS VERBAIS DO SINGULAR E DO PLURAL

A escala de saliência fônica utilizada na análise, proposta por Lemle & Naro (1977) e adotada também por Rodrigues (1987), compreende dois níveis (conforme a intensidade dos segmentos fonéticos que realizam a oposição), e seis classes (conforme a crescente diferença material entre as formas verbais do singular e do plural):

1º NÍVEL (menos saliente): Pares cujos segmentos fonéticos que realizam a oposição são inacentuados em ambos os membros.

CLASSE R: verbos regulares. A diferença entre singular e plural reside na nasalidade (só nasalização ou nasalização e mudança de qualidade vocálica). Exs.: *fala/falam; come/comem; sai/saem*⁴⁵.

CLASSE V: a diferença entre o singular e o plural reside numa vogal final átona, possivelmente nasalada (nasalização e adição de segmento). Exs.: *faz/fazem; quer/querem; quis/quiseram; faltar/faltarem*.

2º NÍVEL (mais saliente): Pares cujos segmentos fonéticos com valor mórfico são acentuados em pelo menos um membro da oposição.

CLASSE L: elemento vocálico tônico oral no singular, em contraste com ditongo tônico nasal no plural (nasalização e mudança de qualidade). Exs.: *está/estão; vai/vão; dá/dão*.

CLASSE E: Pretéritos Perfeitos regulares, independentemente da conjugação; o acento recai na vogal temática. Exs.: *falou/falaram; vendeu/venderam; partiu/partiram*⁴⁶.

CLASSE F: Pretéritos Perfeitos irregulares, com variação no grau de abertura da vogal tônica, em ambas as formas do singular e do plural. Exs.: *trouxe/trouxeram; fez/fizeram; teve/tiveram; veio/vieram; deu/deram*.

CLASSE W: forma completamente distinta para o singular e o plural. Ex.: *é/são*.

⁴⁵ Formas do Pretérito Imperfeito do Indicativo – *falava/falavam, vendia/vendiam, partia/partiam, era/eram* – e do Subjuntivo – *tivesse/tivessem* – foram classificadas como R.

⁴⁶ Classificamos as formas *foi / foram*, como E.

Baseando-nos no princípio da saliência fônica (Lemle & Naro, 1977), estabelecemos a hipótese de que quanto maior ou mais saliente a diferença material entre as formas verbais do singular e do plural, maior a probabilidade de aplicação da concordância, por outro lado, quanto menor ou menos saliente essa diferença, menor a chance de realização da regra. A tabela abaixo contém os resultados encontrados:

Fatores	Frequência	Probabilidade
R – <i>come/comem</i>	32/290 = 11%	.24
V – <i>faz/fazem</i>	2/15 = 13%	.30
L – <i>dá/dão</i>	14/43 = 33%	.79
E – <i>falou/falaram</i>	58/116 = 50%	.89
F – <i>veio/vieram</i>	18/35 = 51%	.87
W – <i>é/são</i>	1/21 = 5%	.16
Total	125/520 = 24%	

Tabela 1: frequência e probabilidade de aplicação da regra de concordância verbal (P3) conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural

Os índices de frequência e probabilidade expostos na tabela em parte confirmam a hipótese formulada: os verbos das classes R e V do 1º nível, menos salientes de um ponto de vista fonológico, inibem a aplicação da concordância⁴⁷. Já os verbos das classes mais salientes L, E, e F do 2º nível, cuja diferença entre singular e plural é mais perceptível, favorecem o uso da regra. A diferença expressiva nos valores de probabilidade de aplicação que se estabelece entre um e outro nível não deixa dúvidas de que a intensidade dos segmentos fonéticos que realizam a oposição exerce um papel fundamental no condicionamento da concordância. Quanto à crescente diferença material entre as formas verbais do singular e do plural, vemos que há uma escala ascendente nos valores de frequência de concordância que se estende até as classes E e F. Nestas duas categorias, que apresentam um padrão semelhante, temos os números mais altos de probabilidade de realização da regra de CV.

Observemos agora os valores da classe W. Esta constitui a categoria mais saliente de um ponto de vista fonológico, já que comporta o grau máximo de diferenciação entre as formas verbais do singular e do plural. Uma vez que a saliência fônica se mostra relevante nas demais classes, esperaríamos que atuasse mais decisivamente neste contexto, fato que não se confirmou⁴⁸. Com índices de frequência e probabilidade de aplicação inferiores aos da categoria R, a classe W foge

⁴⁷ Relembramos que da classe V foram excluídas 27 ocorrências de verbo no infinitivo, que evidenciaram não-concordância categórica. Se estas tivessem sido inclusas, o índice de concordância da classe V cairia para 5%.

⁴⁸ Ainda que considerássemos todas as ocorrências de *é/são* encontradas no *corpus*, incluindo os contextos descritos em Critérios de Exclusão, a frequência de emprego do verbo no plural continuaria muito baixa: 8/76=10%. É interessante notar também que a forma *são* figura somente na fala de 3 informantes: 2 escolarizados e 1 não-escolarizado.

totalmente ao padrão verificado pelos estudos sobre a concordância verbal no português do Brasil, os quais geralmente constataam alta freqüência de realização da regra, quando a oposição entre singular e plural corresponde às formas *é x são*. A título de comparação, vejamos os resultados encontrados por Rodrigues (1987), entre falantes de 3ª faixa etária:

Fatores	Freqüência de não-aplicação	
	Nossos resultados	Rodrigues (1987)
R – <i>come/comem</i>	258/290 = 89%	237/258 = 92%
V – <i>faz/fazem</i>	13/15 = 87%	19/32 = 59%
L – <i>dá/dão</i>	29/43 = 67%	30/76 = 39%
E – <i>falou/falaram</i>	58/116 = 50%	36/101 = 36%
F – <i>veio/vieram</i>	17/35 = 49%	8/23 = 35%
W – <i>é/são</i>	20/21 = 95%	8/29 = 28%

Tabela 2: freqüência de não-aplicação da regra de concordância verbal (P3) conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural, no *corpus* sob análise e entre os falantes de 3ª faixa etária cujos dados compõem parte do *corpus* do português popular de São Paulo (Rodrigues, 1987:231)

Vemos que, com exceção da classe R, que apresenta um índice muito baixo de concordância nos dois *corpora*, as freqüências de não-aplicação da regra verificadas no nosso *corpus* são bem mais altas do que as constatadas por Rodrigues (1987). A diferença mais expressiva se encontra justamente na classe W: 28% (o menor índice de não-aplicação) em Rodrigues (1987), e 95% (o maior índice de não-aplicação) no *corpus* sob análise.

Acreditamos que tais divergências possam ser explicadas em termos sociais, tomando por base a região onde vivem os informantes. Os falantes do português popular de São Paulo cujos dados são estudados por Rodrigues (1987) habitam um espaço urbano, onde o modelo de realização lingüística se encontra mais disponível e formas de não-concordância são estigmatizadas, por isto estes tenderiam a aplicar com maior freqüência a regra padrão de concordância, quando a diferença entre singular e plural é [+saliente]. Os informantes que forneceram os dados da presente pesquisa residem na zona rural, onde possivelmente não existe tal estigmatização. Assim, os valores de não-aplicação da regra tendem a ser bem maiores em sua fala. A diferença nos padrões de concordância encontrados na classe W entre os dois grupos de falantes refletiria justamente uma distinção no valor social atribuído à variante não-aplicação da regra. Utilizando a nomenclatura de Labov (1991), poderíamos dizer que formas como *eles é* corresponderiam a um “*estereótipo*” no meio urbano, mas apenas a um “*indicador*” no meio rural.

Cabe ainda mencionar que embora a saliência fônica favoreça a aplicação da concordância nos contextos L, E e F do 2º nível, tal condicionamento não tem efeito nas instâncias de sujeito posposto ao verbo: a não-concordância é praticamente categórica (95%), não importando a classe

morfológica do verbo (cf. tabela 3). Rodrigues (1987:171) chega às mesmas conclusões: nos casos de sujeito posposto⁴⁹, contexto em que a não-realização da concordância é menos óbvia, o peso do fator saliência fônica é anulado.

Fatores	Frequência
R – <i>come/comem</i>	0/19 = 0%
V – <i>faz/fazem</i>	0/2 = 0%
L – <i>dá/dão</i>	0/5 = 0%
E – <i>falou/falaram</i>	2/21 = 10%
F – <i>veio/vieram</i>	1/6 = 17%
W – <i>é/são</i>	0/2 = 0%
Total	3/55 = 5%

Tabela 3: frequência de aplicação da regra de concordância verbal (P3) com sujeitos pospostos ao verbo conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural

Conforme nota Pezatti (1992), o sujeito posposto não apresenta os traços prototípicos de sujeito, sejam eles de natureza semântica (ser ativo), sintática (estar na posição pré-verbal; ser controlador da CV), ou discursiva (ser Tópico). De fato, notamos que a grande maioria das ocorrências de sujeitos pospostos ao verbo encontradas no *corpus* apresenta o papel semântico de Paciente (45%), como nos exemplos (94-97), ou Inativo (35%), como em (98-99), sendo poucos os casos de sujeito Agente (20%). Todos os sujeitos pospostos exercem a função discursiva de Rema e em 87% dos casos veiculam informação nova. Parece que o fato de não corresponderem ao Tema, que é o ponto de partida da oração e função discursiva natural (não-marcada) do sujeito, junta-se ao fato de não apresentarem o traço semântico atividade, fazendo com que o falante não perceba o SN posposto como sujeito. Com relação aos verbos empregados nestas construções, notamos que 69% correspondem a verbos de um só argumento, como em (94-97), em sua grande maioria de natureza existencial, como em (95-97). Segundo Pezatti (1992) e Camacho (1993), tais verbos tendem a condicionar a ordem VS e a não-aplicação da concordância. Exemplos:

- (94) miorô as istrada num tinha istrada né? (I.5,p.59)
- (95) aí completô seti anu (I.9,p.133)
- (96) cabô us matu qui tinha nu pur nu brasil né? (I.5,p.60)
- (97) otru dia morreu uns par de véio aqui (I.8,p.109)
- (98) essis vinti dia têm validu um TANtu (qui nu tem preço) quanto valeu essis vinti dia (I.13,p.194)
- (99) agora ficô lá minha cunhada lá e a subrinhada (I.10,p.138)

⁴⁹ No *corpus* estudado por Rodrigues (1987:162), a não-concordância com sujeitos pospostos ao verbo também é praticamente categórica: 97% (72/74).

4.2.1.2 PRESENÇA/AUSÊNCIA DO SUJEITO PRONOMINAL

O fator presença/ausência do sujeito pronominal (*eles, elas, vocês*) foi baseado no estudo de Rodrigues (1987) e abrange três categorias:

(i) sujeito não-pronominal – representado por um nome lexical, como em (100); sujeito composto, como em (101); ou outros pronomes⁵⁰, como em (102):

(100) *os fioo tamém cooperaro alguma coisa dipois (I.1,p.11)*

(101) *ela i a mulher du véiu trabaiava um bucadinhu (I.15,p.239)*

(102) *otrus tinha aquela caxoti di madera (I.13,p.200)*

(ii) sujeito pronominal explícito – representado pelos pronomes pessoais *eles, elas, vocês*, exs.:

(103) *eles mi trazem como se eu fosse uma relíquia pra eles nossa senhora (I.1,p.8)*

(104) *ô gente *cêis* desculpa que a gente num sabe conversá muito bem (I.1,p.9)*

(iii) sujeito pronominal não-explícito – sujeito zero, exs.:

(105) *os home da antiguidade durava munto pur causa disso...tinha meno do vício (I.5,p.65)*

(106) *eles mexe com: :: com com q/ que essa: :: com esse: :: éh o giro leiteiro...mexe com esse gado até hoje (I.6,p.86)*

Estabelecemos a hipótese funcional de que o sujeito zero condicionaria o uso de formas verbais marcadas, uma vez que a flexão torna claro o referente da casa vazia do sujeito. Formas verbais não-marcadas, por outro lado, tenderiam a ser utilizadas com sujeito explícito, expresso por nome lexical ou por pronome, de modo a não comprometer a informação. Os resultados são expostos na tabela a seguir:

⁵⁰ No caso de sujeito expresso pelo pronome relativo *que* consideramos o seu referente, classificando-o como sujeito pronominal, ou não-pronominal, conforme retomasse um pronome, ou um nome lexical, exs.: *aqueli conventu ali foi elis qui fizeru (I.10, p.151); é mãe *dessas menina* que tão aí (I.1, p.7).*

Fatores	Freqüência	Probabilidade
n – não-pronominal	17/205 = 8%	.21
p – pronominal explícito	71/175 = 41%	.68
z – pronominal não-explícito	37/140 = 26%	.73
Total	125/520 = 24%	

Tabela 4: freqüência e probabilidade de aplicação da regra de concordância verbal (P3) de acordo com a presença/ausência do sujeito pronominal⁵¹

Observando os valores de probabilidade, vemos que apenas em parte confirmamos a hipótese: o sujeito zero tende realmente a favorecer a realização da concordância, e o sujeito não-pronominal a desfavorece decisivamente. Contudo, o sujeito pronominal explícito – contexto em que a flexão é redundante, uma vez que a presença do pronome identifica o referente do verbo – contribui expressivamente para a realização da regra de CV.

Se examinarmos as freqüências da tabela 4, há dois índices que se destacam: de um lado, os casos de sujeito não-pronominal, com 8% de aplicação da concordância e, de outro, as instâncias de sujeito pronominal explícito, com 41%. Os casos de sujeito zero, por sua vez, se situam em uma faixa intermediária, com uma freqüência não muito significativa de realização da regra (26%).

Ainda que o programa estatístico não tenha selecionado o fator distância do referente do sujeito zero – o que provavelmente se deve ao número reduzido de ocorrências encontradas para alguns contextos que compõem este fator –, notamos que o índice de concordância com sujeito pronominal não-explícito tende a diferir conforme a localização do referente da casa vazia do sujeito. Há maior freqüência de aplicação da regra nas categorias: (b) referente anterior distante na fala do informante (18/46=39%), e (e) referente mencionado posteriormente à ocorrência (5/12=42%), do que nos contextos: (a) referente anterior próximo na fala do informante (11/59=19%), (c) referente anterior na fala do interlocutor (2/17=12%), e (f) referente presente no contexto situacional (1/6=17%)⁵².

Considerando os casos de sujeito zero conforme a função sintático-semântica do sujeito, verificamos que a freqüência de concordância nas instâncias de referente determinado (28/122=23%) é bem menor do que nas de referente indeterminado (9/18=50%). O número

⁵¹ Apesar de a freqüência de concordância ser maior na categoria *p* (sujeito pronominal explícito) do que em *z* (sujeito pronominal não-explícito), o programa estatístico atribuiu maior probabilidade de concordância a (*z*) do que a (*p*). Isto ocorre, conforme já dissemos, porque este leva em conta a relação entre os diversos fatores. Através de dois cruzamentos, verificamos que em alguns contextos favorecedores da concordância, como a classe F do fator saliência fônica (cf. tabela 6) e a categoria (*p*) do fator paralelismo discursivo (cf. tabela 8), a freqüência de realização da regra com sujeito pronominal não-explícito é maior do que a encontrada nos casos de sujeito pronominal explícito, coincidindo com a ordem de classificação dada pelos valores de probabilidade.

⁵² Classificamos o referente do verbo como distante, quando mais de dez sílabas o separam da ocorrência de casa vazia do sujeito. Note-se que o contexto (e), no qual se encontra o maior índice de aplicação da concordância, engloba os únicos casos em que a casa vazia do sujeito constitui informação nova, no sentido de não conhecida, na medida em que seu referente só é esclarecido no discurso subsequente, ex.: *ia só as mulher qui ia tê serviçu* (I.10.p.141).

reduzido de ocorrências deste último contexto, porém, não permite uma conclusão definitiva a este respeito.

A diferença expressiva entre a frequência de realização da regra com sujeito não-pronominal e pronominal explícito, por seu lado, levou a três questões: (i) o alto índice de concordância com sujeito pronominal explícito estaria sendo influenciado pela função sintático-semântica do sujeito, isto é, nos casos de *eles* de referência indeterminada, contexto em que a marca de plural seria também condicionada pelo critério semântico, haveria maior índice de aplicação da concordância? (ii) haveria diferença na taxa de concordância de sujeito pronominal conforme a pessoa do discurso: 2ª (*vocês*) x 3ª (*eles*)? (iii) qual a posição que tais SNs ocupam na oração?

A resposta para a primeira questão é negativa. Através do cruzamento dos fatores presença/ausência do sujeito pronominal e função sintático-semântica do sujeito, verificamos que a concordância com sujeito pronominal explícito tende a ser realizada tanto nos casos de referente determinado, como nos de referente indeterminado. Em tais contextos, considerados separadamente, constatamos os índices de respectivamente 42% (46/110) e 38% (25/65) de aplicação da regra.

Analisando os casos de sujeito pronominal explícito conforme a pessoa do discurso, verificamos igualmente que este não constitui um fator relevante: o índice de realização da concordância se mantém praticamente igual, seja o sujeito representado por *vocês* (15/36=42%), ou por *eles/elas* (56/139=40%).

Na tabela a seguir, expomos os resultados relativos à terceira questão:

Fatores	Frequência	
	n - Sujeito não-pronominal	P - Sujeito pronominal explícito
A – Anteposto contíguo	8/82 = 10%	68/165 = 41%
D – Anteposto distante	6/69 = 9%	3/9 = 33%
P – Posposto contíguo	2/43 = 5%	0/1 = 0%
Z – Posposto distante	1/11 = 9%	-

Tabela 5: frequência de aplicação da regra de concordância verbal (P3) com sujeito não-pronominal e pronominal explícito conforme a posição e distância do sujeito em relação ao verbo

Observemos primeiramente os casos de sujeito pronominal explícito: das 175 ocorrências encontradas no *corpus*, 1 corresponde a sujeito posposto ao verbo, 9 a sujeito anteposto distante do verbo, e a maior parte (165) a sujeito anteposto contíguo ao verbo⁵³, posição que tornaria a relação sujeito-verbo mais saliente para o falante, favorecendo o uso da regra de concordância. Não temos, contudo, como comprovar esta hipótese, na medida em que as ocorrências de sujeito pronominal se concentram em um só contexto.

⁵³ Adotando os critérios de Rodrigues (1987), consideramos o sujeito imediatamente anteposto ao verbo, quando apenas advérbios, como *já, não*, ou pronomes clíticos, como *me, se*, os separam.

Passemos aos casos de sujeito não-pronominal. Ainda que estes englobem quase todas as ocorrências de sujeito posposto ao verbo – contexto que, como vimos, desfavorece decisivamente a concordância –, a tabela 5 evidencia que não foi o fator posição que determinou a baixa frequência de realização da regra com sujeitos não-pronominais. Isto porque os valores de aplicação da concordância com sujeito não-pronominal se mantêm semelhantes nas quatro categorias estabelecidas para o fator posição, não ultrapassando o índice de 10%. Estes dados indicam que o falante popular tende realmente a não-aplicar a regra de CV com sujeitos não-pronominais, independentemente de sua posição em relação ao verbo.

No *corpus* do português popular de São Paulo, Rodrigues (1987:170) igualmente constata valores altos de frequência e probabilidade de não-concordância com sujeito não-pronominal, respectivamente, 82% e .67. Quanto às ocorrências de sujeito pronominal explícito e não-explícito, a autora verifica que enquanto no primeiro caso a frequência de não-realização da regra de concordância é de 72%, no segundo é de apenas 48%. Os valores de probabilidade de não-aplicação nestes dois contextos são de, respectivamente, .45 e .38 – o que leva a autora a concluir que estes interferem de maneira discreta na realização da regra: o sujeito zero tende a condicionar a concordância, ao passo que o sujeito pronominal explícito favorece o uso de formas verbais não-flexionadas. Vemos, portanto, que o padrão encontrado por Rodrigues (1987) é semelhante ao da presente pesquisa, diferindo apenas nos casos de sujeito pronominal explícito, nos quais constatamos tendências opostas.

Rodrigues (1987:172) examina também o padrão de concordância conforme a presença/ausência do sujeito pronominal e o grau de saliência fônica da oposição entre as formas verbais do singular e do plural. A autora verifica que a saliência fônica tem maior peso em instâncias de sujeito pronominal explícito e não-explícito do que nos casos de sujeito não-pronominal. Interessou-nos então analisar este mesmo aspecto no *corpus* sob análise. Os resultados são apresentados a seguir:

Fatores	Frequência		
	n - Sujeito não-pronominal	p - Sujeito pronominal explícito	z - Sujeito pronominal não-explícito
R – <i>come/comem</i>	4/116 = 3%	24/109 = 22%	4/65 = 6%
V – <i>faz/fazem</i>	0/2 = 0%	2/6 = 33%	0/7 = 0%
L – <i>dá/dão</i>	3/21 = 14%	9/11 = 82%	2/11 = 18%
E – <i>falou/falaram</i>	7/44 = 16%	29/37 = 78%	22/35 = 63%
F – <i>veio/vieram</i>	3/12 = 25%	6/10 = 60%	9/13 = 69%
W – <i>é/são</i>	0/10 = 0%	1/2 = 50%	0/9 = 0%

Tabela 6: frequência de aplicação da regra de concordância verbal (P3) segundo a presença/ausência do sujeito pronominal e o grau de saliência fônica da oposição entre as formas verbais do singular e do plural

A tabela 6 evidencia que os valores de freqüência de aplicação encontrados nas classes L, E e F do 2º nível são sempre mais altos do que os das classes R e V do 1º nível, independentemente do tipo de representação do sujeito. Isto mostra que a saliência fônica é realmente relevante para a concordância, ainda que não tenha peso igual em todos os contextos.

Se observarmos somente os índices encontrados nas classes E e F – as quais, como vimos, condicionam decisivamente a concordância – segundo a presença/ausência do sujeito pronominal, vemos que a saliência fônica atua muito mais fortemente nas instâncias de sujeito pronominal explícito e não-explícito do que nos casos de sujeito não-pronominal. A diferença de freqüência de aplicação da regra encontrada nestes contextos confirma, portanto, a tendência de sujeitos pronominais explícitos e não-explícitos favorecerem a CV e sujeitos não-pronominais a desfavorecerem. Voltaremos à questão da concordância com sujeitos pronominais e não-pronominais no item 4.2.1.5, que versa sobre o paralelismo oracional.

4.2.1.3 PARALELISMO DISCURSIVO

Para estabelecer este fator, tomamos por base o trabalho de Scherre & Naro (1993), com algumas modificações nos critérios de análise. O fator compõe-se de três categorias:

- (i) verbo precedido de verbo com marca de plural;
- (ii) verbo precedido de verbo sem marca formal de plural;
- (iii) verbo isolado ou primeiro de uma série.

Consideramos o verbo com a mesma referência que o verbo da construção anterior (mesmo referente do ponto de vista semântico) e não separado da construção anterior por mais de 10 orações, conforme ilustrado nas seqüências abaixo:

- (107-111) *eles chega fala que vai lê pra mim eu num aceito...“não *cêis* pula ...eu que gosto de lê... não *cêis* me engana” (I.1,p.10)*
- (112-113) *lá na roça eu ensinei *as criança* que precisava né?...e *as coitadinha* (num podia) í pra escola e queria aprendê né? (I.1,p.15)*
- (114-115) *morreu *do/treis da família...morreu esse gonçalo a nora...*num sei se *sobrinhu* (I.8,p.109)*
- (116-121) *dipoi eu sortei *us trei vira-lata* qui: :: qui num intendia da arti pa: ::vê si cunsguia né? *elis andaru andaru andaru* pu meu du pastu (mai) num cunsguiru descubrí (I.9, p.123)*

Observamos apenas o discurso do informante. Sabemos que o efeito do paralelismo discursivo pode se manifestar mesmo com mudança de turno entre os interlocutores, mas uma vez que já tínhamos descartado as ocorrências em que o informante repete a forma verbal da pergunta feita pelo documentador, resolvemos considerar o paralelismo somente na fala do informante.

Scherre & Naro (1993) estabelecem como critério que não haja mudança de turno entre as ocorrências. Nesta pesquisa, no entanto, consideramos as ocorrências que estiverem próximas, ainda que separadas apenas por uma breve interrupção do documentador, como em (122-123):

- (122-123) Doc. di dia eles ficam soltos?
 Inf. di dia fica sorto
 Doc.1 à noite (prende)
 Inf. di noite...não as vez dorme no pasto (I.2,p.22)

Algumas vezes o verbo da construção que antecede a ocorrência não foi considerado na contagem (ocorrência descartada), mas para a análise do fator paralelismo discursivo este foi levado em conta apenas como antecedente. Os exemplos a seguir ilustram alguns desses casos: em (124), o verbo descartado (em itálico) é *ter* com valor existencial; em (125), *ser* denotando quantidade; em (126), ocorre truncamento; em (127), o verbo anterior é repetição da forma utilizada pelo interlocutor⁵⁴:

- (124) *tinha* uns *tinha* uns fornecedor que comprava...matos por aí ... (I.4, p.82)
 (125) *era* seis sete carro...passava aqui essas hora (I.7,p.98)
 (126) até que *chegô* os nov/ tá chegando os noventa e um graças a deus (I.1,p.4)
 (127) Doc.1 as crianças vão pras ruas elas voltam? / Inf. *vão*...vai (I.15,p.230)

Seguindo o princípio do paralelismo formal (Scherre & Naro, 1993), formulamos a hipótese de que o verbo precedido de verbo sem marca formal de plural desfavoreceria a realização da concordância, e o verbo precedido de verbo com marca de plural a favoreceria. Vejamos os resultados encontrados:

⁵⁴ Apenas o verbo *ser* na expressão cristalizada *é que* não foi levado em conta como antecedente no fator paralelismo discursivo.

Fatores	Frequência	Probabilidade
p – verbo precedido de verbo com marca de plural	32/57 = 56%	.68
z – verbo precedido de verbo sem marca formal de plural	19/162 = 12%	.29
u – verbo isolado ou primeiro de uma série	74/301 = 25%	.58
Total	125/520 = 24%	

Tabela 7: frequência e probabilidade de aplicação da regra de concordância verbal (P3) segundo o paralelismo discursivo

Os valores de frequência e probabilidade da tabela 7 confirmam a hipótese estabelecida. Existe realmente uma tendência de o falante repetir as últimas formas pronunciadas, de modo que formas verbais não-marcadas tendem a ser seguidas por formas não-marcadas; e formas verbais flexionadas, por formas flexionadas.

Através do cruzamento dos fatores paralelismo discursivo e presença/ausência do sujeito pronominal (cf. tabela 8), verificamos que, no contexto verbo precedido de verbo com marca de plural, a frequência de concordância é bem maior nas instâncias de sujeito pronominal explícito e não-explícito do que nos casos de sujeito não-pronominal. Nas ocorrências de sujeito não-pronominal precedido de uma forma verbal no singular, a não-concordância se mostra categórica.

Fatores	Frequência		
	Sujeito Não-pronominal	Pronominal explícito	Pronominal não-explícito
p – verbo precedido de verbo com marca de plural	2/12 = 17%	12/19 = 63%	18/26 = 69%
z – verbo precedido de verbo sem marca formal de plural	0/43 = 0%	12/40 = 30%	7/79 = 9%
u – verbo isolado ou primeiro de uma série	15/150 = 10%	47/116 = 41%	12/35 = 34%

Tabela 8: frequência de aplicação da regra de concordância verbal (P3) conforme o paralelismo discursivo e a presença/ausência do sujeito pronominal

Labov (1994) interpreta o paralelismo discursivo como uma força de natureza contrafuncional. Segundo o autor, as marcas são utilizadas, quando redundantes; e omitidas, quando necessárias. Pode-se de fato argumentar que uma sucessão de formas verbais flexionadas, como ocorre nos exemplos (118-121) reproduzidos abaixo, veicula informação redundante:

- (118-121) dipoi eu sortei us trei vira-lata qui: :: qui num intendia da arti pa: ::vê si cunsiagua né? elis andaru andaru andaru pu meu du pastu (mai) num cunsiaguru descubrí (I.9, p.123)

Contudo, a tabela 7 evidencia que o número de instâncias de verbo precedido de verbo marcado é reduzido. A maioria das ocorrências se concentra nas duas outras categorias: verbo isolado ou primeiro de uma série – contexto irrelevante para o fator paralelismo discursivo – e verbo precedido de verbo sem marca formal de plural, na qual há apenas 12% de frequência de concordância. Resolvemos então examinar mais de perto este último contexto, com o intuito de verificar até que ponto uma seqüência de verbos não-flexionados comprometeria a informação. Observemos os exemplos (128-138) a seguir:

- (128-131) aí baxava *us mininu* tava chuvendu tá trabaiandu tá sol tá trabaiandu i i num parava tempu ninhum (I.13,p.196)
- (132-134) *as muié* saía naquelas fazenda cum balaio na cabeça lavava ropa pra um lá uma hora duas hora ganhava dissu tudu (I.13,p.198)
- (135-138) *us cumpanheru deli* lá *aquela turma* leva um engradadu di cerveja (...) eu falei “pinga cêis podi bebê pra lá suas pinga di vocêis lá” porque eu (eu nu costumu dexá) nada () eu falei “ó cêis guarda suas pinga di vocêis lá () vorta pra lá u qui cêis quisé bebê cêis bebi” (I.14,p.215)

Vemos que na primeira oração, o falante explicita o sujeito, o qual constitui informação nova. Já nas orações subseqüentes, o sujeito passa a ser informação conhecida pelo interlocutor. Ainda que haja uma sucessão de verbos não-flexionados, a informação relativa ao referente do verbo não é comprometida, uma vez que ou o sujeito é explícito, ou a casa vazia do sujeito retoma referente já mencionado no discurso. Cruzando os fatores paralelismo discursivo e status informacional do sujeito, verificamos que no contexto verbo precedido de verbo sem marca de plural, o sujeito constitui informação dada em 82% das ocorrências. Nos 18% restantes de casos de verbo precedido de verbo não-flexionado, em que o referente constitui informação nova, por sua vez, o sujeito é explícito⁵⁵. Tais instâncias se concentram em três contextos, exemplificados a seguir:

- (i) desdobramento do sujeito, exs.:

- (114-115) morreu do/treis da família...morreu *esse gonçalo a nora*...num sei se *sobrinhu* (I.8,p.109)

⁵⁵ Há apenas uma ocorrência de sujeito zero: *dipois du padri zé duqui já passô multus...veiu passandu...agora tá passandu mais dipressa agora () já passô...cum essi daí já é treis* (I.13,p.204). Neste exemplo, a casa vazia do sujeito tem seu referente esclarecido posteriormente à ocorrência. Do ponto de vista do status informacional, este constitui informação nova. Do ponto de vista do paralelismo discursivo, porém, consideramos que a seqüência de verbos tem o mesmo referente: os padres que trabalharam na igreja.

- (ii) estruturas clivadas que “marcam” o SN sujeito como informação nova, ainda que já tenha sido mencionado no discurso, exs.:

(139-140) *é os pai qui mandava nu (caminhu) eu falei essis dia aqui é *us pai* qui mandava nas fiia i nus fiio (I.12,p.183)*

- (iii) sujeito zero com referente esclarecido na ocorrência subsequente, exs.:

(141-142) *muda muito modifica muito *as coisa* (I.15,p.242)*

Vemos, portanto, que em nenhum desses casos pode-se falar de um comprometimento da informação relativa ao referente do verbo, causado pela não-utilização das marcas de flexão número-pessoal.

4.2.1.4 TRAÇO SEMÂNTICO DO SUJEITO

Este fator foi sugerido pelo estudo de Naro & Scherre (1999). Conforme os autores, estabelecemos a hipótese de que o sujeito cujo traço é [+humano] favoreceria a aplicação da regra de CV, e o sujeito [-humano] a desfavoreceria. Os dados quantificados são exibidos a seguir:

Fatores	Frequência	Probabilidade
A – sujeito [+humano]	84/335 = 25%	.59
O – sujeito [-humano]	7/102 = 7%	.24
Total	91/437 = 21%	

Tabela 9: frequência e probabilidade de aplicação da regra de concordância verbal (P3) de acordo com o traço semântico do sujeito [+humano]⁵⁶

Os números de frequência e probabilidade expostos na tabela validam a hipótese formulada: o sujeito [+humano], como em (143), tende realmente a favorecer a realização da concordância, e o sujeito [-humano], como em (144-147), a inibe decisivamente.

(143) *cêis nunca foru aqui nu hortu? (I.9,p.117)*

(144-145) *os bicho as veiz não queria...num num guentava (I.7,p.99)*

(146) *us matu pudia intrá dendi casa qui eli nu capinava (I.14,p.218)*

(147) *tinha treis fábrica de queijo aqui...mais depois foi cabando (I.6,p.88)*

⁵⁶ Não aplicamos o fator traço semântico do sujeito em 83 ocorrências que correspondem a instâncias de sujeito indeterminado, uma vez que este sempre é [+humano].

Cruzando o fator traço semântico do sujeito [+/-humano] com o fator presença/ausência do sujeito pronominal (cf. tabela 10), verificamos ainda que há sempre maior índice de concordância com sujeito [+humano] do que com [-humano], seja ele não-pronominal, pronominal explícito, ou pronominal não-explícito. Nos casos de sujeito não-pronominal cujo traço semântico é [-humano], como em (144) e (146), contexto que acumula dois fatores inibidores da realização da regra, a não-concordância é categórica.

Fatores	Frequência	
	Sujeito [- humano]	Sujeito [+ humano]
n - não-pronominal	0/59 = 0%	17/146 = 12%
p - pronominal explícito	4/11 = 36%	42/99 = 42%
z - pronominal não-explícito	3/32 = 9%	25/90 = 28%

Tabela 10: frequência de aplicação da regra de concordância verbal (P3) conforme o traço semântico do sujeito [+/-humano] e a presença/ausência do sujeito pronominal

Em nossa análise, examinamos também o traço semântico do sujeito [+/-animado]. Constatamos que quando o sujeito é [-animado], como em (146-147), a não-aplicação da regra é categórica (66/66=100%). Com sujeito [+animado], como em (143-145), temos o mesmo valor de frequência de aplicação constatado com sujeito [+humano]: 25% (91/371). Estes resultados nos levam a crer que os traços [+humano] e [+animado] tornam o sujeito mais saliente para o falante, fazendo com que a regra de CV seja mais frequentemente aplicada. Voltaremos a tratar desta questão no item 4.2.1.6 relativo ao papel semântico do sujeito.

4.2.1.5 PARALELISMO ORACIONAL

O fator paralelismo oracional foi baseado no estudo de Scherre & Naro (1993). Este foi utilizado para verificar se existe uma correlação entre a presença/ausência de marca de plural no último elemento do SN sujeito e a realização/não-realização da concordância.

Na análise inicial dos dados, estabelecemos uma série de categorias de classificação baseadas nas estruturas que foram encontradas no *corpus*:

- 1) presença de -s no último elemento do SN sujeito simples – sujeitos expressos por nome lexical no plural, como em (148), ou pronomes, como em (149-151);

(148) agora *os filhos*: :: casa (I.2,p.23)

(149) *otrus* ficavam nu meu (I.12,p.180)

(150) *cês* vieru dali de taubaté memo ou de ota cidade? (I.4,p.55)

(151) *elis* tratam ela de preta eu não sei o nome dela (I.3,p.43)

2) zero no último elemento do SN sujeito simples – sujeitos expressos por determinante no plural e nome no singular, como em (152-157);

- (152) *meus fio trabaiô muito* (I.13,p.192)
 (153) *não...é só isso aí que os mais véio contava* (I.4,p.46)
 (154) *o meus irmão sabia* (I.12,p.184)
 (155) *tem treis puliça que toma conta lá* (I.5,p.69)
 (156) *os otro meu irmão que estudô mais* (I.8,p.106)
 (157) *muitas minina si perdia naquela época* (I.9,p.118)

3) zero no último elemento do SN sujeito composto – sujeito constituído de: dois ou mais nomes no singular, como em (158), dois nomes no singular com determinante no plural, como em (159), ou nome no singular e pronome, como em (101)⁵⁷;

- (158) *ocê nu conta mais que o que o dito o calecu o joaquim fuma*
 (I.12,p.173)
 (159) *esses caminhãu esses trator desgraçô tudo* (I.7,p.100)
 (101) *ela i a mulher du véiu trabaiava um bucadinho* (I.15,p.239)

4) numeral como último elemento do SN sujeito, como em (160-161);

- (160) *as treis funciona* (I.15,p.240)
 (161) *otus trei num era: :: né? muito chegadu im im busca nãu* (I.9,p.123)

5) zero no último elemento do modificador (sintagma preposicionado (SP) ou sintagma adjetival (SA)) do SN sujeito⁵⁸, como em (162-165);

- (162) *as criança di hoji nu toma nem um biliscãozinho assim* (I.10,p.140)
 (163) *us fiu homi deli apanhava com vinti i um anu si eli visse fumá*
 (I.12,p.172)
 (164) *ota duas mai grandi vai istudá na parti da tardi* (I.2,p.25)
 (165) *eu tenho us cumpanherada meu qui pesca aí até hoji* (I.9,p.126)

⁵⁷Os sujeitos compostos encontrados no *corpus* têm sempre zero no último elemento. Nos exemplos (158-159), os SNs sujeitos poderiam também ser interpretados como sujeitos de lista.

⁵⁸ Em todas as instâncias examinadas, o plural se restringe ao determinante do SN sujeito. Não foram encontradas ocorrências da estrutura presença de -s no último elemento do modificador (SP ou SA).

Com base no princípio do paralelismo formal (Scherre & Naro, 1993), estabelecemos a hipótese de que a presença de marca de plural no último elemento do SN sujeito tenderia a favorecer a realização da concordância, e a ausência de marca a desfavoreceria. Os resultados relativos à quantificação dos dados são mostrados na tabela abaixo:

Fatores	Frequência
1 – presença de -s no último elemento do SN sujeito simples	75/193= 39%
2 – zero no último elemento do SN sujeito simples	7/103 = 7%
3 – zero no último elemento do SN sujeito composto	1/10 = 10%
4 – numeral como último elemento do SN sujeito	0/4 = 0%
5 – zero no último elemento do modificador (SP ou SA) do SN sujeito	2/15 = 13%
Total	85/325 = 26%

Tabela 11: frequência de aplicação da regra de concordância verbal (P3) segundo o paralelismo oracional (1ª categorização)⁵⁹

A tabela 11 evidencia que há um número reduzido de ocorrências para as categorias 3, 4 e 5, e que a maioria das instâncias se concentra nos grupos 1 e 2. Os contextos 2, 3, 4 e 5 têm em comum o fato de não apresentar marca formal de plural no último elemento do SN sujeito, e de condicionar a não-concordância – em contraposição ao contexto 1, em que a marca está presente e a frequência de aplicação da regra se mostra elevada (39%) . Por isto, resolvemos agrupá-los em um só item (número 2), chegando às seguintes categorias: (1) presença da marca formal de plural -s no último elemento do SN sujeito, (2) zero no último elemento do SN sujeito. Os resultados referentes a esta segunda categorização são exibidos a seguir:

Fatores	Frequência	Probabilidade
1 – presença da marca formal de plural -s no último elemento do SN sujeito	75/193= 39%	.64
2 – zero no último elemento do SN sujeito	10/132 = 8%	.31
Total	85/325 = 26%	

Tabela 12: frequência e probabilidade de aplicação da regra de concordância verbal (P3) segundo o paralelismo oracional (2ª categorização)

Os valores expostos na tabela 12 sugerem que a hipótese foi confirmada: há realmente maior frequência e probabilidade de realização da regra de CV quando a marca formal de plural está presente no último elemento do SN sujeito do que quando está ausente.

⁵⁹ Na análise do fator paralelismo oracional, só são examinados os sujeitos expressos antepostos ao verbo. Foram, portanto, excluídas 140 ocorrências de sujeito zero e 55 de sujeito posposto.

A categoria 2 engloba, como vimos, quatro tipos de sujeito, todos representados por nome lexical, já a categoria 1 compreende tanto sujeitos expressos por nome lexical como sujeitos expressos por pronome. Vimos no item 4.2.1.2 que o sujeito pronominal explícito favorece a aplicação da regra de CV, enquanto o sujeito não-pronominal a inibe decisivamente. Resolvemos então examinar mais detalhadamente os resultados do fator paralelismo oracional, cruzando-o com presença/ausência do sujeito pronominal.

Fatores	Frequência	
	1 – Presença da marca formal de plural -s no último elemento do SN sujeito	2 – zero no último elemento do SN sujeito
p - Sujeito pronominal	71/174 = 41%	-
n - Sujeito não-pronominal	4/19 = 21%	10/132 = 8%

Tabela 13: frequência de aplicação da regra de concordância verbal (P3) conforme o paralelismo oracional e a presença/ausência do sujeito pronominal

Vemos, portanto, que nos casos de sujeito pronominal, só há ocorrências da categoria 1, ou seja, sujeitos expressos por *eles, elas, vocês*⁶⁰. Havia 4 ocorrências do tipo *Ele fizeram* no *corpus*, no entanto, estas foram descartadas como sujeito singular e verbo no plural. Ocorrências sem nenhuma marca formal de plural, como *Ele fez*, ainda que pudessem remeter a sujeitos plurais, também não foram incluídas na análise. Na medida em que não podemos contrapor presença x ausência de marca, não é possível concluir se a concordância é condicionada pela marca -s no final do pronome. Acreditamos que neste contexto a realização da regra seja favorecida pelo próprio sujeito pronominal – o qual, como vimos, aparece na maioria das vezes imediatamente anteposto ao verbo, posição que o tornaria mais evidente para o falante. Tal hipótese é reforçada pela diferença significativa de frequência de aplicação da regra nos dois contextos em que a marca de plural no último elemento do SN sujeito está presente: sujeito pronominal (41%) x sujeito não-pronominal (21%).

Tratemos agora dos casos de sujeito não-pronominal. Ainda que a frequência de concordância neste contexto seja em geral baixa, a tabela 13 sugere que o paralelismo oracional é atuante: há maior índice de aplicação quando o último elemento do SN sujeito comporta marca de plural (21%) do que quando tal marca não está presente (8%). O número reduzido de ocorrências de sujeito não-pronominal com marca de plural no último elemento do SN sujeito não nos permite, contudo, fazer uma afirmação categórica a este respeito.

⁶⁰ Conforme mencionado no item 4.2.1.2, pronomes não-anafóricos, como *muitos, outros*, foram classificados como sujeito não-pronominal.

A maior parte das instâncias de sujeito não-pronominal se concentra no contexto 2, no qual o último elemento do sujeito não carrega marca formal de plural. A categoria 2 engloba, como vimos, quatro estruturas diferentes, mas a maioria das ocorrências (103/132=78%) corresponde a casos de sujeitos expressos por determinante no plural e nome no singular, como nos exemplos (152-157). Neste contexto, há somente 7% de frequência de aplicação da regra de concordância (cf. tabela 11). Isto evidencia que o falante popular tende a marcar o plural apenas nos primeiros elementos do SN sujeito, enquanto os demais elementos que compõem esse sujeito, bem como o verbo, não carregam marcas de flexão. Acreditamos que tal fato possa ser interpretado não como de natureza contrafuncional, mas sim funcional: a informação relativa ao referente do verbo não é prejudicada, já que o plural foi explicitado em algum ponto do SN sujeito, e a ausência de flexão nos elementos que o seguem evita a redundância.

4.2.1.6 PAPEL SEMÂNTICO DO SUJEITO

Este fator foi sugerido a partir da leitura das obras de Halliday (1994) e Chafe (1979). Para a classificação do papel semântico do sujeito, seguimos o modelo proposto por Borba (1991). Embora Halliday e Chafe estabeleçam igualmente uma tipologia das funções semânticas exercidas pelos argumentos do verbo, a classificação de Borba se baseia na língua portuguesa, sendo por isto mais adequada para nossa descrição. Transcrevemos a seguir as definições fornecidas por Borba (1991:xx-xxi), juntamente com exemplos retirados do *corpus*:

- (1) “*Sujeito Agente* – é o que por si mesmo desencadeia uma atividade (física ou não) sendo origem dela e seu controlador. É, portanto, o que realiza ou instiga a ação expressa pelo verbo que funciona como núcleo do predicado. Constitui-se de um nome marcado pelo traço animado, realizando-se, na frase, por um nome de animal (*O cavalo corre*); de pessoa (*o tenor canta*); de instituição humana (*O Brasil canta*); de parte do corpo (*seus olhos sorriem*); de atributo dos seres (*A beleza da moça mexe comigo*); de atividade dos seres (*seus passos o conduziam a um lugar estranho*).”

- (166) *cêis foj* até o acampamento? (I.2,p.24)
 (167) *ah elis fala* de sombração fazenu baruio aí pa rua (I.8,p.108)
 (168) *essis missionáriu* qui *andava* pu mundu tudu (I.10,p.148)
 (169) *as professora* (*daqueli tempu*) *trabalhava* mais (I.11,p.161)
 (170) *os bicho saía* memo (I.7,p.99)

- (2) “Sujeito Beneficiário – é aquele que é sede da transferência de posse ou destinatário de um benefício: Lucia tem (ganhou, encontrou) um gato siamês.”

(171) eles já ganharo ainda já: :: exposição aí (I.6,p.86)

- (3) “Sujeito Causativo – é o que provoca um efeito ou, então, é o responsável pela realização do estado de coisas indicado no verbo. Ex.: O vento derrubou a árvore; o medo afugentou o rapaz. Expressa-se por um nome não animado (abstrato/concreto). Superficialmente, pode também realizar-se por uma oração infinitiva (...); ou então por nome animado no lugar de seu atributo (ex.: Pedro me decepcionou [=a atitude de Pedro]; Filhotes de animais atraem as crianças [=a beleza, a graça dos filhotes de animais]).”

(172) então tratá de *otas coisa* que me dá mai um lucrinho aí né? (I.4,p.49)

(173) *desses terremotu* qui dá qui mata (quantia de) pessoa (I.10,p.141)

- (4) “Sujeito Experimentador – é aquele que expressa uma experiência ligada a uma disposição mental (Maria lembra-se com saudade da infância), uma sensação (O sapo vê a cobra), uma emoção (Maria sente a morte da avó), uma cognição (Maria sabe tocar piano).”

(174) itaguara *cêis* sabe onde é que é né? (I.1,p.7)

(175) purqui *us otru* nu gostava di iscola né? (I.10,p.140)

(176) *elis* nu quiriam genti casadu aquela tempu (I.15,p.229)

(177) nu sei *cêis* já ouviram falá num problema que teve aqui em são tiago (I.1,p.4)

(178) *cêis* viram a fita? (I.1,p.6)

(179) mais o boi o senhor pega *elis* aí brabu... o senhor insina *elis* (...) *elis* aprendi (I.7,p.98)

(180) *elis* nu lembram não (I.11,p.160)

- (5) “Sujeito Inativo – suporte de uma propriedade, condição ou situação expressa pelo predicado.”

(181) porque tá tão fácil *as coisa* (I.15,p.221)

(182) *as moça daqui* nu foi muito casadera não (I.1,p.7)

(183) *meus pais* eram daqui (I.1,p.10)

- (184) *elas tava na escola (I.1,p.11)*
 (185) *meus irmão morava lá nu novu osasco (I.12,p.184)*
 (186) *falta treis/ treis meis pra mim enterá setenta anu (I.2,p.31)*
 (187) *comé qui chama essas coisa aí? (I. 1, p.4)*

(6) “Sujeito Paciente – é o afetado por aquilo que o verbo expressa (a rosa branca *murchou*). É o que sofre uma mudança de estado, condição ou posição. Ex.: A manteiga *derreteu*; o pão *esfarelou*.”⁶¹

- (188) *o dia que eles foram sepultado (I.1,p.5)*
 (189) *tive treze filhos...mais criô nove (I.1,p.6)*
 (190) *chegava essas hora assim (I.7,p.100)*
 (191) *i tem *multus carru* qui vai (I.15,p.242)*
 (192) *negóçu de *vinte trinta peão*...carregá de/ defunto na na rede...quase morria de tanto fazê força (I.4,p.53)*

Estabelecemos a hipótese geral de que o sujeito Agente condicionaria o uso da regra de CV, uma vez que o traço atividade o tornaria mais saliente para o falante; o sujeito Inativo, por outro lado, inibiria a aplicação da concordância. Os resultados são expostos na tabela 14:

Fatores	Frequência	Probabilidade
1 – Sujeito Agente	86/279 = 31%	.47
2 – Sujeito Beneficiário	2/4 = 50%	.84
3 – Sujeito Causativo	0/6 = 0%	-
4 – Sujeito Experimentador	9/33 = 27%	.41
5 – Sujeito Inativo	19/122 = 16%	.70
6 – Sujeito Paciente	9/76 = 12%	.30
Total	125/520 = 24%	

Tabela 14: frequência e probabilidade de aplicação da regra de concordância verbal (P3) de acordo com o papel semântico do sujeito

Observando a tabela 14, vemos que a hipótese estabelecida é confirmada pelos índices de frequência, mas não pelos valores de probabilidade. Se examinarmos estes resultados separadamente, temos duas escalas diversas de papéis semânticos favorecedores/inibidores da concordância:

⁶¹ Não encontramos no *corpus* ocorrências de sujeito Factivo, mencionado por Borba (1991).

a) resultado fornecido pelos índices de freqüência:

+CV	/	-CV
Beneficiário - Agente - Experimentador	/	Inativo - Paciente - Causativo

b) resultado relativo aos valores de probabilidade:

+CV	/	-CV
Beneficiário - Inativo	/	Agente - Experimentador - Paciente

Na tentativa de explicar tal padrão, analisamos detalhadamente os resultados encontrados à luz das informações fornecidas por Borba (1991). Realizamos, além disso, o cruzamento do papel semântico com os fatores traço semântico do sujeito e valor semântico do verbo⁶². Os resultados desta análise são expostos nos parágrafos subseqüentes.

Começemos pelos casos de **sujeito Agente**. Este é ativo, e o verbo de ação ou ação-processo que o acompanha indica um fazer por parte desse sujeito. Estabelecemos então a hipótese de que estes traços o tornariam mais perceptível para o falante, favorecendo o uso da regra de CV. Tal suposição é confirmada apenas pela freqüência de concordância, que se mostra um pouco maior do que a taxa geral de aplicação encontrada no *corpus*. O valor de probabilidade, porém, indica que o sujeito Agente não constitui um contexto decisivo, condicionando de forma discreta a não-realização da regra.

Vale ressaltar que o sujeito Agente é sempre [+animado], traço favorecedor da concordância. Em 94% dos casos, é expresso por referente [+humano]. Este último traço, contudo, parece não ser relevante nos casos de sujeito Agente, já que a freqüência de concordância se mantém praticamente igual, seja o sujeito [+humano], ou [-humano] – nestes dois contextos constatamos os índices de respectivamente 31% e 29% de aplicação.

Passemos às instâncias de **sujeito Beneficiário**, o qual figura com verbos de processo. Os resultados, tanto de freqüência como de probabilidade, indicam que este constitui um papel semântico que favorece decisivamente a realização da regra de CV. O número reduzido de ocorrências de sujeito Beneficiário encontradas no *corpus* – apenas 4, todas expressas por referente [+humano] – não nos permite, contudo, fazer uma afirmação categórica a este respeito.

O **sujeito Causativo**, por sua vez, se constrói com um verbo de ação-processo, analogamente ao sujeito Agente. Ao contrário deste último, todavia, ele é sempre portador do traço

⁶² O fator valor semântico do verbo não foi selecionado como estatisticamente significativo. Notamos, todavia, que a freqüência geral de realização da regra com verbos de ação e ação-processo, respectivamente 29% e 33%, é maior do que a encontrada nos casos de verbos de processo (17%) e de estado (17%).

[-animado], o que explica o fato de a não-realização da concordância se mostrar categórica neste contexto.

Com relação às ocorrências de **sujeito Experimentador**, verificamos que apesar de a frequência de aplicação ser ligeiramente maior do que a taxa geral de concordância encontrada no *corpus*, o valor de probabilidade indica que este papel semântico desfavorece a realização da regra.

O sujeito Experimentador é, como o sujeito Agente, sempre portador do traço [+animado]. Em 76% dos casos, é expresso por referente [+humano], contexto no qual há alta frequência de aplicação da regra (36%). Nas poucas instâncias de sujeito Experimentador [-humano], por outro lado, a não-concordância se mostra categórica. Isto indica que, ao contrário do que acontece nos casos de sujeito Agente, o traço [+/-humano] tende a interferir no índice de concordância com sujeito Experimentador.

Este índice também difere conforme se considere o valor semântico do verbo. O sujeito Experimentador pode figurar tanto com verbos de processo (*aprender, ver, ouvir*), como de estado (*saber, conhecer, gostar*). No primeiro caso, a frequência de concordância é de 39%; já no segundo, apenas de 13%.

O **sujeito Inativo**, por seu turno, funciona como mero suporte de propriedades, acompanhando um verbo que expressa somente um estado. Por isto, estabelecemos a hipótese de que a ausência do traço atividade tornaria o sujeito menos saliente para o falante, desfavorecendo o uso da concordância. Tal formulação é confirmada somente pela frequência, que se apresenta menor do que a taxa geral de aplicação da regra encontrada no *corpus*. O valor de probabilidade, por outro lado, indica que este papel semântico favorece a concordância.

Analisando os casos de sujeito Inativo conforme seu traço semântico, verificamos que a maioria é expressa por referente [+animado] (76%) e [+humano] (71%). Há diferença significativa no índice de realização da concordância conforme se trate de referente: [+humano] (21%), [-humano] (3%); [+animado] (20%), [-animado] (0%).

No que se refere às ocorrências de **sujeito Paciente**, por fim, verificamos que os resultados, tanto de frequência como de probabilidade, indicam que este constitui um papel semântico que inibe decisivamente a realização da regra de CV. O sujeito Paciente é afetado por um processo que está fora dele e que o faz sofrer uma mudança de estado, condição, ou posição. Talvez tal função o torne menos saliente para o falante, fazendo com que a regra de CV seja menos frequentemente aplicada. Entre as ocorrências de sujeito Paciente, há 60% de referentes [+animado], e 53% [+humano]. Analogamente ao sujeito Inativo, os valores de frequência de aplicação da concordância com sujeito Paciente diferem substancialmente conforme seu traço semântico: [+humano] (20%), [-humano] (3%); [+animado] (20%), [-animado] (0%).

4.2.2 CONCORDÂNCIA DE 1ª PESSOA DO PLURAL: FATORES LINGÜÍSTICOS

No quadro a seguir, mostramos os fatores lingüísticos utilizados na análise da concordância de 1ª pessoa do plural:

Fatores lingüísticos	Categorias
1) Acentuação da forma verbal	2 - forma paroxítona (ex.: <i>falamos</i>) 3 - forma proparoxítona (ex.: <i>falávamos</i>)
2) Presença/ausência do sujeito pronominal <i>nós</i>	n - não-pronominal p - pronominal explícito z - pronominal não-explícito
3) Distância do referente do sujeito zero	a - referente anterior próximo b - referente anterior distante c - sem referente explícito no discurso anterior
4) Posição e distância do sujeito em relação ao verbo	A - anteposto contíguo (SV) D - anteposto distante (S ...V) P - posposto contíguo (VS) Z - posposto distante (V... S)
5) Função discursiva do sujeito	T - Tema R - Rema
6) Status informacional do sujeito	D - Dado N - Novo
7) Paralelismo discursivo (entre verbos)	p - verbo precedido de verbo com marca de plural z - verbo precedido de verbo sem marca formal de plural u - verbo isolado ou primeiro de uma série
8) Valor semântico do verbo	A - Ação P - Processo 2 - Ação-Processo E - Estado C - Copulativo
9) Papel semântico do sujeito	1 - Agente 2 - Beneficiário 4 - Experimentador 5 - Inativo 6 - Paciente

Quadro 6: fatores lingüísticos utilizados na análise da concordância de 1ª pessoa do plural (P1)

O quadro abaixo exhibe os fatores selecionados e descartados pelos programas estatísticos (os fatores selecionados são apresentados por ordem de importância). Nos itens subsequentes, trataremos dos fatores que se mostraram numericamente significativos.

Fatores selecionados	Fatores descartados
1) Acentuação da forma verbal	2) Presença/ausência do sujeito pronominal <i>nós</i>
7) Paralelismo discursivo	4) Posição e distância do sujeito em relação ao verbo
6) Status informacional do sujeito	5) Função discursiva do sujeito
3) Distância do referente do sujeito zero	8) Valor semântico do verbo
	9) Papel semântico do sujeito

Quadro 7: fatores lingüísticos (P1) selecionados e descartados pelos programas estatísticos do Varbrul

4.2.2.1 ACENTUAÇÃO DA FORMA VERBAL

Este fator foi embasado nos estudos de Rodrigues (1987) e Bortoni-Ricardo (1985). Conforme as autoras, estabelecemos a hipótese de que formas verbais paroxítonas (ex.: *falamos*) favoreceriam a aplicação da regra de concordância, e formas proparoxítonas (ex.: *falávamos*) a desfavoreceriam. A tabela 15 contém os resultados encontrados:

Fatores	Freqüência	Probabilidade
2 – forma verbal paroxítona	103/154 = 67%	.94
3 – forma verbal proparoxítona	2/163 = 1%	.07
Total	105/317 = 33%	

Tabela 15: freqüência e probabilidade de aplicação da regra de concordância verbal (P1) de acordo com a acentuação da forma verbal

Vemos que a acentuação da forma verbal exerce um papel decisivo no condicionamento da concordância. Quando a forma verbal é proparoxítona, a não-realização da regra é praticamente categórica (99%), independentemente da escolaridade do informante⁶³. Já quando é paroxítona, a freqüência de aplicação se mantém alta (67%), ainda que os escolarizados apresentem um índice bem maior de concordância (70%) do que os analfabetos (47%).

⁶³ É interessante notar que as únicas ocorrências de verbo flexionado com acentuação proparoxítona (2/163) ocorrem com o verbo *ser* na fala de duas informantes escolarizadas: *nós éramos oito irmãos* (I.1,p.4); *nói éramu oitu* (I.10,p.138).

Rodrigues também constata que a desinência *–mos* é “*mais perceptível, e por isso mesmo mais usada, nas formas paroxítonas, em contraste com seu menor grau de perceptibilidade nas formas proparoxítonas, ainda que, nos dois contextos, corresponda a sílabas átonas finais*” (1987:151). A autora verifica igualmente que, quando a forma verbal esperada é proparoxítona, a não-concordância se mostra praticamente categórica no português popular: 99% (182/184)⁶⁴.

Conforme atestam Ferreira Netto (2001) e Câmara Jr. (1985), as palavras paroxítonas resultam das regras de acentuação próprias da língua portuguesa e são, na sua maioria, originárias do latim vulgar. Já as formas proparoxítonas fogem ao padrão acentual, tendo origem em empréstimos lexicais tardios, posteriores à formação das regras acentuais do português. Alguns autores (cf. Câmara Jr, 1985; Lemle 1978, In: Rodrigues, 1987) notam que na variedade popular há uma tendência de redução de palavras proparoxítonas a paroxítonas, pela perda de um ou mais segmentos fonéticos. No *corpus* sob análise, encontramos exemplos desta redução também em substantivos: (i) supressão da vogal postônica associada à perda da consoante que a segue (*árvore>arve*); (ii) não-realização da sílaba final (*química>quimi*). Segundo Rodrigues (1987), a redução que atinge as formas verbais constitui um caso especial, já que se trata de um segmento silábico completo, com valor mórfico, que deixa de ser realizado.

Uma vez que a concordância praticamente não se realiza com formas verbais proparoxítonas, uma hipótese funcional estabelece que o falante explicitará o sujeito mediante outras estratégias. Voltaremos a tratar desta questão nos dois itens subseqüentes.

4.2.2.2 PARALELISMO DISCURSIVO

Na análise do fator paralelismo discursivo⁶⁵ em ocorrências de sujeitos de 1ª pessoa do plural, seguimos exatamente os mesmos critérios adotados para a 3ª pessoa do plural, estabelecendo hipótese idêntica: verbo precedido de verbo no plural tenderia a favorecer a aplicação da regra de concordância, e verbo precedido de verbo sem marca formal de plural a desfavoreceria. Os resultados são exibidos na tabela a seguir:

⁶⁴ Em formas verbais paroxítonas, a freqüência de não-concordância entre falantes urbanos é de apenas 27% (Rodrigues, 1987:151).

⁶⁵ Conforme mencionado no item 4.2.1.3, o fator paralelismo discursivo foi baseado no estudo de Scherre & Naro (1993), o qual o analisa somente em ocorrências de 3ª pessoa do plural.

Fatores	Frequência	Probabilidade
p – verbo precedido de verbo com marca de plural	33/46 = 72%	.87
z – verbo precedido de verbo sem marca formal de plural	14/113 = 12%	.43
u – verbo isolado ou primeiro de uma série	58/158 = 37%	.41
Total	105/317 = 33%	

Tabela 16: frequência e probabilidade de aplicação da regra de concordância verbal (P1) segundo o paralelismo discursivo⁶⁶

Os valores de frequência e probabilidade confirmam a hipótese: o verbo precedido de verbo no plural condiciona decisivamente a realização da concordância; e o verbo precedido de verbo no singular, a não-realização da regra. Observemos agora os exemplos abaixo, os quais ilustram de maneira clara como o fator paralelismo discursivo atua no *corpus*:

(193-206) mais nós ia di manhã cedu... i à tardi armá a urupuca eu cu meu irmão essi um qui faleceu agora faiz treis anu...uns treis anu mais o menu...i nós ia à tardi levá cumida pa pô dibaxu da urupuca armava a urupuca...otu dia cedu ...((corte na fita)) daí ela: :: di manhã cedu (nóis) () sem tomá café sem nada () quandu meis di julho () tava cheiu di jada... ai nós levantava nós dois ia lá paqueles fundão lá visitá a urupuca quandu dava achava passarinhu dibaxu da urupuca nós ficava feliz da vida era: :: (lambura) juruti debaxu da urupuca... i daí nós ficava contenti...quandu é di tardi.../Inf.2 (brincadeira) sem graça dimai/ Inf.nóis ficava contenti quandu achava: :: passarinhu dibaxo da urupuca...até nós:: tirava u passarinhu com cuidadu pa nu iscapá... i depoi quandu era di tardi nós tornava í otra vez levava cumida i tornava armá () a urupuca (I.10,p.138)

(207-211) curimbatá é u pexi mai gordo qui injisti na terra...cê qué sabê...nói fomu lá cuma penera () tava danu dificultadi pa nói qui qui nói fizemu? eu e u otru ermãu meu...rumemu um ramu di: :: gaiu di made/...di arvi sabi? i cumecemu puxá pra lá i pra cá...daqui a pocu eles começaru ficá tonto naquela água suja né? eles saíru pra cima...ali nói pegava cua mãe (I.9,p.128)

⁶⁶ A categoria (u) não constitui um contexto relevante para a análise do paralelismo discursivo. O índice relativamente alto de concordância encontrado (37%) se deve provavelmente a um outro fator: o status informacional do sujeito, já que 97% dos referentes que acompanham verbo isolado ou primeiro de uma série veiculam informação nova (cf. item 4.2.2.3). O programa estatístico, levando em conta esta sobreposição entre fatores, atribuiu à categoria (u) uma probabilidade baixa de concordância.

Note-se que em (193-206), há uma seqüência de formas verbais não-marcadas, todas de acentuação proparoxítona. Já em (207-211), ocorre a utilização sucessiva de verbos marcados de acentuação paroxítona; o último verbo de acentuação proparoxítona, no entanto, não aparece flexionado. Estes exemplos evidenciam que o fator paralelismo discursivo só é relevante se a forma verbal for paroxítona. Conforme visto nas observações relativas ao fator acentuação, quando a forma verbal é proparoxítona, a não-concordância é praticamente categórica, não importa se o verbo é precedido de plural, singular, ou se é único, como demonstra a tabela 17 abaixo:

Fatores	Frequência	
	Forma verbal proparoxítona	Forma verbal paroxítona
p – verbo precedido de verbo com marca de plural	0/10 = 0%	33/36 = 92%
z – verbo precedido de verbo sem marca formal de plural	0/83 = 0%	14/30 = 47%
u – verbo isolado ou primeiro de uma série	2/70 = 3%	56/88 = 64%

Tabela 17: frequência de aplicação da regra de concordância verbal (P1) conforme o paralelismo discursivo e a acentuação da forma verbal

Os valores expostos nesta tabela oferecem uma visão mais clara dos dados. Estes mostram que o paralelismo discursivo tem peso decisivo no contexto forma verbal paroxítona precedida de verbo no plural, no qual constatamos alta frequência de aplicação da concordância (92%). Nas formas verbais paroxítonas precedidas de verbo no singular, a frequência de aplicação da concordância é bem menor, ainda assim se mantém em um índice elevado (47%).

Conforme visto na análise das ocorrências de P3, estabelecemos a hipótese de o paralelismo discursivo não ser de natureza contrafuncional. Nos exemplos (193-206) mencionados acima, a utilização de uma seqüência de verbos não-marcados – que é, como vimos, condicionada pela acentuação da forma verbal, mas ainda assim constitui uma repetição de formas sem flexão – não compromete a informação, uma vez que o referente de tais verbos é explicitado diversas vezes no trecho do discurso. Mediante o cruzamento dos fatores paralelismo discursivo e status informacional do sujeito, verificamos que nestas instâncias de verbo precedido de verbo no singular, o sujeito constitui informação dada em 76% dos casos.

Pode-se argumentar, por outro lado, que em exemplos como (207-210), a cadeia de verbos marcados veicula informação redundante. Todavia, examinando as ocorrências, notamos que o fato de um referente se manter como Tema na estrutura narrativa do discurso não o torna necessariamente informação dada. No contexto verbo precedido de verbo com marca de plural, há 54% de casos em que o sujeito constitui informação nova. Voltaremos a esta questão no item 4.2.2.3 a seguir.

4.2.2.3 STATUS INFORMACIONAL DO SUJEITO

Este fator foi sugerido pela obra de Halliday (1994). Para determinar se o sujeito constitui informação dada ou nova, seguimos os critérios adotados pelo autor, considerando os contextos lingüístico e situacional, e a proeminência tônica que caracterizaria o elemento novo.

Entendemos que a informação relativa ao referente de um verbo pode ser veiculada tanto através da flexão número-pessoal, como através do sujeito explícito. Nos casos de sujeito Dado, ambas as estratégias são desnecessárias. Já nas instâncias de sujeito Novo, ao menos uma das duas formas de explicitação do referente terá de ser utilizada.

Na análise do status informacional do sujeito, procuramos fazer com que este fator não se sobrepusesse sobre o fator presença/ausência do sujeito pronominal. A questão foi então colocada do seguinte modo: o referente de um determinado verbo – independentemente de o sujeito ser explícito ou zero, independentemente de o verbo estar ou não flexionado – constitui informação conhecida ou nova? Um teste que se mostrou válido para determinar o status informacional do sujeito consiste em tirar a flexão do verbo e o sujeito (se estiverem presentes). Se ainda assim for possível depreender o referente do verbo, este constitui informação dada. Observa-se o discurso como um todo e ao mesmo tempo a unidade discursiva analisada.

Mais adiante cruzaremos o status informacional do sujeito com o fator presença/ausência do sujeito pronominal. Por este motivo, exemplificaremos os critérios adotados para a classificação do referente do verbo como informação dada ou nova, tomando por base as três categorias do fator presença/ausência do sujeito pronominal: (i) sujeito não-pronominal, (ii) sujeito pronominal explícito, (iii) sujeito pronominal não-explícito.

(I) Sujeito não-pronominal

Há somente 7 casos de sujeito não-pronominal (i.e. sujeito composto de 1ª pessoa do plural) no *corpus*, a maioria constituindo informação nova, pelo fato de ser a primeira vez em que aparecem no discurso – há apenas uma instância em que o sujeito composto já tinha figurado no discurso, equivalendo portanto a informação dada.

(II) Sujeito pronominal explícito

Nas ocorrências em que o pronome pessoal *nós* é anafórico, ou já foi mencionado no discurso anterior próximo, como em (212), este constitui informação dada:

(212) abóbora...*nóis* planta...*nóis* planta abóbora aí (I.2,p.34)

Em grande parte das ocorrências de sujeito pronominal encontradas no *corpus*, porém, *nós* veicula informação nova. Isto ocorre em três casos: (1) o pronome não tem um antecedente de 1ª pessoa do plural presente no discurso; (2) o pronome tem um antecedente, mas a não-realização do pronome (ou da flexão) implicaria ambigüidade na interpretação do referente; (3) o pronome é entonacionalmente ou estruturalmente marcado.

A categoria (1) se compõe de instâncias em que não há um SN de 1ª pessoa do plural expresso no discurso anterior à ocorrência de sujeito pronominal. Isto ocorre nos casos de:

- (i) *nós* catafórico⁶⁷, remetendo a um referente posterior, como em (213);

(213) *nóis* ia im itapecerica *eu i a madalena* fazê compra (I.12,p.187)

- (ii) *nós* exofórico, referindo-se às pessoas presentes na situação, como em (214);

(214) tá só *nóis* aqui não tem problema de falá (I.9,p.123)

- (iii) *nós* anafórico remetendo a dois SNs no singular expressos anteriormente, mas não articulados em sujeito composto, como em (215). Note-se que, neste caso, *nós* veicula informação nova, ainda que seja anafórico. Em (216), é justamente o pronome explícito *nós* que esclarece que o referente anterior constitui um sujeito composto ligado pela partícula *com* (e não um sujeito de 1ª pessoa do singular), constituindo, portanto, informação nova.

(215) *eu eu* fui *eu* fui sê testemunha de um casamento...*a moça* me me convidô pa sê testemunha dela e tudo e ... e o pessoal dela num tava querendo muito o casamento não...*ái nós* fomo na casa do sítio do pai do pai dela (I.6,p. 91)

(216) *eu cu joaquim meu irmão* *nóis* vivia remendanu eli (I.12,p.180)

- (iv) *nós* sem referente discursivo. Em tais instâncias, o pronome não remete a um SN no plural expresso no discurso anterior ou posterior à ocorrência, como no exemplo (217), em que *nós* se refere ao informante e a outra pessoa que não é explicitada, podendo ser sua esposa ou alguém que o ajuda na roça. É interessante notar que o falante pode inclusive mudar a referência de 1ª pessoa do singular para o plural, sem explicitar o outro referente que se junta ao *eu*, como acontece em (218);

⁶⁷ Em nossa análise, levamos em conta o tempo do discurso, encarando este último como um processo. Isto implica que só são considerados os referentes anteriores à ocorrência, para classificar como Dado e Novo.

(217) *nói corta* argum pa gastá (I.4,p.48)

(218) Doc. o senhor não vai pra ubatuba?/Inf. vô: :: sempri *nóis vai* (I.5,p.59)

Cabe ainda notar que do ponto de vista exclusivo da informação, poderíamos considerar *a gente* como correferente de *nós*, assim se *nós* tivesse como antecedente *a gente* referindo-se ao mesmo sujeito, seria Dado. Como estamos estudando a concordância verbal, e *a gente*, embora semanticamente seja 1ª pessoa do plural, constitui de um ponto de vista morfológico um sujeito de 3ª pessoa do singular, este não foi considerado como antecedente de *nós*, ex.:

(219) mandô trazê *a gente* pra qui...oito hora da noite quando *nói chegamo* aqui (I.1,p.15)

A categoria (2) engloba instâncias em que o pronome tem um antecedente de 1ª pessoa do plural expresso no discurso, mas a realização do pronome evita uma possível ambigüidade na interpretação do referente do verbo. Isto pode acontecer nos casos em que:

- (i) há mudança de referente entre uma e outra ocorrência do pronome. Mesmo sendo anafórico, caso a última menção do pronome, ou do sujeito composto ao qual o pronome remeta, esteja distante da ocorrência considerada e houver mudança de referente, muitas vezes faz-se necessária a explicitação do referente do verbo novamente, seja através do pronome explícito, ou da flexão, ex.:

(220) *nóis num/ tem um sobrinhu* que mora lá ...*nóis num vendeu* lá (I.8,p.106)

- (ii) há dois referentes possíveis para o verbo em questão, como em (221), em que se o pronome não fosse explícito, o referente do verbo poderia ser interpretado como *eu*:

(221) achu qui nessa primera veiz *queu* vi eli... () lobisomi que minha mãe sem/ sempre falava pa *nóis...nóis* nu nu acreditava (I.9,p.131)

- (iii) a ausência do sujeito pronominal dá margem para uma interpretação do referente do verbo como indeterminado. Em (222), por exemplo, se o falante dissesse *matou duas*, a casa vazia do sujeito teria sido interpretada como de referência indeterminada e a ocorrência sequer teria sido incluída na análise – note-se que, do ponto de vista do paralelismo discursivo, trata-se de uma ocorrência de verbo precedido de verbo com

marca de plural, mas o sujeito não constitui informação dada e tem necessariamente de ser explicitado através do pronome ou da flexão do verbo:

(222) aqui na estrada memo otu dia umas par de veis nós matemo cascavel nessa estrada aqui vem do artu pa bebê água nu riberãoné? lá de vorta ela no barro ela num sobi...ali memo fai uns doi meis *nói* matemo dua nu pé daqueli pau lá naquele naquele postu lá (I.4,p.57)

(iv) o pronome estabelece contraste com outro referente, ex.:

(223) *eli* dava leiti pa nós *nóis* fornecia a pai di mio (I.12,p.180)

O contexto (3), por fim, compõe-se de casos em que o pronome é marcado pela entonação, como em (224), ou pela estrutura clivada: *ser* SN *que* X / SN *é que* X / SN *que* X, como em (225-227). Entendemos que estes constituem dois modos de marcar o SN sujeito como informação nova. É interessante notar que neste caso o sujeito tem necessariamente de ser explícito⁶⁸.

(224) “ói esse lumbriguero...nós temos que dá nossos filho amanhã...i *nós* **tamém** vamo tomá” (I.1,p.4)

(225) *foi foi nós* *que* fizemo (I.2,p.34)

(226) *nóis é que* tinha que fazê o (nosso) separado (I.3,p.40)

(227) fico comendo o churrasco que *nói* paguemo ... ((risadas)) *nói* *que* paguemo (I.5,p.72)⁶⁹

(III) Sujeito pronominal não-explicito

Quando a casa vazia do sujeito retoma SN de 1ª pessoa do plural já mencionado, o sujeito geralmente constitui informação dada, seja o verbo flexionado, como em (228), ou não-flexionado, como em (229):

(228) *nói* vendia muito ovo agora nu: :: nu tamo lidando mais com galinha botadera (I.2,p.21)

(229) *nói* lida só pa bebê... fai quezo ora de viz em quano (I.5,p.68)

⁶⁸ Note-se que se o pronome *que* figurar em uma oração relativa, retomando um referente da oração principal, ele veicula informação dada, ex.: são tiago hoje é uma cidade boa pra *nóis* *que* somo aqui...desse lugarzinho pobre né? (I.1,p.15).

⁶⁹ Encontramos exemplos de estrutura clivada também com a utilização do sujeito composto: *eu ca minha muié* *qui* *qui* lutava (I.14, p.212).

Há instâncias em que a flexão do verbo se faz necessária para evitar a ambigüidade na interpretação do referente da casa vazia do sujeito. Neste caso, o sujeito constitui informação nova, ainda que tenha sido mencionado anteriormente. Em (230), por exemplo, se o verbo estivesse na 3ª pessoa do singular, poderíamos interpretar seu referente como *a cidade de São Tiago*. É interessante notar que se analisássemos as duas ocorrências sucessivas de verbo no plural exclusivamente do ponto de vista do paralelismo discursivo, concluiríamos erradamente que a marca do segundo verbo é redundante:

- (230) *são tiago* hoje é *uma cidade* boa pra *nóis* que somo aqui...*desse lugarzinho pobre* né? graças a deus temos bons professores boas escola...bons padre (I.1,p.15)

Em outros casos, a flexão evita uma interpretação do referente do verbo como indeterminado – conforme já mencionamos, há na verdade um número considerável de instâncias em que se não houvesse o pronome explícito ou a flexão do verbo quando o pronome não é explícito, a ocorrência sequer teria sido considerada na contagem, já que o sujeito teria sido interpretado como indeterminado; ex.:

- (231) Inf.é...isso canarvial (essas coisa) nem tem...mais pra consumo do gado memo qui *nói* lida com ele/ Doc.1 do gado a senhora () / Inf.tira leite...qué dizê o gado nossa vendemo tudo (I.2,p.19)

Quando não há um SN de 1ª pessoa do plural expresso no discurso anterior à ocorrência, o sujeito constitui informação nova que é explicitada mediante a flexão⁷⁰. Este corresponde ao contexto (c) do fator distância do referente do sujeito zero (item 4.2.2.4 adiante) e se compõe de 13 ocorrências. Destas, 11 são de verbo flexionado. Note-se que o verbo é quase que categoricamente marcado, porque se não o fosse, não teria sido incluído entre as ocorrências. As únicas instâncias em que o verbo está no singular, transcritas em (234-235), correspondem a uma seqüência, e foram interpretadas como informação dada. Os casos de sujeito zero sem referente de 1ª pessoa do plural explícito no discurso anterior são exemplificados a seguir:

⁷⁰ Na 3ª pessoa do plural, a casa vazia do sujeito mais verbo no plural, sem referente mencionado no discurso ou presente na situação, foi descartada por tratar-se de sujeito indeterminado. Na 1ª pessoa do plural, no entanto, a estrutura *Sujeito Zero + Verbo marcado* foi incluída entre as ocorrências.

(i) menção do pronome de 1ª pessoa do singular e de um nome lexical, mas não-articulados como sujeito composto – a casa vazia do sujeito mais o verbo marcado de 1ª pessoa do plural remetem aos dois referentes mencionados, que ainda não tinham sido combinados em um só sujeito;

(232) eles acha que *eu* mereço melhor né? (...) graças a deus...vivê cinquenta e quatro ano casada...com *um maridinho BRAvo* mais graças a deus vivemo muito bem (I.1,p.8)

(ii) menção implícita do sujeito – a casa vazia do sujeito mais o verbo têm um referente implícito, mas não mencionado no discurso, como em (233-235). Estas são as únicas instâncias em que o verbo aparece sem marca. Interpretamos a flexão como menção do referente, classificando o sujeito do primeiro verbo da seqüência como informação nova, e o do segundo e terceiro verbos, como informação conhecida:

(233-235) Doc. o senhor lembra da festa do casamento? /Inf. ah nu fui nu tivi festa/
Doc. ah nu feiz?/ Inf. ah nu fizemu ah casa muito pobri nu tinha qui fazê nada né? (I.15,p.237)

(iii) sujeito exofórico – a casa vazia do sujeito mais o verbo marcado de 1ª pessoa do plural têm seus referentes presentes no contexto situacional: o falante utiliza esta estrutura para incluir aos documentadores e a si mesmo como referente do verbo (seguindo o critério descrito no item anterior, consideramos o referente do primeiro verbo como informação nova, e o do segundo, como informação dada);

(236-237) nu meu tempu era ótimo tinha SAÚ: ::di....hoje im dia já tamu tamu podi duenti (I.5,p.59)

(iv) referente mencionado posteriormente à ocorrência;

(238) então *eu* fui nu casamentu da bevelina (...) aí fomu nu casamentu ...quando casô a gente foi...*u vosso avô....meu bisavô* ...foi *u grigório* (...) i foi *u bevelinu* (I.12,p.183)

Estabelecidos os critérios, formulamos a hipótese funcional de que o sujeito Dado condicionaria o uso de formas verbais não-marcadas; e o sujeito Novo, a aplicação da concordância. Os resultados relativos ao fator status informacional do sujeito são exibidos na tabela 18:

Fatores	Frequência	Probabilidade
D – sujeito Dado	20/111 = 18%	.22
N – sujeito Novo	85/206 = 41%	.67
Total	105/317 = 33%	

Tabela 18: frequência e probabilidade de aplicação da regra de concordância verbal (P1) conforme o status informacional do sujeito

Os valores de frequência e probabilidade mostrados na tabela evidenciam que a hipótese foi confirmada. Vemos que a concordância é favorecida, quando o sujeito constitui informação nova, e decisivamente desfavorecida, se o referente equivale a informação dada.

Já mencionamos que procuramos fazer com que este fator não se sobrepusesse sobre o fator presença/ausência do sujeito pronominal. Em outras palavras, o referente de um verbo poderia ser considerado Dado ou Novo independentemente de o sujeito ser ou não explícito. Ora, a informação relativa ao referente do verbo pode ser veiculada tanto pelo sujeito formalmente expresso na oração, como através da flexão número-pessoal. Interessou-nos então verificar qual era o padrão de realização do sujeito, conforme seu status informacional. Vejamos os resultados:

Fatores	Frequência	
	Sujeito Dado	Sujeito Novo
n - sujeito não-pronominal	1/1 = 100%	2/6 = 33%
p - sujeito pronominal explícito	7/33 = 21%	55/171 = 32%
z - sujeito pronominal não-explícito	12/77 = 16%	28/29 = 97%
Total	20/111=18%	85/206 = 41%

Tabela 19: frequência de aplicação da regra de concordância verbal (P1) conforme o status informacional do sujeito e a presença/ausência do sujeito pronominal

Examinando a tabela, o primeiro aspecto que nos chama a atenção é a diferença significativa no índice de realização do sujeito, conforme este represente informação dada ou nova. Vemos que nas instâncias de sujeito Novo, 86% (177/206) são formalmente expressos na oração. Já nos casos de sujeito Dado, temos apenas 31% (34/111) de instâncias de sujeito explícito. Estes dados sugerem que o falante popular tenderia a optar pelo sujeito formalmente expresso na oração, quando este equivale a informação nova, e pelo sujeito zero, quando o referente constitui informação conhecida pelo interlocutor.

A tabela 19 mostra também que há um número reduzido de instâncias de sujeito não-pronominal (sujeito composto) no *corpus*. A maior parte dos casos de sujeito explícito é representada pelo pronome pessoal *nós* (204/211=97%). Neste contexto, a informação não é comprometida independentemente de o verbo estar ou não flexionado. Ainda assim, o falante tende a empregar a concordância com uma frequência um pouco maior, quando o sujeito pronominal representa informação nova (32%) do que quando equivale a informação dada (21%).

Passemos aos casos de sujeito pronominal não-explícito. Destes, 73% (77/106) equivalem a informação dada, e 27% (29/106), a informação nova. Nos casos de sujeito zero que corresponde a informação dada, temos somente 16% de frequência de aplicação da regra de concordância, um índice ainda menor do que os que foram constatados com sujeito pronominal explícito. Os exemplos mencionados a seguir ilustram o padrão de concordância encontrado nas instâncias de sujeito zero equivalente a informação dada:

- (239-241) *nói* pega i desci du onibu i vai à pé até lá nu (nu fim) du bairru...pra: :: recordá o tempu qui qui andava pelas istrada (I.10,p.139)
- (242-245) essi aí *nói* fazia serva pa pegá memu...pa criá né? num era pa: :: assim pa pa matá...fazia a serva dipoi ia lá armava a arapuca i pegava (I.9,p.125)
- (246-251) *nóis* crescemo juntu i: :: andava juntu...pulandu pa trepandu im arve pulandu pegava o bambu assim a/ afirmava o bambu nu chão i assim pegava o bambu caía LÁ longe dipoi tornava subí outra vei pa *nói* torná pulá di novu otra veiz (I.10,p.139)
- (252-256) tava chuvendu *nóis* tava batendu a enxada nu parava ia nu ranchu ranchu di capim trocava a ropa botava aquela enxugá lá nu fogu i ia trabaiá (I.13,p.192)
- (257-262) *nóis* plantava lavora nu ia plantá milho ...cincu seis veiz num lugar como planta hoji não...plantava treis anu num lugar largava aquela passava pu otru a fazenda era muito grandi... num anu...nu sigundu anu nu capinava roça não... só chegava terra nu milhu uns pocuzinhu di (terra) (I.13,p.198)

Tais instâncias exemplificam uma estrutura frequentemente utilizada pelo falante popular: *Nós + verbo no singular ou no plural, Zero + verbo no singular, Zero + verbo no singular*. Nestes casos, a hipótese de ordem funcionalista é confirmada: o sujeito que constitui informação dada leva

a uma seqüência de verbos não-marcados com casa vazia do sujeito, evitando a redundância, já que, conforme destaca Rodrigues (1987), a informação relativa ao referente já foi transmitida no pronome explícito da primeira oração.

Já nas 29 instâncias de sujeito zero equivalente a informação nova, a frequência de concordância é de 97%⁷¹. Este contexto engloba, como vimos, ocorrências em que não há um referente de 1ª pessoa do plural expresso no discurso anterior, tornando-se a flexão o único elemento que identifica o sujeito do verbo. Os demais casos correspondem a instâncias em que há um antecedente de 1ª pessoa do plural, mas a flexão do verbo se faz necessária para evitar a ambigüidade na interpretação do referente da casa vazia do sujeito, ou para evitar uma interpretação do referente do verbo como indeterminado. Conforme nota Guy (1981:245), tais casos não podem ser tomados como argumentos a favor da hipótese funcionalista, porque é possível que ocorrências de 1ª pessoa do plural totalmente não-marcadas tenham passado despercebidas, sem terem sido incluídas na análise.

4.2.2.4 DISTÂNCIA DO REFERENTE DO SUJEITO ZERO

O fator distância do referente do sujeito zero⁷² compõe-se de três itens:

- a) referente anterior próximo;
- b) referente anterior distante;
- c) sem referente explícito no discurso anterior.

O procedimento de análise consiste em procurar o referente anterior à ocorrência de casa vazia do sujeito. Havendo um referente anterior no plural, constituído do pronome de 1ª pessoa do plural (*nós*), ou correspondendo a um sujeito composto de pronome de 1ª pessoa do singular mais um nome lexical (*eu ca minha mãe*), classificamos a ocorrência como: (b) se mais de 10 sílabas separarem o referente da forma verbal com sujeito zero, ou (a) se estiver a até 10 sílabas de distância.

⁷¹ Há apenas uma ocorrência de verbo não-marcado: *quando nós viemo praqui é... é nós três né? (I.11,p.165)*. Esta se insere num contexto discursivo em que o documentador pergunta à informante quantas pessoas moravam em sua casa. O sujeito constitui então a informação nova requisitada, sendo explicitado na função de Rema da oração subsequente. Interpretamos a casa vazia do sujeito como informação nova no sentido de não conhecida, na medida em que seu referente só é esclarecido posteriormente.

⁷² Para estabelecer este fator, tomamos por base os estudos de Naro & Scherre (1999), Rodrigues & Campos (2002), e Pereira & Rodrigues (2003), os quais o utilizam para a análise da concordância de 3ª pessoa do plural.

A categoria (c) compõe-se de casos em que não há um SN de 1ª pessoa do plural expresso no discurso anterior à ocorrência, conforme ilustrado nos exemplos (232-238) mencionados no item anterior.

A hipótese de ordem funcionalista estabelecida é a de que o referente distante da forma verbal com sujeito zero favoreceria a CV, uma vez que esta facilitaria a ligação entre o verbo e o sujeito extra-sentencial. A casa vazia do sujeito cujo referente está próximo à ocorrência, por outro lado, condicionaria a não-concordância. Os dados quantificados são apresentados na tabela 20:

Fatores	Frequência	Probabilidade
a – referente anterior próximo	2/24 = 8%	.15
b – referente anterior distante	27/69 = 39%	.64
c – sem referente explícito no discurso anterior	11/13 = 85%	.57
Total	40/106 = 38%	

Tabela 20: frequência e probabilidade de aplicação da regra de concordância verbal (P1) de acordo com a distância do referente do sujeito zero⁷³

Os resultados expostos na tabela 20 confirmam a hipótese funcionalista: o sujeito zero com referente anterior próximo, como em (263-266), inibe decisivamente a concordância, ao passo que o sujeito zero com referente anterior distante, como em (267-269), a favorece. Note-se que, no primeiro caso, a utilização de uma forma verbal não-marcada não compromete a informação, uma vez que o referente se encontra muito próximo da ocorrência considerada; já no segundo, a aplicação da regra de concordância contribui para a identificação do referente distante e para a construção da coerência do discurso.

- (263) di primero *nói* prantava muito ma agora nu pranta mais (I.2,p.19)
- (264) sempri *nói* tá roçano pasto aí mata cada um bruta cascavelzão (I.4,p.57)
- (265) ainda onti *nóis* trabalhô... ó apanhô o café (I.11,p.166)
- (266) *nóis* chegava lá durmia na casa dela (I.12,p.181)
- (267) Inf. dipois...deus ajudô...ele adoeceu...ele adoeceu e *nós* tivemos que mudá di repente aqui pra são tiago/ Doc.1 ((risos)) imagine / Inf. chegamo aqui oito hora da noite com muito sacrifício (I.1, p.11)

⁷³ Note-se que, apesar de a frequência de concordância ser bem maior no contexto (c) do que em (b), o programa estatístico atribuiu maior probabilidade de aplicação a (b) do que a (c). Isto porque levou em conta a relação entre estas categorias e o status informacional do sujeito: em (c), a maioria dos sujeitos (10/13=77%) veicula informação nova; já em (b), predominam os casos de sujeito Dado (51/69=74%).

- (268) aí quando nós saímo do cinema tinha um tinha um barzinho aberto...eh: :: eu fui falei com meu tio falei “eu vô eu vô passá ali...vam/ vamo levá quarquê coisa pa *nóis* tomá café”...eh: :: e aquele tempo o povo num era muito de bebê não né? era mais era café mesmo...aí paramo lá no no no no boteco (I.6,p.80)
- (269) *nói* eramu oitu...oitu...já morreru: :: sei qui agora até já já morreru quatu já agora tamu im quatu (I.10,p.138)

Com relação ao contexto (c) – o qual engloba os casos de sujeito zero sem referente de 1ª pessoa do plural explícito no discurso anterior –, porém, não podemos fazer uma afirmação categórica sobre os resultados encontrados, já que grande parte dos dados corresponde a formas verbais flexionadas. Conforme já mencionamos no item relativo ao status informacional do sujeito, é possível que formas de 1ª pessoa do plural totalmente não-marcadas, com casa vazia do sujeito e verbo no singular, tenham passado despercebidas.

4.2.3 FATORES SOCIAIS

O quadro abaixo contém os fatores sociais (relativos às características sociais dos informantes) que foram utilizados na análise da concordância de 3ª e 1ª pessoas do plural.

Fatores Sociais	Categorias
1) Procedência	G – Minas Gerais S – São Paulo
2) Sexo	M – masculino F – feminino
3) Escolaridade	P – escolarizados (até 4ª série primária) 0 – não-escolarizados

Quadro 8: fatores sociais utilizados na análise da concordância de P3 e P1

O programa do Varbrul selecionou como estatisticamente significativos os seguintes fatores sociais: sexo, procedência, e escolaridade, para as ocorrências de sujeitos de 3ª pessoa do plural; e escolaridade, para as ocorrências de sujeitos de 1ª pessoa do plural.

Relembramos que a nossa pesquisa abrange um número reduzido de informantes. Por este motivo, as relações estabelecidas entre a realização da concordância e os fatores sociais são apenas indicações de um padrão que provavelmente seria encontrado em um grupo mais amplo de falantes. Os resultados são apresentados nos itens subseqüentes.

4.2.3.1 ESCOLARIDADE

Rodrigues (1987:206) analisa a relação entre o grau de escolaridade do informante e a concordância, verificando que este fator social tem peso relevante na 1ª pessoa do plural: falantes analfabetos apresentam maior frequência de não-realização da regra de CV (56%) do que falantes escolarizados (38%). Com relação a 3ª pessoa do plural, porém, o fator não se mostra atuante – a frequência de não-concordância entre falantes escolarizados e não-escolarizados é praticamente a mesma, respectivamente, 70% e 72%.

Baseando-nos neste estudo, estabelecemos a hipótese de que os falantes escolarizados, ainda que possuam apenas de 1 a 4 anos de escolarização⁷⁴, apresentariam maior uso da regra formal de concordância, ao passo que falantes de escolaridade nula tenderiam a utilizar com maior frequência formas verbais não-marcadas. Confirmando a hipótese, o fator se mostrou significativo tanto para a 1ª, como para a 3ª pessoa do plural, conforme evidenciado na tabela a seguir:

Fatores	Sujeitos de 3ª pessoa do plural		Sujeitos de 1ª pessoa do plural	
	Frequência	Probabilidade	Frequência	Probabilidade
P – escolarizados	84/281 = 30%	.57	58/152 = 38%	.62
0 – não-escolarizados	27/174 = 16%	.39	24/117 = 21%	.34
Total	111/455 = 24%		82/269 = 30%	

Tabela 21: frequência e probabilidade de aplicação da regra de concordância verbal (P3 e P1) segundo o grau de escolaridade do informante⁷⁵

Vemos que independentemente da pessoa gramatical, há maior probabilidade de aplicação da regra entre falantes escolarizados do que entre os analfabetos, ainda que a diferença entre um e outro grupo se mostre mais significativa na 1ª pessoa do plural. Este fato indicaria que a influência normativa da escola tem um papel importante na aquisição de formas verbais marcadas de número-pessoa.

⁷⁴ Entre os escolarizados, apenas um informante chegou a fazer o supletivo de 5ª série.

⁷⁵ Foram excluídas 113 ocorrências (65 de 3ª pessoa do plural e 48 de 1ª pessoa do plural) encontradas na fala de uma informante cujo grau de escolaridade era desconhecido.

4.2.3.2 SEXO

Estudos sociolinguísticos em geral normalmente verificam que a mulher tende a se aproximar mais da norma padrão do que os homens. Estudos específicos sobre a concordância verbal, como os de Rodrigues (1987) e Bortoni-Ricardo (1985), comprovam, no entanto, que os homens empregam mais freqüentemente a regra do que as mulheres. No *corpus* estudado por Bortoni-Ricardo, o sexo constitui fator social relevante no condicionamento da concordância tanto de 3ª, como de 1ª pessoa do plural. Já em Rodrigues (1987), o fator só é pertinente na 1ª pessoa do plural.

No *corpus* sob análise, este fator foi relevante apenas para ocorrências com sujeitos de 3ª pessoa do plural. Vejamos os resultados encontrados:

Fatores	Freqüência	Probabilidade
F – feminino	82/276 = 30%	.57
M – masculino	43/244 = 18%	.42
Total	125/520 = 24%	

Tabela 22: freqüência e probabilidade de aplicação da regra de concordância verbal (P3) conforme o sexo do informante

Os valores expostos na tabela evidenciam que as mulheres têm maior probabilidade de aplicar a regra de concordância do que os homens. A confiabilidade destes resultados é, no entanto, prejudicada pelo fato de todas as mulheres cujos dados estudamos – com exceção de uma informante cujo grau de escolaridade era desconhecido – serem escolarizadas. Ainda que tenham permanecido por pouco tempo na escola, possuindo de 1 a 4 anos de escolarização, este período provavelmente condicionou um maior uso de formas verbais marcadas de 3ª pessoa do plural em sua fala. Cruzamos então o fator grau de escolaridade com sexo do informante, chegando aos seguintes resultados:

Fatores	Freqüência	
	Mulheres	Homens
Escolarizados	68/211 = 32%	16/70 = 23%
Não-escolarizados	-	27/174 = 16%

Tabela 23: freqüência de aplicação da regra de concordância verbal (P3) de acordo com o sexo e o grau de escolaridade do informante

Comparando o índice de concordância entre falantes escolarizados de ambos os sexos, vemos que as mulheres ainda superam os homens no emprego da regra. A diferença não se mostra numericamente significativa, mas sugere a hipótese de que a escola exerceria maior influência sobre as mulheres no que toca à realização da regra padrão de concordância verbal.

4.2.3.3 PROCEDÊNCIA

Rodrigues (1987:213) verifica que os falantes nascidos na capital paulistana tendem a aplicar com maior frequência a regra de CV do que os falantes procedentes de outras regiões (Noroeste de São Paulo, Norte do Paraná, Norte de Minas Gerais, Sul da Bahia, e região Nordeste). O fator é significativo tanto na 3ª, quanto na 1ª pessoa, mas tem peso maior em P1.

No *corpus* sob análise, o fator procedência se mostra relevante somente nas ocorrências com sujeitos de 3ª pessoa do plural. Os dados quantificados são apresentados na tabela 24:

Fatores	Frequência	Probabilidade
G – Minas Gerais	49/254 = 19%	.31
S – São Paulo	76/266 = 29%	.68
Total	125/520 = 24%	

Tabela 24: frequência e probabilidade de aplicação da regra de concordância verbal (P3) segundo a procedência do informante

Os resultados relativos à procedência dos informantes mostram que os falantes de São Paulo têm uma probabilidade maior de aplicar a concordância do que os falantes de Minas Gerais. Através do cruzamento dos fatores sociais, verificamos que a diferença mais expressiva entre um e outro grupo de informantes se dá entre os de escolaridade nula: enquanto os procedentes de São Paulo apresentam uma frequência de 26% de aplicação da concordância, os que residem em Minas Gerais empregam a regra somente em 7% dos casos⁷⁶. Tal padrão pode estar relacionado ao grau de interação dos informantes de São Paulo com falantes de outras variedades lingüísticas.

No quadro a seguir, comparamos a frequência geral de não-aplicação da concordância encontrada entre os informantes da zona rural cujos dados estudamos com a constatada por Rodrigues (1987) entre informantes de 3ª faixa etária residentes no meio urbano:

	Frequência de não-aplicação	
	Sujeito de 3ª pessoa do plural	Sujeito de 1ª pessoa do plural
Informantes da zona rural (<i>corpus</i> do Projeto Filologia Bandeirante)	395/520 = 76%	212/317 = 67%
Informantes da zona urbana (<i>corpus</i> do português popular de São Paulo)	338/519 = 65%	91/200 = 45%

Tabela 25: frequência geral de não-aplicação da regra de concordância verbal (P3 e P1) no *corpus* sob análise e entre os falantes de terceira faixa etária cujos dados integram o *corpus* do português popular de São Paulo (Rodrigues, 1987:218)

⁷⁶ Nas ocorrências de P1, também verificamos que os informantes analfabetos residentes em São Paulo apresentam uma frequência muito maior de aplicação da concordância (39%) do que os informantes analfabetos de Minas Gerais (3%).

Vemos que os falantes da zona rural empregam mais freqüentemente formas verbais não-marcadas do que os falantes da região urbana, o que confirma uma de nossas hipóteses. A diferença entre um e outro grupo de informantes se manifesta nas duas pessoas, mostrando-se mais significativa em P1 (22%) do que em P3 (11%).

Conforme destaca Rodrigues (1987), a tendência de não concordar o verbo com o sujeito é refreada pelo padrão de realização lingüística dos centros urbanos, principalmente quando o sujeito corresponde a 1ª pessoa do plural, contexto em que a não-realização da concordância seria mais saliente e mais estigmatizada. Já na zona rural, a concordância não-padrão parece não veicular um significado social negativo.

Tanto os falantes da zona urbana, como os falantes da zona rural, apresentam maior índice de não-concordância com sujeitos de 3ª pessoa do plural do que com sujeitos de 1ª pessoa do plural. A diferença entre um e outro contexto lingüístico é significativa na fala de informantes da região urbana (20%). Já entre os informantes da zona rural, não se mostra tão expressiva (9%). Isto igualmente confirma nossa hipótese de que os falantes rurais apresentariam “a mesma” freqüência de não-concordância na 3ª e na 1ª pessoas do plural. Cabe esclarecer que quando estabelecemos esta hipótese, não esperávamos encontrar índices exatamente iguais, mas sim valores próximos, como os que foram constatados.

CONSIDERAÇÕES FINAIS

Nesta pesquisa, buscamos analisar, de um ponto de vista sincrônico e quantitativo, a variação entre aplicação x não-aplicação da regra de concordância verbal de 3ª e 1ª pessoas do plural, na variedade do português popular falada por pessoas idosas residentes na zona rural dos estados de São Paulo e Minas Gerais, na área correspondente às trilhas das bandeiras paulistas.

O capítulo 1 compreendeu a apresentação dos pressupostos teórico-metodológicos adotados, baseados na Sociolinguística e na Linguística Funcional. Nele discorremos principalmente sobre a metodologia sociolinguística utilizada por Labov (1991, 1994) para o estudo de fenômenos linguísticos variáveis, e sobre a proposta de análise funcionalista de Halliday (1994, 1987, 1974), as quais serviram de base para o estudo da concordância verbal no português popular.

No capítulo 2, apresentamos uma resenha de estudos que tratam da concordância verbal no português do Brasil. Foram expostos os fatores linguísticos e sociais que condicionam a aplicação/não-aplicação da regra, bem como as hipóteses formuladas para explicar a origem da variação da concordância verbal.

No capítulo 3, descrevemos o *corpus* sob análise, apresentando informações sobre: o Projeto Filologia Bandeirante, as características sociais dos informantes, a variedade linguística popular, a coleta de dados, e as normas de transcrição.

O capítulo 4 abrangeu a descrição, análise e interpretação dos dados. Na primeira parte deste capítulo, expusemos os contextos de variação da concordância verbal encontrados no português popular, bem como os critérios para seleção de ocorrências. Na segunda, apresentamos os resultados da análise, os quais foram estabelecidos a partir dos cálculos fornecidos pelos programas computacionais do Varbrul.

Os resultados confirmaram todas as hipóteses formuladas na Introdução deste trabalho. A baixa frequência de realização da concordância verbal – 24%, na 3ª pessoa do plural, e 33%, na 1ª pessoa do plural – mostrou que os falantes do português popular registrado nas trilhas das bandeiras se inclinam realmente a não aplicar a regra padrão estabelecida pela gramática normativa, evidenciando ainda que casos de não-concordância tendem a ocorrer com praticamente a mesma frequência tanto na 3ª, como na 1ª pessoa do plural⁷⁷.

⁷⁷ Apesar do baixo índice de concordância constatado, os dados mostram que o falante popular tende a optar pelo sujeito fisicamente presente na oração. Em P3, 73% (380/520) das ocorrências correspondem a sujeitos expressos, e em P1, 67% (211/317).

Examinado o desempenho individual de cada falante pudemos observar que, com relação às ocorrências de 3ª pessoa do plural, a frequência de não-aplicação da regra de concordância superou a de aplicação da regra, na fala de todos os informantes, com apenas uma exceção. Na 1ª pessoa do plural, somente três informantes apresentaram maior frequência de realização do que de não-realização da regra⁷⁸. Isto aconteceu precisamente na fala de pessoas que tinham alguma atividade ou característica social que os diferenciavam do falante tipicamente rural e que podem ter levado à aquisição de formas verbais flexionadas. Entre estas, podemos citar como exemplo: hábitos de leitura, costume de assistir a jornais televisivos, curso primário completo, maior grau de interação com falantes de outras variedades lingüísticas através de trabalho fora da zona rural, etc.

Houve alguns inquéritos em que a não-concordância se mostrou categórica. Isto se deu na fala de dois informantes, na 3ª pessoa do plural, e de quatro informantes, na 1ª pessoa do plural. Tais pessoas são representantes legítimos do falante rural, pois são analfabetas, ou semi-escolarizadas, tendo trabalhado somente na roça, em atividades como criação de gado, cultivo de plantações, e transporte de mercadorias com carro de boi.

Ainda que tais resultados indiquem uma tendência de não-realização da regra padrão de concordância verbal no português popular falado por informantes rurais, a análise estatística dos dados demonstrou que existem determinados contextos que favorecem, ou inibem, decisivamente a sua aplicação. Isto igualmente confirma a hipótese estabelecida de que a variação entre aplicação x não-aplicação da regra de concordância verbal não é aleatória, mas sim condicionada tanto por fatores lingüísticos (semânticos, gramaticais, e discursivos), como por fatores de natureza social.

Retomamos a seguir os fatores lingüísticos selecionados como estatisticamente significativos no condicionamento da concordância de 3ª pessoa do plural:

1) Grau de saliência fônica da oposição entre as formas verbais do singular e do plural

Os dados confirmaram que há maior frequência e probabilidade de aplicação da regra de concordância, quando a diferença entre singular e plural é materialmente maior, ou mais saliente, envolvendo segmentos acentuados (classes L, E, F do 2º nível). Já nos verbos em que a oposição entre singular e plural é menor, realizando-se através de segmentos não-acentuados (classes R e V do 1º nível), a aplicação da concordância é decisivamente desfavorecida.

⁷⁸ Um destes informantes utilizou somente verbos flexionados de 1ª pessoa do plural. Este foi o único caso encontrado de aplicação categórica da regra.

Na classe W, a mais saliente de um ponto de vista fonológico, não se confirmou, contudo, a hipótese da saliência fônica: a frequência e a probabilidade de aplicação da regra se mostraram extremamente baixas. Comparando os resultados obtidos com os encontrados por Rodrigues (1987) entre falantes de 3ª faixa etária do meio urbano, constatamos que a diferença mais marcante entre informantes da região urbana e da zona rural se encontra justamente na classe W: enquanto os primeiros tendem a aplicar com alta frequência a regra de concordância quando a oposição entre singular e plural corresponde às formas *é* x *são*, no segundo grupo, a não-concordância é praticamente categórica. Hipotetizamos que este fato estaria relacionado à diferença de significado social atribuído à variante não-aplicação da regra: negativo no meio urbano, mas não no meio rural.

Assim como Rodrigues (1987), verificamos que a saliência fônica não se mostra atuante em instâncias de sujeito posposto ao verbo, contexto no qual a não-concordância é praticamente categórica.

2) Presença/ausência do sujeito pronominal (*eles, elas, vocês*)

Os resultados evidenciaram que a concordância é fortemente favorecida nos casos de sujeito exposto pelos pronomes *eles, elas, vocês* – os quais quase sempre figuram imediatamente antepostos ao verbo. O sujeito pronominal não-explicito constitui um outro contexto em que há grande probabilidade de realização da regra, ainda que a frequência de aplicação tenda a diferir conforme a localização do referente da casa vazia do sujeito e a função sintático-semântica do sujeito. O sujeito não-pronominal, por outro lado, condiciona decisivamente o uso de uma forma verbal não-marcada, independentemente da posição que ocupe em relação ao verbo. Vimos que estes resultados são semelhantes aos encontrados por Rodrigues (1987), diferindo apenas nos casos de sujeito pronominal explícito, contexto no qual a autora constata alta frequência de não-concordância. Assim como Rodrigues (1987), verificamos que a saliência fônica tem maior peso em instâncias de sujeito pronominal explícito e não-explicito do que nos casos de sujeito não-pronominal.

3) Paralelismo discursivo

Os dados confirmaram a hipótese do paralelismo discursivo: se o verbo é precedido de um verbo marcado, a concordância é fortemente favorecida; se, por outro lado, um verbo não-marcado precede a ocorrência, a regra de CV tende a não ser aplicada. A tendência de uma forma verbal marcada ser seguida por outra igualmente marcada é bem mais forte em instâncias de sujeito

pronominal explícito e não-explícito do que nos casos de sujeito não-pronominal. Quando antecedido por uma forma verbal no singular, o sujeito não-pronominal condiciona categoricamente a não-concordância. Questionamos, todavia, a posição de Labov (1994), segundo a qual o paralelismo seria de natureza contrafuncional. Através do cruzamento dos fatores paralelismo discursivo e status informacional do sujeito, verificamos que a utilização de uma sucessão de formas verbais não-flexionadas não compromete a informação, uma vez que em 82% das instâncias de verbo precedido de verbo não-flexionado, o referente constitui informação dada, já apresentada ao interlocutor; e nos 18% restantes de casos de verbo precedido de verbo sem marca, em que o referente corresponde a informação nova, o sujeito é explícito.

4) Traço semântico do sujeito

A análise dos dados evidenciou que a concordância é favorecida quando o sujeito é [+humano], ou simplesmente [+animado]. O sujeito [-humano], por outro lado, inibe decisivamente a aplicação da regra. A não-concordância se mostrou categórica nos casos de: (i) sujeito [-animado], e (ii) sujeito não-pronominal [-humano] – contexto que acumula dois fatores inibidores da CV. Hipotetizamos que os traços semânticos [+humano] e [+animado] tornam o sujeito mais saliente para o falante, fazendo com que a regra seja mais freqüentemente aplicada.

5) Paralelismo oracional

O exame inicial dos dados mostrou que a concordância é favorecida, quando o último elemento do SN sujeito carrega marca de plural, e desfavorecida, quando tal marca não está presente. Mediante o cruzamento dos fatores paralelismo oracional e presença/ausência do sujeito pronominal, verificamos, todavia, que esta formulação somente pode ser aplicada a instâncias de sujeito não-pronominal. Nestas, predomina a tendência funcional de marcar o plural apenas nos primeiros elementos do SN sujeito, enquanto os demais elementos que o compõem, bem como o verbo, não carregam marcas de flexão. Nos pouquíssimos casos em que o último elemento do sujeito não-pronominal comporta marca de plural, a freqüência de concordância tende a ser maior. Nas ocorrências de sujeito pronominal, por outro lado, não há como contrapor presença x ausência de marca, uma vez que só foram incluídos na análise pronomes pessoais no plural. Embora a freqüência de realização da regra seja alta neste contexto, acreditamos que isto se deva não à marca de plural no final do pronome, mas ao próprio pronome. Tal hipótese é reforçada pela comparação

entre a taxa de aplicação da concordância nos dois contextos em que a marca de plural no último elemento do SN sujeito está presente: sujeito pronominal (41%) x sujeito não-pronominal (21%).

6) Papel semântico do sujeito

A hipótese inicial estabelecida de que o sujeito Agente, sendo ativo, seria mais saliente para o falante, favorecendo o uso da concordância, e de que o sujeito Inativo, por oposição, inibiria a aplicação da regra, foi confirmada pelos índices de frequência, mas não pelos valores de probabilidade. Enquanto os primeiros colocam os sujeitos Beneficiário, Agente e Experimentador como papéis semânticos favorecedores da concordância, e os sujeitos Inativo, Paciente e Causativo como desfavorecedores; os últimos sugerem que os sujeitos Beneficiário e Inativo condicionariam a realização da regra, ao passo que os sujeitos Agente, Experimentador e Paciente a desfavoreceriam. Observando o traço semântico do sujeito [+/-animado], verificamos que os sujeitos Agente, Beneficiário e Experimentador são todos portadores do traço [+animado], o que explicaria a maior frequência de concordância encontrada nestes contextos. Nos casos de sujeito Inativo e Paciente, nos quais há variação entre referente [+animado] x [-animado], a frequência de realização da regra já é bem menor. O sujeito Causativo, sendo sempre [-animado], condiciona categoricamente a não-concordância. O traço [+/-humano], por sua vez, se mostra relevante nas instâncias de sujeito Experimentador, Inativo e Paciente, mas não nas de sujeito Agente, contexto no qual a frequência de realização da regra é praticamente igual, seja o sujeito [+humano] ou [-humano].

Na 1ª pessoa do plural, os fatores selecionados como numericamente relevantes no condicionamento da concordância foram os seguintes:

1) Acentuação da forma verbal

Os dados confirmaram que formas verbais paroxítonas são mais perceptíveis para o falante, condicionando a realização da regra de CV. Constatamos, da mesma forma que Rodrigues (1987), que quando a forma verbal é proparoxítona, a não-concordância se mostra praticamente categórica no português popular⁷⁹.

⁷⁹ Note-se que, tanto na 3ª pessoa, como na 1ª pessoa do plural, o fator selecionado como mais importante no condicionamento da concordância constitui um fator de natureza fonológica.

2) Paralelismo discursivo

O fator paralelismo discursivo se mostrou significativo também em ocorrências de 1ª pessoa do plural, confirmando a tendência de repetição das últimas formas verbais utilizadas, de modo que verbos marcados levam a verbos marcados; e verbos não-flexionados, a verbos não-flexionados. Tal princípio somente tem efeito, contudo, em verbos de acentuação paroxítona; já quando a forma verbal é proparoxítona, o paralelismo discursivo não se mostra atuante. Mediante o cruzamento deste fator com o status informacional do sujeito, questionamos novamente a formulação de Labov (1994) segundo a qual o paralelismo discursivo seria de natureza contrafuncional. O uso de uma cadeia de formas verbais não-flexionadas não constitui um impedimento para a apreensão do sujeito, o qual constitui, em grande parte das ocorrências, informação conhecida pelo interlocutor. A utilização de uma seqüência de verbos marcados, por sua vez, nem sempre veicula informação redundante, já que muitas vezes faz-se necessária a explicitação do referente do verbo novamente, seja através do pronome explícito ou da flexão, para evitar uma possível ambigüidade na interpretação do sujeito.

3) Status informacional do sujeito

Os dados confirmaram a hipótese funcional estabelecida, mostrando que há maior frequência e probabilidade de concordância, quando o sujeito constitui informação nova. O sujeito equivalente a informação dada, por sua vez, inibe decisivamente a realização da regra. Cruzando este fator com presença/ausência do sujeito pronominal, vimos que a maior parte dos casos de sujeito explícito corresponde ao pronome pessoal *nós*. Ainda que, neste contexto, a presença física do pronome identifique o referente do verbo, a frequência de realização da regra é um pouco maior nos casos de sujeito Novo do que nos de sujeito Dado. No que se refere às ocorrências de sujeito zero, verificamos que há uma forte tendência de não-realização da regra, quando este veicula informação dada. O sujeito zero correspondente a informação nova, por outro lado, condiciona decisivamente o uso de formas verbais flexionadas. Vimos, no entanto, que este último caso não pode ser tomado como argumento a favor da hipótese funcionalista, porque é possível que ocorrências totalmente não-marcadas não tenham sido inclusas na análise.

4) Distância do referente do sujeito zero

Os dados demonstraram a validade da hipótese funcionalista: o sujeito zero, cujo referente se encontra distante da ocorrência considerada, condiciona a aplicação da concordância, facilitando a ligação entre o verbo e o referente extra-sentencial. O sujeito zero com referente próximo à ocorrência, por outro lado, inibe decisivamente a realização da regra. Nos casos de sujeito zero sem referente explícito no discurso anterior, constatamos alta frequência de utilização de verbos flexionados. Não pudemos, todavia, fazer uma afirmação categórica sobre este último contexto, já que é possível que formas não-marcadas de 1ª pessoa do plural, com casa vazia do sujeito e verbo no singular, tenham passado despercebidas.

Com relação aos fatores sociais, pudemos constatar a relevância dos que seguem:

1) Escolaridade

O fator escolaridade se mostrou relevante para a concordância tanto de 1ª como de 3ª pessoa do plural, evidenciando que a frequência e a probabilidade de realização da regra são maiores entre falantes escolarizados do que entre falantes de escolaridade nula. Este padrão é condizente com o encontrado no estudo de Rodrigues (1987), o qual constata maior frequência de realização da regra de CV de 1ª pessoa do plural, entre falantes escolarizados. Tais resultados sugerem que a escola exerce um papel importante na aquisição de formas verbais flexionadas.

2) Sexo

O sexo do informante se mostrou significativo somente para a concordância de 3ª pessoa do plural. Os resultados relativos à análise deste fator demonstraram que, ao contrário do que havia sido constatado em dados de informantes urbanos estudados por Rodrigues (1987) e Bortoni-Ricardo (1985), são as mulheres da zona rural que apresentam maior frequência e probabilidade de realização da concordância do que os homens. Vimos, todavia, que quase todas as mulheres cujos dados estudamos são escolarizadas, fato que pode ter influenciado tal resultado. Cruzando os fatores sexo e escolaridade, comparamos a frequência de concordância encontrada entre falantes escolarizados de ambos os sexos, verificando que as mulheres superam os homens no emprego da regra. Embora esta diferença não seja numericamente significativa, sugerimos que a escola

exerceria maior influência sobre as mulheres do que sobre os homens, no que toca à realização da concordância.

3) Procedência

O fator procedência, atuante apenas na concordância de 3ª pessoa do plural, mostrou que os falantes de São Paulo apresentam maior frequência e probabilidade de realização da concordância do que os falantes de Minas Gerais. A comparação dos resultados da presente pesquisa com os encontrados por Rodrigues (1987), por sua vez, indicou que há de fato uma diferença nos padrões de concordância encontrados no português popular, conforme este registre uma variedade urbana ou uma variedade rural. Confirmamos a hipótese de que os falantes da zona rural, cujos dados estudamos, apresentam uma frequência maior de formas de não-concordância do que os falantes do português popular de São Paulo (Rodrigues, 1987), os quais habitam uma comunidade urbana. A diferença entre um e outro grupo de informantes se manifestou nas duas pessoas gramaticais, contudo, mostrou-se mais significativa na 1ª do que na 3ª pessoa do plural.

Esperamos que nossa pesquisa tenha contribuído para uma melhor compreensão das estruturas de variação da concordância verbal no português popular. Apesar do grande número de estudos existentes, o tema certamente não está esgotado. Os próprios dados estudados no presente trabalho sugerem algumas questões para estudos futuros, dentre as quais podemos citar como exemplo: a concordância com sujeito representado pelo pronome indefinido *tudo* retomando SN de 3ª pessoa do plural, ou substantivo coletivo no singular; a concordância em estruturas com o verbo *ser*, etc. Estudos posteriores podem também verificar em que medida os fatores selecionados como estatisticamente significativos em nossa pesquisa atuam em outros *corpora* representativos das variedades lingüísticas do português do Brasil.

REFERÊNCIAS BIBLIOGRÁFICAS

- ASSIS, R. M. (1988) Variações lingüísticas e suas implicações no ensino do vernáculo: uma abordagem sociolingüística. In: LIRA, S. e VANDRESEN, P. (eds.) *Ilha do Desterro* 20, p.59-81.
- BAXTER, A. N. & LUCCHESI, D. (1993) Processos de descrioulização no sistema verbal de um dialeto rural brasileiro. *Papia – Revista de Crioulos de base Ibérica*. Brasília, 2 (2): 59-71.
- BECHARA, E. (2001) *Moderna Gramática Portuguesa*. 37ª ed. Rio de Janeiro, Ed. Lucerna.
- BENVENISTE, E. (1976) “O homem na língua”. In: BENVENISTE, E. *Problemas de Lingüística Geral*. São Paulo, Ed. Nacional/EDUSP.
- BORBA, F. S. (org.) (1991) *Dicionário gramatical de verbos do Português Contemporâneo*. São Paulo, Ed. Unesp.
- BORTONI-RICARDO, S. M. (1981) A concordância verbal em português: um estudo de sua significação social. In: COUTO, H.H. (ed.) *Ensaio de Lingüística Aplicada*. Brasília, Thesaurus.
- _____ (1985) *The urbanization of rural dialect speakers: a Sociolinguistic study in Brazil*. Cambridge: Cambridge University Press.
- CAMACHO, R. G. (1993) Concordância verbal e verbos existenciais. *Estudos Lingüísticos XXII*. Anais de seminários do GEL. Vol 1. Ribeirão Preto:GEL/Inst. Moura Lacerda, p.631-638.
- CÂMARA JR., J. M. (1985) *História e Estrutura da Língua Portuguesa*. Rio de Janeiro, Padrão Livraria Editora.
- CASTILHO, A. T. (1997) “A gramaticalização”. In: *Estudos Lingüísticos e Literários [UFBA]* 19: março de 1997, p.25-63.
- _____ (no prelo) “Teoria gramatical: Formalismo/Funcionalismo. Postulados da Sintaxe Funcional”. In: CASTILHO, A.T. “Sintaxe Funcional da Língua Portuguesa”.
- CHAFE, W. L. (1979) *Significado e Estrutura Lingüística*. Rio de Janeiro, Livros Técnicos e Científicos.
- COHEN, M. A. A. (2000) Pressupostos Teórico-metodológicos do Projeto Filologia Bandeirante. In: MEGALE, H. (org.) (2000) *Filologia Bandeirante: Estudos 1*. São Paulo, Humanitas, FFLCH/USP, p.211-222.
- COHEN, M. A. A. M. et al. (1997) Filologia Bandeirante. In: *Filologia e Lingüística Portuguesa*, São Paulo, Humanitas, 1:79-94.
- CUNHA, C. & CINTRA, L. F. L. (1985) *Nova Gramática do Português Contemporâneo*. 2ª ed. Rio de Janeiro, Ed. Nova Fronteira.

- DANES, F. (1974) "Functional Sentence Perspective and the organization of the text". In: DANES, F. (ed.) *Papers on Functional Sentence Perspective*. Prague: Academic Publishing House of the Czechoslovak Academy of Sciences.
- DIK, S. C. (1997) "Towards a Functional Grammar of Discourse". In: DIK, S.C. *The Theory of Functional Grammar*. Dordrecht-Holland/ Providence RI: USA: Foris Publications.
- DUARTE, M. E. (1993) Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I. e KATO, M. (orgs.) *Português brasileiro: uma viagem diacrônica*. Campinas, Ed. da Unicamp, p.107-128.
- FERREIRA NETTO, W. (2001) *Introdução à Fonologia da Língua Portuguesa*. São Paulo, Hedra.
- FERREIRA NETTO, W. & RODRIGUES, A. C. S. (2001) Repensando a questão da transcrição de inquéritos orais. In: *Estudos Lingüísticos*. Anais de Seminários do GEL, Assis.
- GALVES, C. (1993) O enfraquecimento da concordância no português brasileiro. In: ROBERTS, I. e KATO, M. (orgs.) *Português brasileiro: uma viagem diacrônica*. Campinas, Ed. da Unicamp, p. 387-408.
- GUY, G. R. (1981) *Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax and language history*. University of Pennsylvania. PhD Thesis.
- HALLIDAY, M. A. K. (1974) "The place of 'Functional Sentence Perspective' in the system of linguistic description". In: DANES, F. (ed.) *Papers on Functional Sentence Perspective*. Prague: Academic-Publishing House of the Czechoslovak Academy of Sciences.
- _____ (1987) *Language as social semiotic – the social interpretation of language and meaning*. Great Britain, Edward Arnold.
- _____ (1994) *An Introduction to functional grammar*. Great Britain, Edward Arnold.
- HERRERAS, J. M. M. (1983) Presentación de Labov. In: LABOV, W. *Modelos Sociolingüísticos*. Madrid. Cátedra.
- ILARI, R (1992) *Perspectiva Funcional da Frase Portuguesa*. 21ª ed. Campinas, S.P., Editora da Unicamp.
- KOCH, I. G. V. (2000) Tematização e rematização no português falado no Brasil. In: GÄRTNER, E.; HUNDT, C.; SCHÖNBERGER, A. (eds.) *Estudos de lingüística textual do português*. Frankfurt am Main: TFM, 2000, p. 127-148.
- LABOV, W. (1991) *Sociolinguistic Patterns*. Philadelphia, University of Pennsylvania Press, 11ªed.
- _____ (1994) *Principles of Linguistic Change: Internal Factors*. Oxford –UK / Cambridge-USA, Blackwell Publishers, vol. 1.
- LAVANDERA, B. R. (1984) *Variación y Significado*. Buenos Aires, Hachette.

- LEMLE, M. & NARO, A. J. (1977) *Competências básicas do português*. Rio de Janeiro: Fundação Movimento Brasileiro de Alfabetização/Fundação Ford.
- MATTOS E SILVA, R. V. (1991) Caminhos de mudança sintático-semântica no português arcaico. *Cadernos de Estudos Lingüísticos* 20. Campinas, UNICAMP, p.59-74.
- _____ (2001) De fontes sócio-históricas para a história social lingüística do Brasil: em busca de indícios. In: MATTOS E SILVA, R. V. (org.) (2001) *Para a História do Português Brasileiro*. São Paulo, Humanitas, FFLCH/USP, vol. II, tomo II, p. 275-301.
- MELLO, H. R. (1996) *The genesis and development of Brazilian Vernacular Portuguese*. New York: City University of NY, tese de doutoramento.
- MENDES, S. T. P. (2000) Que língua falavam os bandeirantes na região de Minas Gerais? In: MEGALE, H. (org.) (2000) *Filologia Bandeirante: Estudos 1*. São Paulo, Humanitas, FFLCH/USP, p.65-75.
- MICHELETTI, H. (2000) *A indeterminação do sujeito: um estudo do pronome de terceira pessoa do plural*. São Paulo, USP, Dissertação de Mestrado.
- NARO, A. J. (1981) The social and structural dimensions of a syntactic change. *Language* 57 (1):63-98.
- NARO, A. J. & LEMLE, M. (1977) Syntactic Diffusion. *Ciência e Cultura* 29 (3):259-268.
- NARO, A.J. & SCHERRE, M. M. P. (1993) Sobre as origens do português popular do Brasil. *Delta*, 9, nº especial, p.437-454.
- _____ (1999) Variable concord in Portuguese: the situation in Brazil and Portugal. In: MCWHORTER, John (ed.) *Current issues in pidgin and creole linguistics*. Amsterdam, Benjamins.
- NEVES, M. H. M. (2001) *A Gramática Funcional*. São Paulo, Martins Fontes.
- OLIVEIRA, M. (2000) Para um programa de análise lingüística do português falado na trilha dos bandeirantes. In: MEGALE, H. (org.) (2000) *Filologia Bandeirante: Estudos 1* São Paulo, Humanitas, FFLCH/USP, p. 223-235.
- PEREIRA, D. C. & RODRIGUES, A. C. S. (2003) *Algumas observações sobre a concordância verbal na fala de idosos do Projeto Filologia Bandeirante*. Comunicação apresentada no GEL (Grupo de Estudos Lingüísticos do Estado de São Paulo), realizado em Taubaté, em maio de 2003.
- PEZATTI, E. G. (1992) *A ordem das palavras em português: aspectos tipológicos e funcionais*. Araraquara: UNESP. Tese de Doutorado.
- POSSENTI, S. (1979) Discurso: objeto da Lingüística. In: *Sobre o discurso*. Série Estudos 6. Uberaba: Faculdades Integradas.

- ROBERTS, I. (1993) O português brasileiro no contexto das línguas românicas. In: ROBERTS, I. e KATO, M. (orgs.) *Português brasileiro: uma viagem diacrônica*. Campinas, Ed. da Unicamp, p. 409-425.
- RODRIGUES, A. C. S. (1987) *A concordância verbal no português popular em São Paulo*. São Paulo: USP. Tese de Doutorado.
- _____ (1989) Concordância verbal e estratégias de pronominalização. In: *Estudos Lingüísticos*. Anais de Seminários do GEL. Lorena/Prefeitura Municipal/GEL, p.546-554.
- _____ (2000) Concordância verbal e saliência social no português popular em São Paulo. In: GARTNER, E., HUNDT, C. e SCHOMBERG (eds.) *Estudos de sociolingüística brasileira e portuguesa*. Frankfurt am Main: TFM (Biblioteca luso-brasileira, vol 15) p.41-61.
- RODRIGUES, A. C. S. & CAMPOS, O. G. L. A. S. (2002) *Reflexões sobre fatos de não-concordância verbal no português culto brasileiro*. São Paulo, Comunicação apresentada no Congresso da ALFAL, Costa Rica, 2002.
- RODRIGUES, A. C. S & FERREIRA NETTO, W. (2000) Transcrição de Inquéritos: problemas e sugestões. In: MEGALE, H. (org.) (2000) *Filologia Bandeirante: Estudos 1*. São Paulo, Humanitas, FFLCH/USP, p.171-193.
- SAID ALI, M. (1971) *Gramática Histórica da Língua Portuguesa*. Rio de Janeiro, Ed. Melhoramentos.
- SCHERRE, M. M. P. (1988) *Reanálise da concordância nominal em Português*. Rio de Janeiro: UFRJ: Tese de doutorado.
- _____ (1993) *Introdução ao pacote Varbrul para microcomputadores*. Rio de Janeiro: UFRJ.
- SCHERRE, M. M. P. & NARO, A. J. (1993) “Duas dimensões do paralelismo formal na concordância verbal no português popular do Brasil”. In: *Delta*, vol.9, n.1, 1993, p.1-14.