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

LÍVIA CAROLINA BAENAS BARIZON

REALIZAÇÃO/ NÃO-REALIZAÇÃO DOS PRONOMES SUBJETIVOS NO PORTUGUÊS POPULAR DA CIDADE DE SÃO PAULO: UM CASO DE VARIAÇÃO

VERSÃO CORRIGIDA
Exemplar original se encontra no CAPH da FFLCH (Centro de Apoio à Pesquisa Histórica)

SÃO PAULO
2012

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

LÍVIA CAROLINA BAENAS BARIZON

REALIZAÇÃO/ NÃO-REALIZAÇÃO DOS PRONOMES SUBJETIVOS NO PORTUGUÊS POPULAR DA CIDADE DE SÃO PAULO: UM CASO DE VARIAÇÃO

Dissertação apresentada ao Programa de Pós-Graduação em Filologia e Língua Portuguesa do Departamento de Letras Clássicas e Vernáculas 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

VERSÃO CORRIGIDA
Exemplar original se encontra no CAPH da FFLCH (Centro de Apoio à Pesquisa Histórica)

SÃO PAULO
2012

A minha mãe Jovita, pelo amor incondicional e por me ensinar a dar valor aos estudos.
Ao meu esposo Leonardo, pela compreensão, carinho e amor nos momentos difíceis.
A Jasmin, com carinho.

AGRADECIMENTOS

Agradeço primeiramente a Deus, fonte de vida e luz, que conduz meus passos e guia a minha vida, nunca me deixando só.

Agradeço a minha mãe Jovita, em primeiro lugar, por ter me dado a vida, em segundo lugar, por ter me criado com tanta dedicação e zelo em minha formação enquanto ser humano, me transmitindo seus valores o que resume ao que sou hoje. E em terceiro lugar, por nunca ter desistido de mim.

Agradeço ao meu marido Leonardo, presente de Deus que sempre está ao meu lado, sendo companheiro, cuidando e acreditando em mim.

Agradeço ao meu padrasto Gú, pela amizade, pelo carinho, pela força e pelas palavras necessárias de perseverança.

Agradeço ao meu irmão Vinicius e minha cunhada Carolina, pelos sorrisos e carinho.

Agradeço ao meu sogro Barizon e minha sogra Cássia pelas orações, pelo apoio e amizade.

Agradeço a minha orientadora Dra. Angela C. de S. Rodrigues, pelo auxílio, dedicação, incentivo e amizade, por sempre estar disposta a me ouvir, pela sua experiência e sabedoria, me ensinado o verdadeiro valor não somente da pesquisa, como também das coisas da vida.

Agradeço a professora Dra. Deize Crespim Pereira, por todo incentivo, pelas leituras e discussões realizadas, por sua amizade, sabedoria e prestatividade.

Agradeço a professora Dra. Léa Silvia Braga de Castro Sá que foi a minha orientadora durante a graduação e responsável pelos meus primeiros passos trilhados no caminho do conhecimento.

Agradeço aos amigos que conquistei durante o período da minha vida que estive em São Paulo: Giovana, Iva, Raquel, Neusa, Cristina, Claudia,...

Agradeço a todas as pessoas que de forma direta ou indireta contribuíram para a realização desta pesquisa.

“Peçam, e lhes será dado; busquem, e encontrarão; batam, e a porta lhes será aberta. Pois todo o que pede, recebe; o que busca, encontra; e àquele que bate, a porta será aberta.”

Mateus 7:7-8

RESUMO

Esta pesquisa tem como objetivo descrever, em sincronia, os padrões de variação de preenchimento do sujeito por pronomes pessoais, encontrados na variedade popular do Português falado na cidade de São Paulo, especificamente, das favelas da zona Norte (Jardins Carombé e Paulistano). O *corpus* de análise é constituído por 15 inquéritos de 30 minutos cada, gravados na segunda metade da década de 1980, estratificados por sexo, escolaridade e idade do informante. Buscou-se, sobretudo, explicar a variável realização/ não-realização do pronome subjetivo, levando-se em conta os pressupostos teórico-metodológicos da Sociolinguística Variacionista de William Labov e da Linguística Funcional de Michael Halliday, no intuito de identificarmos os fatores linguísticos e sociais que estão condicionando a variável dependente. Os resultados obtidos, através do estudo em tempo aparente, não nos possibilitam afirmar que está ocorrendo mudança linguística em progresso, ou seja, que a língua esteja caminhando em direção às línguas "não-pro-drop". O Status Informacional demonstrou ser o fator mais significativo, prevalecendo sobre os demais fatores e que condiciona a realização/ não- realização dos pronomes subjetivos.

Palavras-chave: Pronomes Subjetivos. Variação Linguística. Português Popular. Sociolinguística. Linguística Funcional.

ABSTRACT

This research aims to describe, in synchrony, the variation patterns of filling of the subject by personal pronouns found in the popular variety of Portuguese spoken in São Paulo, specifically, the slums of the North (Jardins Carombé and Paulistano). The *corpus* of analysis consists of 15 surveys of 30 minutes each, recorded in the second half of the 1980s, stratified by sex, age and education of the informant. It´s above all explain the variable occurrence of the subjective pronoun, taking into account the theoretical and methodological assumptions of Sociolinguistics Variationist by William Labov and Functional Linguistics by Michael Halliday, in order to identify linguistic and social factors conditioning that are the dependent variable. The results obtained, through the apparent time study, do not allow us to assert that linguistic change in progress is occurring, ie, that language is moving in a direction to the "non-pro-drop" languages. The Information Unit proved to be the most significant factor, prevailing over other factors, that is conditioning the occurrence of subjective pronouns.

.

Key Words: Subjective Pronouns. Linguistic Variation. Popular Portuguese. Functional Linguistics. Sociolinguistics.

LISTA DE QUADROS

Quadro 1- Formalismo e Funcionalismo..... 16
Quadro 2 - Gramática Formal x Gramática Funcional..... 16
Quadro 3 - Metafunções..... 23
Quadro 4 - Informantes discriminados conforme o fator Idade 48
Quadro 5 - Fatores Linguísticos e Social 55
Quadro 6 - Categorias referentes ao fator Pronome Pessoal Subjetivo 56
Quadro 7- Categorias referentes ao fator Forma Verbal 58
Quadro 8 - Categorias referentes ao fator Polaridade 61
Quadro 9 - Categorias do fator Indeterminação do Referente 69
Quadro 10- Fatores linguísticos selecionados e descartados pelo programa estatístico Goldvarb 2001 76
Quadro 11 - Análises do Step Up e Step Down referentes a três rodadas: com todos os fatores; sem o fator Idade; sem o fator Forma Verbal, respectivamente..... 85

LISTA DE GRÁFICOS

Gráfico 1 - Ocorrência de sujeito nulo na fala espontânea segundo a pessoa gramatical e a faixa etária (%)..... 42
Gráfico 2 - Resultado geral de realização/ não-realização dos pronomes subjetivos 75
Gráfico 3 - Frequência de realização dos pronomes subjetivos conforme a pessoa discursiva 81
Gráfico 4 - Frequência de realização/ não-realização dos pronomes subjetivos, de acordo com as faixas etárias dos informantes 87
Gráfico 5 - Frequência de realização dos pronomes subjetivos por indivíduo..... 92

LISTA DE TABELAS

Tabela 1 - Correlação de frequência relativa de ocorrência das pessoas do discurso e frequência de omissão do pronome..... 30
Tabela 2 - Evolução nos paradigmas flexionais do Português 31
Tabela 3 - Retenção vs. Apagamento pronominal em várias categorias sintáticas: corpus de 1981 (TARALLO, 1991)..... 33
Tabela 4 - Retenção pronominal segundo função sintática e período de tempo (dados de 1725 a 1880) (TARALLO, 1991)..... 33
Tabela 5 - Resumo da porcentagem de retenção pronominal..... 34
Tabela 6 - Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo de realização de acordo com o Status Informacional do Sujeito..... 76
Tabela 7 - Frequência de realização do pronome de 1ª pessoa do plural, de acordo com o Status Informacional do Sujeito, em nossa amostra e no *corpus* de Pereira (2004) 77
Tabela 8 - Frequência de não-realização dos pronome subjetivo de primeira pessoa do plural, conforme o Status Informacional do Sujeito, em nosso *corpus* e no *corpus* de Pereira (2004)..... 78
Tabela 9 - Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo de realização, conforme o Paralelismo Formal..... 79
Tabela 10 - Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo de realização, conforme o fator Pronome Pessoal..... 80
Tabela 11 - Frequência do pronome subjetivo, resultante do cruzamento dos fatores Pronome Pessoal e Status Informacional do Sujeito 82
Tabela 12 - Comparação da frequência de não-realização dos Pronomes de nosso *corpus* e do *corpus* de Silva (1993)..... 82
Tabela 13 - Frequência de realização/ não-realização do pronome subjetivo e peso relativo de realização, conforme o fator Polaridade..... 83
Tabela 14 - Frequência de realização/ não-realização do pronome de primeira pessoa do singular, conforme a Polaridade, no *corpus* sob análise e no *corpus* de Neves (2006) 84
Tabela 15 - Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo de realização, conforme a Faixa Etária..... 86
Tabela 16 – Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo de realização, conforme a Forma Verbal 88

Tabela 17 - Frequência de realização dos pronomes de acordo com o cruzamento dos fatores Forma verbal e Status informacional do sujeito 89
Tabela 18 - Frequência de realização/ não-realização dos pronomes subjetivos conforme a Indeterminação do Referente 89
Tabela 19 - Frequência de realização/ não realização dos pronomes subjetivos, conforme o Valor Semântico do Verbo 90
Tabela 20 - Frequência da realização dos pronomes subjetivo, considerando-se o cruzamento dos fatores Valor Semântico do verbo e Status Informacional 91
Tabela 21 - Frequência de realização/ não-realização dos pronomes subjetivos, conforme o sexo dos informantes 92

SUMÁRIO

INTRODUÇÃO **12**

CAPÍTULO 1: Pressupostos teórico-metodológicos **14**

1.1 A Sociolinguística Variacionista 17

1.2 A Linguística Funcional 20

CAPÍTULO 2: Expressão dos pronomes subjetivos na literatura linguística **24**

2.1 Fatores linguísticos 25

2.2 Fatores sociais 41

CAPÍTULO 3: O *corpus* sob análise **43**

3.1 O Português Popular em São Paulo 43

3.1.1 Características do Português Popular 45

3.1.2 Perfil Social dos Informantes 47

3.1.3 Normas de transcrição 48

3.1.3.1 Normas de transcrição convencional do Projeto Caipira 48

3.2 Coleta de Dados 51

CAPÍTULO 4: Seleção e Caracterização dos Dados **52**

4.1 Critérios de seleção de dados 52

4.2 Critérios de exclusão de dados 53

4.3 Realização/ não-realização dos pronomes subjetivos: a variável 54

4.3.1 Variáveis Independentes 55

4.3.1.1 PRONOME PESSOAL SUBJETIVO 56

4.3.1.2 FORMA VERBAL E AMBIGUIDADE MORFOLÓGICA 56

4.3.1.3 PARALELISMO FORMAL 59

4.3.1.4 POLARIDADE 61

4.3.1.5 STATUS INFORMACIONAL DO SUJEITO 61

4.3.1.6 INDETERMINAÇÃO DO REFERENTE 69

4.3.1.7 VALOR SEMÂNTICO DO VERBO 71

4.3.1.8 IDADE DO INFORMANTE 73

CAPÍTULO 5: Apresentação e discussão dos resultados 74

5.1 Notas Introdutórias 74

5.2 Resultado Geral 75

5.3 Fatores selecionados e descartados..... 75

5.3.1 Status Informacional do Sujeito 76

5.3.2 Paralelismo Formal 78

5.3.3 Pronome Pessoal Subjetivo / Pessoa discursiva 80

5.3.4 Polaridade 83

5.3.5 Idade do Informante 85

5.3.6 Forma Verbal 88

5.3.7 Indeterminação do Referente 89

5.3.8 Valor Semântico do verbo 90

5.3.9 Análise da Realização dos pronomes por indivíduo 91

Considerações Finais..... 94

REFERÊNCIAS..... 100

INTRODUÇÃO

O objetivo desta pesquisa é descrever quais são os padrões de variação de preenchimento do sujeito por pronome, em todas as pessoas do singular e plural, encontrados na variedade do português popular falada na cidade de São Paulo, especificamente, das favelas da zona Norte (Jardins Carombé e Paulistano). Trata-se de um estudo de natureza sincrônica e quantitativa que busca, sobretudo, analisar e explicar a variável realização/ não-realização do pronome subjetivo, tendo como aporte os pressupostos teórico-metodológicos da Sociolinguística Variacionista e da Linguística Funcional de Michael Halliday (1994).

As seguintes hipóteses orientam a presente pesquisa:

(1) a variação entre realização/ não-realização do pronome subjetivo não seria aleatória, mas condicionada por fatores internos, linguísticos e por fatores extralinguísticos, de natureza social;

(2) "A compensação funcional" estaria atuando no *corpus*, ou seja, o sujeito explícito tenderia a ocorrer mais em contextos ambíguos, quando este vincula informação nova no discurso, e menos quando sua realização não se fizer necessária para a compreensão da mensagem, ou seja, quando a informação é dada;

(3) fatores de natureza antifuncional, como o paralelismo formal, poderiam influenciar a variável realização/ não-realização do pronome sujeito;

(4) o pronome sujeito explícito ocorreria com maior frequência quando houvesse a alternância de referente (mudança de tópico discursivo), isto é com referente Novo; por outro lado o pronome sujeito ocorreria com menos frequência quando o referente fosse mantido, em que é Dado.

(5) o falante do português popular estaria utilizando mais o pronome do que zero, sendo possível perceber que os idosos utilizariam mais o zero e os mais novos utilizariam mais o pronome. Assim, o Português do Brasil estaria deixando de ser uma língua ***pro-drop*** passando para ***não-pro-drop.***

(6) a polaridade poderia estar atuando no *corpus*: a sentença afirmativa favoreceria o uso do pronome pessoal explícito, enquanto a negativa favoreceria o uso de zero (Ø);

(7) as formas verbais marcadas em pessoa e número (como: 1ª pessoa do singular), por não apresentarem ambiguidade morfológica, favoreceriam a omissão do pronome sujeito; as formas não marcadas em pessoa e número favoreceriam a presença do pronome;

(8) haveria maior frequência de preenchimento do pronome subjetivo quando o referente fosse indeterminado;

(9) poderia haver relação entre o preenchimento do sujeito e o tipo de verbo (de acordo com suas propriedades sintático-semânticas).

O presente trabalho está dividido em cinco capítulos. No primeiro, são apresentados os pressupostos teórico-metodológicos utilizados, que compreendem os da Sociolinguística Laboviana (1972) e Linguística Funcional de Halliday (1994).

O segundo capítulo corresponde a uma resenha detalhada dos trabalhos elaborados anteriormente sobre o preenchimento do sujeito por pronome no Português do Brasil, sendo ressaltados os fatores considerados relevantes para elaboração desta pesquisa, ou seja, os fatores linguísticos e sociais que condicionam a realização X não-realização do pronome subjetivo, bem como a formulação das hipóteses que expliquem a variável.

No terceiro capítulo, são apresentados esclarecimentos sobre o *corpus* e as características sociais dos informantes.

No quarto capítulo, discorremos sobre os critérios de seleção e exclusão de dados, apresentação e explicação da variável dependente e das variáveis independentes (grupos de fatores utilizados).

O quinto capítulo versa sobre a análise e interpretação dos dados e discussão dos resultados fornecidos pelo programa computacional Goldvarb 2001. Assim, discorremos neste capítulo a respeito dos fatores considerados condicionantes da realização dos pronomes subjetivos.

Encerramos nossa pesquisa ressaltando de forma breve os resultados obtidos e sugerindo propostas para elaboração de pesquisas futuras.

Pretendemos, com esta pesquisa, contribuir, de algum modo, para uma melhor caracterização e compreensão da variação da expressão dos pronomes subjetivos da Língua Portuguesa, possibilitando a elaboração de um quadro descritivo mais completo do Português popular falado na cidade de São Paulo, na medida em que são poucos os trabalhos que contemplam essa variedade.

Acreditamos que o uso do pronome sujeito não é aleatório, a sua omissão não é categórica, nem sua explicitação, mas podem ser motivadas por fatores linguísticos, ou seja, prosódicos, sintáticos, semânticos, discursivos e pragmáticos e por fatores sociais.

Para comprovar as hipóteses explicitadas anteriormente, foi necessário realizar um estudo detalhado e uma análise estatística das ocorrências de realização / não-realização dos pronomes subjetivos de todas as pessoas do singular e plural presentes no *corpus*.

CAPÍTULO 1: Pressupostos teórico-metodológicos

As bases teórico-metodológicas deste trabalho são fornecidas pela Sociolinguística Variacionista Laboviana e pela Linguística Funcional de Michael A. K. Halliday (1994).

A Sociolinguística Variacionista, iniciada por William Labov (1972), também conhecida como "Teoria da Variação Linguística", estuda a língua em uso, em situações reais de comunicação, dentro de uma comunidade de fala específica, levando em conta a influência de fatores externos à língua (sociais) e fatores internos (linguísticos) para o estudo da variação e mudança linguísticas, além de operar com programas estatísticos para quantificar os dados coletados.

Sendo o estudo variacionista proposto por Labov de natureza quantitativa, foi necessária a utilização do programa computacional GoldVarb 2001 para o tratamento estatístico dos dados. Por meio desse programa, foi possível efetuar cálculos de frequência e peso relativo da variável realização/ não-realização do pronome subjetivo, cruzar fatores e analisar o peso relativo de cada fator condicionante. Desse modo, torna-se viável a identificação dos fatores internos e externos que estão favorecendo ou inibindo a variação, e estabelecer correlações entre comportamento linguístico e estrutura social.

Também foram empregados os pressupostos da Linguística Funcionalista, utilizando especificamente o funcionalismo de Halliday (1994), para a definição das hipóteses e análise dos resultados. Acreditamos que a Teoria Funcionalista de Michael Halliday se relaciona perfeitamente com a Sociolinguística Variacionista, pois ambas se propõem a analisar o uso efetivo da linguagem em seu contexto social e dinâmico: *"It is functional in the sense that it is designed to account for how the language is used."* (HALLIDAY, 1994: xiii).

As duas teorias também apontam para a importância de se estudar a língua falada cotidiana, em situações reais de uso. Segundo LABOV: "a sociedade é o contexto da língua, só na primeira adquire a segunda significação" (1983:18). Halliday (1987:187), por sua vez destaca*: "if we seek to explain the internal workings of language we are forced to take into consideration its external relation to the social context".*

As duas teorias integradas se enquadram numa análise linguística que podemos chamar de funcionalista em contraposição à análise formal. De acordo com Castilho (2010: 64), há uma regularidade na identificação das duas principais correntes teóricas da linguística contemporânea: o formalismo e o funcionalismo.

Para Hymes e Dik (1974/ 1978/1981: 5, 1989, apud CASTILHO, 2010: 64), o *formalismo* vê a língua como um conjunto de orações, em que um falante utiliza sua

competência, ou seja, sua capacidade de produzir, interpretar e julgar sentenças consideradas gramaticais. Desse modo, as orações são descritas independentemente do seu contexto, e neste caso a sintaxe é autônoma com relação à semântica e à pragmática. Os dados da língua interna são analisados distintivamente dos dados da língua externa. Já para o *funcionalismo*, a língua é vista como um instrumento de interação social, que o falante utiliza para estabelecer comunicação e aguçar sua competência comunicativa, que é a capacidade de o falante saber utilizar uma variedade de língua adequada ao contexto, em outras palavras, o falante ser capaz de se comunicar, por meio da linguagem, em diferentes situações de interação. Assim, a descrição da língua é feita levando-se em conta as funções que as expressões linguísticas adquirem em determinados contextos, tendo como foco a pragmática, dentro da qual se estuda a semântica e a sintaxe.

A sintaxe formal compreende a sintaxe estrutural e a sintaxe gerativa, estas duas últimas não consideram as estruturas usadas em diferentes situações sociais. A sintaxe funcional estuda a língua contextualizada em uma determinada situação de interação, analisando o modo como a língua se gramaticaliza, como as categorias gramaticais e cognitivas são representadas (CASTILHO, 2010: 65).

Para muitos autores esses modelos teóricos não são vistos como divergentes, mas sim como complementadores; o que os diferencia é a estratégia de abordagem do fenômeno linguístico em cada um dos sistemas: gramática; léxico; semântica e discurso (CASTILHO, 2010:64).

No quadro abaixo, temos as principais características de cada modelo:

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 linguístico deve ter lugar no interior do sistema de usos linguísticos.
As sentenças de uma língua devem ser descritas independentemente do contexto em que ocorreram.	A descrição dos elementos linguísticos 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	A criança descobre o sistema que subjaz à língua e ao uso linguístico ajudada por

estímulo é pobre.	*inputs* de dados linguísticos extensos e altamente estruturados, presentes em contextos naturais.
Os universais linguísticos são propriedades inatas do organismo biológico e psicológico dos homens.	Os universais linguísticos 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.

Quadro1: Formalismo e Funcionalismo segundo Dik (1978/ 1981:4, In: CASTILHO, 2010: 66)

GRAMÁTICA FORMAL	**GRAMÁTICA FUNCIONAL**
Orientação sintagmática (“*chain grammar*”).	Orientação paradigmática (“*choice grammar*”).
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 linguísticos.	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 segundo Halliday (1994: xxviii, In: PEREIRA, 2004)

A utilização das teorias Sociolinguística e Funcionalismo, explicitadas acima, é altamente justificável, pelas seguintes razões: (a) o material analisado corresponde ao português popular falado em situações reais de comunicação, podendo este ser descrito com base nos parâmetros de ambas as teorias; (b) a possibilidade de se descrever e se analisar, por meio da metodologia da Sociolinguística, a variável realização/não-realização dos pronomes subjetivos, buscando resultados e explicando-os utilizando como aporte teórico a Linguística Funcional.

A seguir, discorremos a respeito dos pressupostos teórico-metodológicos da Sociolinguística Variacionista e da Gramática Funcional Sistêmica de Halliday.

1.1 A Sociolinguística Variacionista

A teoria da Sociolinguística Variacionista surge em 1964, em um congresso na Universidade da Califórnia em Los Angeles, onde se reuniram diversos pesquisadores da área da linguagem, entre eles, William Labov, que constataram a possibilidade de haver covariações entre sistemas linguístico e social de uma dada comunidade linguística.

Por meio da publicação de *"As dimensões da Sociolinguística",* texto introdutório de William Bright referente aos trabalhos do congresso, em 1966, o termo Sociolinguística é instituído e essa nova área da linguagem adquire seu objeto de estudo: a língua falada em seu contexto social.

Considerar a língua falada como ponto de partida para estudos linguísticos promove uma mudança dos parâmetros de análise linguística até então estipulados, pois essa nova vertente aflora as discussões sobre a oposição entre imanência da língua e contexto social.

Perceber que a língua é um sistema heterogêneo e que varia se tornou o principal postulado da Sociolinguística, sendo este termo considerado como equivocado por Labov, que afirma ser impossível desvincular a língua da sociedade. *"Por vários anos resisti ao termo sociolinguística, já que ele implica que pode haver uma teoria ou prática linguística bem sucedida que não é social."* (LABOV, 2008: 13).

A Sociolinguística Variacionista, também conhecida como "Teoria da Variação Linguística", caracteriza-se por ser um modelo teórico-metodológico que estuda a língua em uso, em situações reais de comunicação, a língua falada no cotidiano, daí o seu caráter empirista, ou seja, Labov, ao propor um modelo teórico, quis provar que se pode partir da observação dos dados da língua falada para se chegar à análise teórica.

> Para lidar com a língua, temos de olhar para os dados da fala cotidiana o mais perto e diretamente possível, e caracterizar seu relacionamento com as teorias gramaticais do modo mais acurado que pudermos, corrigindo e adequando a teoria para que ela se ajuste ao objeto visado (LABOV, 2008:236).

Para desenvolver estudos de variação dentro da teoria linguística, Labov postula uma série de conceitos teórico-metodológicos que servem de norteadores e são de extrema relevância para o desenvolvimento desta pesquisa. Segundo o autor, toda gramática de uma língua é constituída de *regras categóricas* e *regras variáveis*.

Define-se *regra categórica* como sendo as regras internas e *invariantes* das línguas, ou seja, as regras que não podem ser infringidas, pois resultaria em uma incompreensão dos

enunciados. Como exemplo disso, temos o uso do artigo em português que sempre antecede o substantivo. Vale ressaltar que este é justamente o ponto em comum entre a teoria gerativista e a sociolinguística, ou seja, ambas as teorias assumem o conceito de *regra categórica*.

Porém, além das *regras categorias* ou *invariantes*, existem as *regras variáveis*, que se aplicam quando estamos diante de duas ou mais formas distintas de se dizer a mesma coisa, com o mesmo valor de verdade, mesmo conteúdo informativo. A essas diversas maneiras de se dizer a mesma coisa, em um mesmo contexto, dá-se o nome de *variantes linguísticas*. O conjunto de variantes linguísticas constitui, portanto, a variável linguística. Desse modo, o preenchimento do sujeito por pronome trata-se de uma *regra variável*, em que duas variantes estão em competição: **realização** X **não-realização** do pronome subjetivo.

Para a Sociolinguística, a variação não é aleatória, mas condicionada por fatores internos à língua, conhecidos também por fatores linguísticos e fatores externos à língua, ou seja, fatores sociais. Como exemplo de fatores sociais, podemos citar: a história da comunidade ou da vida do falante, a classe social a qual pertence, o sexo, nível de escolaridade, idade do falante, etc. A esse grupo de fatores condicionantes (internos e externos à língua) dá-se o nome de *variáveis independentes*. A regra variável, ou seja, o fenômeno linguístico analisado, é denominada *variável dependente*.

A Sociolinguística Laboviana possui uma metodologia específica, de natureza quantitativa, em que os dados são selecionados, codificados e quantificados. Desse modo é possível, por meio do programa computacional GoldVarb 2001, desenvolver índices de frequência e peso relativo de aplicação de uma determinada variável, identificando os fatores condicionantes mais significativos e os fatores irrelevantes, sendo possível ainda estabelecer cruzamento de fatores (sociais e linguísticos). Desse modo, com a utilização de dados concretos, torna-se possível ao linguista identificar os fatores sociais e linguísticos que estão influenciando a variável analisada e estabelecer correlações entre comportamento linguístico e social.

O ponto central da Sociolinguística, como já foi dito anteriormente, é o modo de ver a língua como um sistema heterogêneo e dinâmico, que está em constante variação. Porém, essa variação muitas vezes é interpretada como um "caos", daí a necessidade de estabelecer mecanismos que demonstrem haver regularidade e ordem por de trás da heterogeneidade.

Mas se a língua varia e essa variação promove mudanças na língua, que acontecem enquanto um falante a utiliza, a grande questão formulada por Weinreich, Labov e Herzog (2006: 35) é: "*como é que as pessoas continuam a falar enquanto a língua muda, isto é,*

enquanto passa por períodos de menor sistematicidade?" E como um linguista pode analisar e comprovar essa mudança linguística em curso, durante o processo?

Para responder à segunda questão, Labov postula duas formas de analisar a língua: o estudo *em tempo real* (diacrônico) e em *tempo aparente* (sincrônico), ambos com a finalidade de analisar se está ocorrendo mudança linguística em progresso. A presente pesquisa desenvolve uma análise em *tempo aparente*, de modo que se torna possível detectar se está ocorrendo mudança dos pronomes subjetivos.

> O problema da transição, ou seja, analisar os estágios da mudança linguística, aqui proposto dependerá de análise mais acurada da distribuição de formas linguísticas no *tempo aparente* _ isto é, ao longo das dimensões formadas por faixas etárias da população atual. Tal análise é possível somente porque a descrição simples original da mudança em *tempo real* nos permite distinguir entre a gradação etária da população atual e os efeitos da mudança linguística (LABOV, 2008: 194).

O autor propõe a junção das formas de análise sincrônica e diacrônica a partir do *Princípio da Uniformidade*, ou seja, as mesmas forças que atuaram nas mudanças linguísticas do passado podem ser percebidas no presente.

Para o autor, os princípios empíricos da mudança linguística estão fundamentados em cinco problemas: (1) *fatores condicionantes, restritivos*– verificar o conjunto de mudanças possíveis e condições possíveis para a mudança; (2) *transição* – analisar os estágios da mudança, percurso através do qual uma mudança se realiza; (3) *encaixamento* - como a mudança linguística está encaixada no sistema linguístico e na estrutura social; (4) *avaliação* – entender como as mudanças são avaliadas pelos membros de uma comunidade de fala e quais significados estão sendo veiculados; (5) *implementação* – analisar os motivos pelos quais uma dada mudança ocorre em um determinado momento e em uma determinada comunidade de fala e não ocorre em outra. (WEINREICH, LABOV & HERZOG, 2006: 121-125).

A mudança linguística começa dentro de uma comunidade de fala, toma uma direção e assume o caráter de uma diferenciação ordenada. Nem toda variação implica mudança, mas toda mudança implica variação. Vale ressaltar que os fatores linguísticos e sociais estão intimamente inter-relacionados no desenvolvimento da mudança linguística e devem ser considerados para desenvolver um estudo empírico satisfatório da língua. (WEINREICH, LABOV & HERZOG, 2006: 126).

É muito difícil explicar as mudanças por forças internas apenas, é preciso considerar os fatores sociais. Analisando a língua em situações reais de uso é que podemos entender e tentar desvendar a complexidade da variação e mudança linguísticas.

Essa perspectiva de se analisar a língua viva, falada no cotidiano e principalmente a língua estigmatizada por falantes do português padrão tornou-se o foco central desta pesquisa, pois os falantes do português que se encontram na base da escala social também inovam e propagam mudanças na língua.

1.2 A Linguística Funcional

A língua em uso, em situações reais de interação, constitui o objeto de estudo da Linguística Funcional, tendo como foco a análise de como as pessoas efetivamente se comunicam por meio da língua, ou seja, como *"as estruturas semânticas e gramaticais funcionam no ato de comunicação"* (DANES In TLP1, In: HALLIDAY, 1974: 44), considerando tanto o uso quanto o sistema linguístico.

Os termos função e funcional são muito utilizados pela teoria funcionalista, porém vários autores atribuem diferentes significados a esses termos. Conforme Danes (1987:4 apud NEVES, 1994:112), a maioria dos autores da escola linguística de Praga usaram o termo função para designar "tarefas" que a linguagem ou seus componentes desempenham, ou propósito ao qual eles servem.

Para nossa pesquisa utilizamos especificamente o funcionalismo de Michael Halliday (1994), denominado pelo próprio autor como uma Teoria Funcionalista Sistêmica, baseada na teoria de Firth (Robins, 1964: 290, apud NEVES, 1994:116). É sistêmica, pois vê a língua como uma rede sistêmica que compõe significados, sendo que seu usuário faz escolhas.

O funcionalismo de Halliday é moderado, pois além de apontar a inadequação do formalismo, também analisa a estrutura, propondo uma interpretação linguística diferente da proposta formalista, já que interpreta o uso linguístico correlacionado ao sistema linguístico. De acordo com as afirmações de Halliday (1978: 40) "*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'".* Desse modo, o autor propõe analisar a língua partindo do uso para o conceito.

Seu modelo de estudo sistêmico-funcional parte da premissa de que o estudo da língua deve ser baseado nas relações paradigmáticas. Conforme Neves (1994: 117), "uma gramática

sistêmica é, acima de tudo, paradigmática, isto é, coloca nas unidades sintagmáticas apenas a realização, reservando, para o nível abstrato e profundo, as relações paradigmáticas".

Em sua obra *An Introduction to Functional Grammar*, Halliday apresenta uma visão geral de sua gramática funcional, esclarecendo os termos utilizados para definir os níveis ou estratos de língua, que são: Semântica, Gramática e Fonologia. O autor utiliza o termo Sintaxe como sendo oposto a Semântica, em que a Pragmática pode atuar como um terceiro termo. Para o autor, a Sintaxe é apenas uma parte da Gramática, na medida em que esta é constituída por uma Sintaxe, um Léxico e uma Morfologia.

Halliday (1994: xiv) afirma que o termo Sintaxe não é adequado, pois este sugere uma direção particular, tal qual a língua é interpretada como um sistema da forma, no qual os significados são designados. Na história da linguística ocidental, desde os primórdios na Grécia Antiga, primeiramente a forma da palavra era estudada (Morfologia); então, para explicar a forma das palavras, os gramáticos começaram a explorar a forma da sentença (Sintaxe); e quando a forma estava estabilizada, a questão foi: o que esta forma significa? Assim, depreendiam o significado (Semântica). Na Linguística Funcional, por outro lado, a direção é inversa, a língua é interpretada como um sistema de significados, acompanhado de formas através das quais os significados são realizados. Assim, a gramática funcionalista busca estudar os significados da língua, vendo a forma como um meio para um fim e não como um fim em si mesma. *"this puts the forms of a language in a different perspective: as means to an end rather than as an end in themselves"*. (HALLIDAY, 1994: xiv).

A teoria de Halliday é conhecida como sistêmica. Para o autor (1994: xiv), uma teoria sistêmica é uma teoria do significado e da escolha, porque a língua ou qualquer outro sistema semiótico é interpretado como uma rede interligada de opções que estabelecem sentido. Halliday propõe uma gramática que possibilite analisar de forma útil os textos tanto na modalidade escrita quanto falada, daí uma gramática voltada ao uso linguístico; além de seu teor interdisciplinar, pois sua gramática parte da análise textual e pode ser relacionada a áreas como: etnografia, literatura, educação, pedagogia, estudo da fala e outras. Pode também ser usada na análise da composição escrita da criança; análise da interação entre professor-aluno; análise de livros didáticos; análise de fala espontânea, ou seja, em diversas variedades linguísticas.

Na parte da análise do discurso, a gramática de Halliday apresenta dois níveis de estudo: um é o nível de compreensão do texto, que propõe analisar e interpretar como e por que o texto significa o que significa. Nesse processo, são ressaltados os diversos significados de um

texto, analisando a ambiguidade, metáforas, etc. O segundo nível é o da avaliação: a análise linguística pretende verificar se um texto cumpre de forma efetiva o seu propósito.

A língua é vista, pelo autor, como um sistema pelo qual se promovem significados, um sistema semântico dentro de outros sistemas para codificação dos significados que produz. A semântica não se refere somente ao significado da palavra: é todo sistema de significados de uma língua, expressado por uma gramática juntamente com seu vocabulário.

Para o autor, a relação existente entre a Semântica ("*meaning*") e a Gramática ("*wording*") não é arbitrária. *"the relation between the meaning and the wording is not, however, an arbitrary one; the form of the grammar relates naturally to the meanings that are being encoded."* (HALLIDAY, 1994: xvii).

A língua adulta construiu estruturas semânticas que nos possibilitam pensar sobre nossa experiência e o sistema de significações que, por sua vez, gerou estruturas léxico-gramaticais, como a distinção entre classes de palavras, verbo e substantivo, que refletem as análises das experiências em processos e participantes. Desta forma, a criança adquire a linguagem e constrói sua gramática: elas podem estabelecer uma ligação entre as categorias gramaticais, a realidade que a circunda e que está dentro de sua cabeça. (HALLIDAY, 1994: xvii).

Conforme Halliday (1994, apud Neves 1994: 118), "uma gramática funcional é essencialmente uma gramática 'natural', no sentido de que tudo nela pode ser explicado, em última instância, com referência a como a língua é usada".

De acordo com Halliday (1994, apud Neves 1994: 118), os componentes essenciais do significado na língua são funcionais. Todas as línguas se organizam por meio de dois principais significados: o "ideacional" (reflexivo); e o "interpessoal" (ativo). Esses componentes, também conhecidos como metafunções, são as manifestações, no sistema linguístico, dos dois principais propósitos que estabelecem todos os usos da linguagem: a) entender o ambiente (ideacional) e b) influir sobre os outros (interpessoal). Relacionado a esses dois, há um terceiro componente metafuncional, o "textual", que lhes confere relevância.

Cada componente de uma língua é explicado por meio de sua função incorporada no sistema linguístico. Deste modo, Halliday postula as metafunções da linguagem.

O quadro a seguir apresenta as metafunções da linguagem postuladas por Halliday.

Metafunção	Definição (tipo de sentido)	Status na sentença	Tipo de estrutura
1. Ideacional (Experiencial)	Construindo um modelo da experiência	Oração como representação	Estrutura de transitividade: Processos, participantes e circunstâncias.
2. Interpessoal	Desempenhando relações sociais	Oração como troca	Estruturas de modo
3. Textual	Criando relevância para o contexto	Oração como mensagem	Estrutura temática e informacional
4. Lógica	Construindo relações lógicas	----------	

Quadro 3: Metafunções.
Fonte: Halliday (1994:36, In: PEREIRA, 2004).

Em primeiro lugar a linguagem serve para expressão do conteúdo, isto é, exteriorizar a experiência que o falante tem do mundo real e do seu mundo interior. Esta função é denominada de **Ideacional**. Dentro desta função reconhecem-se duas subfunções: Experiencial e Lógica.

De acordo com Neves (1994: 111*)*, *"*é por meio dessa função que o falante e o ouvinte organizam e incorporam na língua sua experiência dos fenômenos do mundo real, o que inclui sua experiência dos fenômenos do mundo interno da própria consciência, ou seja, suas reações, cognições, percepções".

Em segundo lugar, a linguagem serve também para manter relações sociais, de modo que um falante consiga alcançar seus objetivos através da interação com outra pessoa, trata-se da função **Interpessoal** da linguagem. "Com esta função o falante usa a linguagem como meio de participar do evento de fala: ele expressa seu julgamento pessoal e suas atitudes, assim como as relações que estabelece entre si próprio e o ouvinte, em particular, o papel comunicativo que assume" (NEVES, 1994:111).

A função **Textual** estabelece relação entre a linguagem e a situação em que ela é utilizada. Abrange a capacidade de construir textos, ou discursos apropriados à situação e também possibilita o receptor da mensagem ou leitor a distinguir o texto de um conjunto aleatório de orações. Estabelece no discurso relação de coesão entre as orações. Esta função, conforme Neves (1994: 111), "diz respeito à criação do texto; por ela, a linguagem contextualiza as unidades linguísticas, fazendo-as operar no cotexto e na situação".

Cada estrutura liga-se a uma determinada função da linguagem.

> Assim, o sistema de transitividade que especifica os papéis dos elementos da oração, como "ator" e "meta", codifica a experiência do mundo, e liga-se, pois, com a função ideacional; o sistema de modo (do qual deriva o de modalidade),

> especificando funções como "sujeito", "predicador", "complemento", "finitude", diz respeito aos papéis da fala e liga-se, pois, com a função interpessoal; os sistemas de tema e de informação, especificando as relações dentro do próprio enunciado, ou entre o enunciado e a situação, dizem respeito à função linguisticamente intrínseca, a função textual (NEVES, 1994: 117).

Desse modo, cada elemento da língua é explicado de acordo com a função que cumpre no sistema linguístico. Nesse caso, uma gramática funcional é a que compõe todas as unidades de uma língua, isso é, as orações e expressões, como formas orgânicas de funções, interpretando cada parte como funcional em relação ao todo.

CAPÍTULO 2: Expressão dos pronomes subjetivos na literatura linguística

Há, na literatura linguística, vários estudos referentes à variação da expressão dos pronomes subjetivos no Português do Brasil. Tais estudos comprovam que a variação entre realização/ não-realização do pronome sujeito é um fenômeno recorrente nas diversas variedades do português brasileiro, sendo encontrada em textos que documentam: a variedade falada não-culta (cf. SILVA, 2003)[1]; a variedade escrita culta (cf. SILVA, 1993 [2]; DUARTE 1993 [3]; TARALLO, 1991) e a variedade falada culta (cf. DUARTE, 1996; NEVES, 2007; TARALLO, 1991).

A maior parte destes estudos recai sobre a análise de todas as pessoas do discurso (cf. DUARTE, 1993, 1996; SILVA, 1993; TARALLO, 1991). Apenas Silva (2003) descreve especificamente a variação do uso dos pronomes de 2ª pessoa do singular e Neves (2007) focaliza o estudo da expressão do pronome de primeira pessoa do singular.

Vale ressaltar também que grande parte dos estudos são sincrônicos (cf. DUARTE, 1996; NEVES, 2007; SILVA, 1993 e 2003), outros diacrônicos (cf. DUARTE, 1993; GALVES, 1993) e há trabalhos em que se contrastam resultados de estudo sincrônico e diacrônico (cf. TARALLO, 1991).

---

[1]No trabalho de SILVA (2003) foi utilizado o *corpus* do Projeto PEUL/ UFRJ. Segundo PAIVA & SCHERRE(1999, p 203): o grupo de pesquisa responsável por esse projeto *" vem se caracterizando pelo interesse no estudo da variedade não culta do português falado no Rio de Janeiro e pela busca de uma teoria mais geral da heterogeneidade e da mudança linguísticas".*

[2]O *corpus* utilizado por SILVA (1993) contém dados de escritores estratificados em3 níveis de escolaridade: os que concluíram o ensino médio, estudantes universitários e graduados.

[3]DUARTE (1993) utilizou um *corpus* de peças teatrais.

É importante notar também que, em todos os estudos, os linguistas utilizam cálculos estatísticos ao desenvolverem análise de fatores linguísticos diversos (funcionais e/ou antifuncionais) que favorecem a realização/ não-realização dos pronomes subjetivos (cf. DUARTE, 1993; NEVES, 2007; SILVA, 2003; SILVA, 1993). O estudo de Duarte (1996) é o único, dentre os estudos analisados, em que fatores sociais são considerados relevantes.

Alguns linguistas se propõem a explicar as motivações da variação, baseando-se, por um lado, em pressupostos da Sociolinguística Variacionista e do Funcionalismo (NEVES, 2007; SILVA, 1993; SILVA, 2003), por outro, da Gramática Gerativa (GALVES, 1993). Há ainda estudos em que se combinam aspectos das teorias Sociolinguística e Gerativista (DUARTE, 1993, 1996; TARALLO, 1991).

Desse modo, faremos, a seguir, uma apresentação de alguns destes estudos destacando os aspectos que contribuíram de alguma forma para o desenvolvimento da presente pesquisa. Em 2.1 e 2.2, discorreremos, respectivamente, sobre as variáveis linguísticas e sociais que podem favorecer ou desfavorecer o uso dos pronomes subjetivos.

2.1 Fatores linguísticos

Muitos são os fatores linguísticos sugeridos pelos pesquisadores que condicionam a variação entre realização/ não-realização dos pronomes subjetivos. Tais fatores podem ser de natureza **funcional**, como: ***compensação funcional; ambiguidade; ênfase; conexão discursiva; referenciação*** ou **antifuncional**, como por exemplo: ***paralelismo formal.*** Encontramos outros fatores como ***polaridade; valor semântico dos verbos*** e também sugestões de análise da teoria gerativista, como a questão do parâmetro ***"pro-drop".*** A seguir apresentamos os estudos que abordam estes fatores.

Silva (2003) desenvolveu um estudo sobre a variação no uso de pronomes de segunda pessoa do singular (as formas *você, cê* e *zero*) na fala carioca.

Ao se referir a pesquisas feitas em outras línguas como o espanhol, a autora apresenta a discussão de que a omissão do pronome sujeito ou a eliminação de marcas de concordância nos sintagmas verbal e nominal podem ser norteadas por dois aspectos: podem ser funcionais, na medida em que eliminam as marcas de redundância, preservando as necessárias para manter clara a informação a ser veiculada; ou pelo contrário, serem antifuncionais, já que podem aparecer repetições de marcas justamente quando menos necessárias, ou seja, marcas redundantes, como, por exemplo, o paralelismo formal, conceituado ao longo do texto.

Silva (2003) nos apresenta uma série de hipóteses de natureza funcional para sua pesquisa, como: "condição de distintividade" de Kiparsky (1972, apud SILVA, 2003: 123): "há uma tendência para a informação semanticamente relevante ser mantida na estrutura superficial". Poplack (1980, apud SILVA, 2003: 123) acrescenta que a perda de informação tende a não ocorrer em contextos que apresentam ambiguidade, o que Hochberg (1986, apud SILVA: 124) coloca como ***"compensação funcional",*** ou seja, formas explícitas aparecem mais em contextos ambíguos.

De acordo com as ideias do texto, a ambiguidade pode ser morfológica (ex: ***você/ ele*** fal***a***); ou textual (elementos referentes no texto que concorrem), assim ela é, muitas vezes, esclarecida por elementos do texto, no caso, pela presença do pronome sujeito, como no exemplo abaixo:

Achei ótima a ideia do José casar com a Carlota, já que há algum tempo ***ela tava*** *a fim de sair de casa.* (SILVA, 1993: 36)

Para verificar se uma sentença, com a forma verbal de 3ª pessoa é ambígua, retira-se o pronome sujeito da oração e se examina a existência de outro candidato a sujeito da frase. No exemplo acima, se retirarmos o pronome sujeito ***ela***, o verbo ***tava*** pode ser interpretado com referência ***a José, eu, você e Carlota***.

Silva (2003) comprovou sua hipótese, ou seja, nos contextos ambíguos o sujeito pronominal[4] explícito apareceu com maior frequência.

A ***ambiguidade*** foi uma das três variáveis utilizadas por Silva (1993) em seu estudo quantitativo sobre a omissão/ realização dos pronomes sujeitos de 1ª, 2ª e 3ª pessoas, em um *corpus* de língua escrita, constituído por 70 cartas pessoais.

Nesse trabalho, com relação à ***ambiguidade***, Silva (1993) quantificou a taxa de omissão do pronome em relação às três pessoas do discurso e constatou que de 1.650 ocorrências de verbo de primeira pessoa, 1.271 (77%) não tinham o pronome sujeito expresso. Isso ocorre devido ao fato de a desinência ser suficiente para identificar o referente. Com a segunda pessoa, das 435 ocorrências, apenas em 130 (30%) o pronome sujeito não é realizado. E quanto a terceira pessoa do singular, somente referentes expressos no discurso imediatamente anterior foram considerados, pois a troca por um novo referente pressupõe sua realização com um novo substantivo. Dos 702 referentes de 3ª pessoa expressos anteriormente, 349 (50%) foram omitidos.

[4] A expressão "sujeito pronominal" refere-se, neste trabalho, aos pronomes pessoais das três pessoas gramaticais do singular e do plural.

A autora percebe que o pronome de primeira pessoa do singular é o mais omitido, por sua forma verbal ter menos ***ambiguidade*** (morfológica e/ou textual) e também por ser mais predizível.

Uma hipótese funcional apresentada no texto é o princípio da iconicidade de Givón (1983, 1990, apud SILVA, 2003: 125): quanto mais predizível uma informação, menos codificação ela recebe, ou seja, menos marcas redundantes serão explicitadas.

Já para a realização do pronome, a autora nota que o sujeito mais explicitado é o de segunda pessoa do singular, tanto no *corpus* de cartas como no de língua falada. Isso deve ocorrer devido à ambiguidade morfológica das formas verbais constante entre a segunda e terceira pessoas do português falado e escrito no Rio de Janeiro, ou seja, em contextos ambíguos, a frequência de realização do pronome sujeito é muito alta.

Desse modo, na análise de meus dados levo em conta a seguinte hipótese fundamentada por Silva (2003): em contextos ambíguos ocorreria com maior frequência o pronome sujeito explícito, e em contextos não-ambíguos, o pronome tenderia a ser omitido.

Além da ambiguidade, fator funcional, Silva (2003) analisou, nos *corpora,* um fator antifuncional: "o princípio do processamento paralelo" ou "***paralelismo***", definido por Scherre & Naro (1991, 1998, apud SCHERRE & NARO, 2007: 63) como a seguinte tendência: formas explicitamente marcadas conduzem a formas explicitamente marcadas, e formas não explicitamente marcadas conduzem a formas não explicitamente marcadas, ou seja, marcas que ocorrem, por exemplo, nos primeiros elementos de um sintagma nominal promoverão o aparecimento de outras marcas, como as marcas de concordância nominal de plural: a**s** menina**s** bonita**s.** O paralelismo também pode ocorrer em textos falados, como na ocorrência retirada do *corpus* de língua falada do Português Popular da cidade de São Paulo, em que a primeira pessoa do singular aparece no início da sentença e se repete ao longo do texto:

Inf *é ...**eu** fui criada di manera diferenti... agora... dessi:... dessi tempu da gora... intão a minha barrera maió foi essa aí... **eu** tinha um medu di homi horrivi... qui mamãe insinava nóis assim... intão **eu** tinha um medu di homi horrivi... foi ondi trapalhô muitu a minha vida aqui... i foi u pió... i foi u pió qui **eu** daí... cumu **eu** sei qui **eu** cheguei aqui nova né?...* (Inf. I M, 13b)

O ***paralelismo formal*** é um fator de natureza antifuncional, pois ocorre repetição de marcas consideradas redundantes no processo de comunicação. Desse modo, do ponto de vista da funcionalidade, o estudo de Silva (2003) traz resultados contraditórios já que fatores tanto de ordem funcional (como ambiguidade) quanto antifuncional (paralelismo formal)

estão atuando nos *corpora e* motivando o uso/ omissão do pronome sujeito de 2ª pessoa do singular.

Em um de seus estudos, Silva (1993) quantifica e analisa a influência das variáveis: ênfase e conexão discursiva, na variável omissão/ realização dos pronomes sujeitos de 1ª, 2ª e 3ª pessoas.

A variável ***ênfase*** é analisada nos dados, sendo que a autora utiliza dois subcomponentes para conceituá-la: contrastividade e reforço.

Quanto à noção de contrastividade, a autora fundamenta-se em Chafe (1976, apud SILVA, 1993: 41) para explicar que nas línguas em que a desinência verbal é suficiente para indicar o sujeito, pronomes são usados, independentemente, por contrastividade. As orações que expressam contraste são identificadas por Silva pelos seguintes critérios estabelecidos: uso de conjunções e advérbios: *mas, embora, ainda assim*; uso de verbos com sentido opositivo: *acordar/ dormir*; construções de *negação/ afirmação* e quando há diferentes complementos para o mesmo verbo: *comprei um Volkswagen zero quilômetros para ela... **eu comprei** para mim um Fiat 80.* (SILVA, 1993: 42).

O conceito de reforço não requer a existência de uma relação entre dois elementos mencionados. O reforço atua, geralmente, em expressões, tais como: *pelo menos; realmente.* Por exemplo*: As ruas ficam mais bonitas, mesmo com nevoeiro. **Eu pelo menos** acho.* (SILVA, 1993: 42).

Com relação à variável ***ênfase***, resultados de Silva (1993) mostram que com a primeira e segunda pessoas, situações de ênfase desfavorecem a omissão do pronome sujeito e contextos não enfáticos favorecem a omissão do pronome sujeito. Em outras palavras, quando há ênfase, há maior probabilidade de realização do pronome sujeito e quando não há, maior é a probabilidade de sua omissão.

No caso da ***conexão discursiva*** (relação de conexão entre o item analisado e o contexto discursivo precedente), a autora notou que a variável mudança de referência correlaciona-se à variável omissão/ realização do pronome sujeito. Desse modo, se há mudança de referente, a probabilidade de realização do pronome subjetivo é maior, e se há manutenção do mesmo referente, é maior a probabilidade de omissão do pronome subjetivo.

O grau de ***conexão discursiva*** condiciona a variável omissão/realização do pronome. A autora estipulou uma escala de conexão discursiva, com seis graus de conexão, partindo do primeiro grau, em que há maior conexão, e o tópico é mantido no discurso, até o sexto grau em que ocorre mudança do tópico discursivo.

Os graus de conexão discursiva são estabelecidos por Silva (1993: 43), pelas seguintes categorias:

1. primeiro grau (ótima conexão): casos em que o mesmo sujeito, o mesmo tempo e modo dos verbos são mantidos; geralmente referem-se a uma sequência de acontecimentos relacionados ao mesmo referente, como nas sequências narrativas. Sob o ponto de vista sintático, trata-se de uma sequência de orações coordenadas:

Ø telefonei para você, Ø continuei a copiar o trabalho. Ø Fiquei de saco cheio, Ø parei de copiar e Ø fui juntar-me aos meus familiares na praia. (SILVA, 1993: 43).

2. segundo grau (boa conexão): o mesmo sujeito é mantido, mas há alguma mudança no tempo, aspecto, ou modo do verbo, que reflete uma mudança no nível discursivo entre o primeiro e segundo plano ou mudanças similares em outros tipos de discurso encontrados no *corpus* (real/ irreal; fato/opinião, etc):

*Ultimamente eu não tenho tido muitas novidades pra te contar, porque **EU ESTAVA** vivendo em completa alienação.* (SILVA, 1993: 43).

3. terceiro grau (conexão levemente enfraquecida): o mesmo sujeito não é mantido, o material entre a ocorrência sob análise e a ocorrência anterior não contém um outro possível candidato para sujeito (por exemplo, a oração de intervenção pode ser impessoal):

*Ø Fiquei estudando ontem até às 3:30 h. A.M. Ainda são 11hs. **Ø vou** ficando por aqui.* (SILVA, 1993: 43).

4.quarto grau (conexão enfraquecida): o mesmo sujeito não é mantido, e sua última ocorrência exercia outra função sintática (objeto direto ou indireto, possessivo, etc):

*Desculpe-me pela minha reclamação infundada. **EU ESTAVA** brincando.* (SILVA, 1993: 43).

5. quinto grau (conexão fortemente enfraquecida): o mesmo sujeito não é mantido, e entre um dado sob análise e sua última ocorrência há uma alternativa para candidato a sujeito:

*Vou ter que estudar segunda-feira mesmo. A Vera vai ter que ser rápida e direta, pois **EU TENHO** mais o que fazer.* (SILVA, 1993: 43).

6. sexto grau (conexão interrompida): o tópico do discurso é alterado, com ou sem mudança do sujeito:

*Diga-me o preço total e envie o número de tua conta aí nos EUA para que possa te enviar a soma correspondente. Quanto ao livro, **EU SEGUI** tuas instruções.* (SILVA, 1993: 43).

Assim, a autora constata que à medida que o grau de conexão discursiva decresce, aumenta a probabilidade de realização do pronome sujeito.

Os resultados quantificados da ausência do pronome de primeira pessoa cruzados com os graus de conexão discursiva mostram que em contextos com maior conexão discursiva, maior é a frequência de omissão do pronome sujeito, e em contextos com menor conexão discursiva, maior é a frequência de realização do pronome sujeito de primeira pessoa.

Como padrão de variação da 1ª, 2ª e 3ª pessoas, a autora constatou que a primeira pessoa é a mais usada e o pronome de 1ª. pessoa é o mais frequentemente omitido o que ocorre devido ao tipo de texto ser de cartas pessoais, tendo como centro o escritor.

Por outro lado, os pronomes de terceira pessoa são usados menos frequentemente, aparecem, em geral, em sequências narrativas e sua taxa de omissão é de 50%. A segunda pessoa ocorreu apenas ocasionalmente e sua omissão foi de 30%.

É o que podemos verificar na tabela a seguir:

Tabela 1: Correlação de frequência relativa de ocorrência das pessoas do discurso e frequência de omissão do pronome.

Categoria	**Total de ocorrências**	**Omissão**
Primeira pessoa	1650/2787 (59%)	1271/1650 (77%)
Segunda pessoa	435/2787 (16%)	130/435 (30%)
Terceira pessoa	702/2787 (25%)	349/702 (50%)

Fonte: SILVA, 1993: 46.

Silva (1993) afirma que há grande correlação entre a frequência relativa de pessoas do discurso e a frequência de omissão do pronome correspondente. Observa-se que a autora enfatiza explicações de natureza funcionalista, no sentido de o pronome ocorrer mais frequentemente quando é mais necessário para a comunicação ser efetivada, ou seja, a autora dá muita importância ao nível discursivo, afirmando que a motivação para o fenômeno morfossintático sob análise é encontrada fundamentalmente nas funções discursivas.

Vale ressaltar que Silva (1993) chama atenção para a questão do processamento paralelo.

Silva (1993: 46) explica que o fenômeno do processamento paralelo ocorre devido à relação de dependência entre o elemento variável (por exemplo: marcas de concordância nominal) e um elemento precedente que atua como um gatilho desencadeador. Por exemplo: a variável desinência de plural no verbo depende da ocorrência de um sujeito no plural (gatilho desencadeador) para provocá-la.

Em seu estudo, Silva (1993) percebe que isso não ocorre, pois o próprio pronome subjetivo é em si mesmo o gatilho desencadeador, não dependendo de algum elemento prévio para desencadeá-lo.

A autora atenta para que em alguns casos o sujeito possa ser considerado como um elemento dependente, assim o processamento paralelo poderia funcionar no sentido contrário, ou seja, para favorecer a presença do sujeito com base na presença categórica da desinência. Porém, estudos prévios (NARO, 1981, apud SILVA, 1993: 46) mostraram que isso não ocorre de fato.

Temos, também o estudo de Duarte (1993), sobre os pronomes sujeitos do Português do Brasil, em que a autora, utilizando um *corpus* de peças teatrais dos séculos XIX e XX, postula 3 paradigmas de formas flexionais dos verbos, ou seja, ela constata que as formas verbais que antes eram seis (paradigma 1) evoluíram para uma paradigma de quatro formas com a perda das segundas pessoas diretas do singular e plural (paradigma 2). Este último paradigma coexiste com um terceiro, em que há a diminuição para três formas flexionais, em consequência da perda do pronome de primeira pessoa do plural: *nós* substituído por *a gente* (paradigma 3). Conforme ilustração a seguir:

Tabela 2: Evolução nos paradigmas flexionais do Português.

Pessoa	**Número**	**Paradigma 1**	**Paradigma 2**	**Paradigma 3**
1ª	Singular	Cant-o	Cant-o	Cant-o
2ª direta	Singular	Canta-s	_____	_____
2ª indireta	Singular	Canta-0	Canta-0	Canta-0
3ª	Singular	Canta-0	Canta-0	Canta-0
1ª	Plural	Canta-mos	Canta-mos	Canta-0
2ª direta	Plural	Canta-is	_____	_____
2ª indireta	Plural	Canta-m	Canta-m	Canta-m
3ª	Plural	Canta-m	Canta-m	Canta-m

Fonte: DUARTE, 1993: 109.

A partir de um estudo diacrônico, a autora analisa se essa simplificação do sistema flexional está influenciando o uso do sujeito pleno ou nulo, verificando também, com base na teoria gerativista, se o Português do Brasil está se tornando uma língua [+ pro-drop] ou [-pro-drop].

Como resultado, ela percebe que no decorrer dos anos, os sujeitos de 1ª e 2ª pessoas foram representados cada vez mais frequentemente por pronome explícito, por outro lado, o pronome sujeito de 3ª pessoa foi omitido com maior frequência.

O fato de o pronome sujeito de 3ª. pessoa poder ser apagado com maior frequência está associado à referência do sujeito, ou seja, quando o referente for esperado, usa-se mais o pronome nulo, caso contrário, o pronome pleno. Então, se não há mudança de referente, o pronome tende a ser omitido, se há mudança de referente, o pronome tende a ser realizado.

Duarte (1993) afirma que no percurso do uso de sujeitos explícitos, durante os séculos XIX e XX, a característica de língua ***"pro-drop"***[5] da Língua Portuguesa foi alterada, sendo que a língua evoluiu de uma tendência de "pro-drop" para "não-pro-drop", ou seja, para o uso do pronome explícito.

Em seu outro estudo a respeito da perda do princípio "Evite Pronome", Duarte (1996) constatou que o português do Brasil passa por um período de mudança no que diz respeito à identificação do sujeito nulo referencial, que coincide com a simplificação do paradigma flexional em consequência da redução ocorrida no quadro pronominal.

Não há em PB[6] uma relação direta entre presença de marca distintiva de pessoa no verbo e ocorrência de sujeito nulo. A diminuição da frequência do sujeito nulo espalha-se por todo o paradigma, o que sugere que este perdeu a sua *riqueza funcional.*

A autora baseia-se na hipótese de Rizzi (RIZZI, 1988: 15, apud: DUARTE, 1996: 88) "o sujeito pronominal, sendo focalizado ou contrastivo, precisa ser foneticamente realizado". Além do mais, observa-se clara preferência do falante pelo uso da forma pronominal plena. Atualmente, as formas pronominais nula e plena são intercambiáveis, e o uso desta não compromete a aceitabilidade da sentença. A autora percebeu que o PB perdeu o princípio "Evite Pronome" e caminha, em virtude dessa perda, em direção das línguas ***não-pro-drop.***

Esses estudos de Duarte dialogam com o de Tarallo (1991) na medida em que o autor diagnostica a emergência de uma língua portuguesa brasileira, com características que a diferenciam do português europeu.

O autor apresenta em seu estudo quatro grandes mudanças sintáticas do português brasileiro. Para nosso trabalho, interessa-nos apenas a seguinte mudança: "reorganização do sistema pronominal que teve como consequências a implementação de objetos nulos no sistema de um lado, e sujeitos lexicais mais frequentes de outro". (TARALLO, 1991: 70)

---

[5] O termo "pro-drop" trata-se de um parâmetro, da gramática gerativa, em que as línguas tendem a "deixar cair" seus argumentos sentenciais.

[6] PB é a abreviação para Português do Brasil.

Tarallo (1991) contrasta os resultados de dois de seus estudos sobre o uso dos pronomes: um sincrônico, com *corpus* de entrevistas de 1981; outro diacrônico, de língua escrita (cartas, diários e peças teatrais), dos séculos XVIII e XIX, divididos em 4 períodos de 50 anos: tempo I (*circa* 1725), tempo II (*circa* 1775), tempo III (*circa* 1825) e tempo IV (*circa* 1880).

No estudo sincrônico, os resultados de retenção vs. apagamento pronominal com relação às categorias sintáticas mostram maior tendência ao preenchimento da posição de sujeito e menor frequência de preenchimento pronominal da posição objeto direto, conforme tabela abaixo:

Tabela 3: Retenção vs. Apagamento pronominal em várias categorias sintáticas: *corpus* de 1981.

POSIÇÃO	sujeito	objeto direto	objeto indireto	oblíquo	genitivo
Retenção	572	72	58	60	46
%	79.4%	18.2%	40.8%	36.1%	79.3%
Apagamento	148	324	84	106	12
total	720	396	142	166	58

Fonte: TARALLO, 1991: 82.

.

No estudo diacrônico, o autor apresenta a porcentagem de retenção pronominal, segundo a categoria sintática e o período do tempo. Os resultados mostram que a regra de retenção se mantém quase categórica nos sintagmas preposicionais nos tempos I e II, mas a partir do tempo III começa haver decréscimo. Além disso, nota-se que à medida que as porcentagens para os sintagmas preposicionados e objeto direto decaem, aumentam as porcentagens de retenção pronominal no sujeito. É o que podemos constatar na tabela 4:

Tabela 4: Retenção pronominal segundo função sintática e período de tempo (dados de 1725 a 1880).

circa	I 1725	II 1775	III 1825	IV 1880
Sujeito %	28/120 23.3%	41/154 26.6%	25/152 16.4%	57/154 32.7%
Objeto direto %	83/93 89.2%	51/53 96.2%	36/43 83.7%	59/98 60.2%
Objeto indireto %	49/49 100%	41/41 100%	48/49 97.9%	49/62 79%
Oblíquo %	18/20 90%	25/25 100%	4/4 100%	24/43 55.8%
Genitivo %	17/18 94.4%	27/28 96.4%	43/51 84.3%	35/43 81.4%

Fonte: TARALLO, 1991: 83.

Tarallo (1991: 83) sugere uma correlação entre o decréscimo de frequência de realização do objeto direto e sintagmas preposicionais e o aumento da realização do sujeito. O autor afirma que a perda de referência pronominal faz com que o sistema se rearranje, marcando outros argumentos sentenciais mais frequentemente.

O autor, com base nos resultados das tabelas 3 e 4, conclui que para as 3 posições principais, **sujeitos, objetos diretos e sintagmas preposicionais**, ocorreu uma mudança na hierarquia para retenção pronominal.

Nos dados diacrônicos, havia a seguinte hierarquia: **SPs > objetos diretos > sujeito,** contudo por volta de 1880 ocorre a mudança no sistema pronominal, sendo que a frequência de retenção pronominal começa a decrescer para SPs (em menor escala) e objetos diretos (em grande escala) enquanto a porcentagem de realização dos sujeitos começa a crescer, tendo a seguinte hierarquia para os dados de 1981: **Sujeitos > SPs > objetos diretos**. É o que mostra a tabela abaixo:

Tabela 5: Resumo da porcentagem de retenção pronominal.

TEMPO **Função:**	**1725**	**1775**	**1825**	**1880**	**1981**
Sujeito	23.3%	26.6%	16.4%	32.7%	79.4%
Objeto direto	89.2%	96.2%	83.7%	60.2%	18.2%
(SPs)	96.5%	98.9%	91.3%	72.9%	44.8%

Fonte: TARALLO, 1991: 84.

Assim, os resultados dos dois trabalhos, sincrônico e diacrônico, comprovam que, ao longo do tempo, o pronome sujeito tem sido mais explicitado, enquanto os pronomes com função de objeto direto mais omitidos.

Outro estudo essencial para a realização desta pesquisa é o de Neves (2007) sobre a predicação e os predicados e a realização do sujeito. A autora parte da ideia de que falamos por predicações, acredita que a unidade básica de análise é a oração, e dá grande importância ao contexto. Essa ideia condiz com a proposta funcionalista de Halliday, de que a sistematização vai além dos limites da oração, ou seja, abrange o texto.

De acordo com Halliday (1985, apud NEVES, 2007: 38), apesar de a oração ser a unidade fundamental de realização da linguagem, pois cumpre funções na linguagem, o estudo da linguagem não deve parar nos limites da oração (relação construcional que norteia as gramáticas normativas), pois a sistematização pode chegar às formas de organização não-

construcionais, isso é, fora dos limites da oração, ou seja, ao lado dela (no domínio da entonação e do ritmo), em torno dela (no domínio da coesão e do discurso) e além dela (no domínio dos modos metafóricos de expressão). Para Halliday (1985, XXII, apud NEVES, 2007: 38), "o texto é o objeto da linguística, juntamente com o sistema".

Para Neves, nas propostas funcionalistas o verbo é o centro da oração, e todos os termos que preenchem a valência do verbo serão argumentos, que mantém relações de dependência e hierarquização. Os argumentos (ou complementos) preenchem lugares vazios diferentes, portanto, todos eles são diferentes entre si, estando entre eles o sujeito, ou seja, um argumento de diferente estatuto, já que ele é o escopo da predicação que se opera na oração, ou seja, o argumento mais relevante.

O estudo de Neves (2007) se fundamenta na análise da valência dos verbos, isto é, os argumentos (sujeito e objeto) que preenchem as casas que participam da estrutura argumental.

Para Neves (2007: 40) "a estrutura argumental é um fenômeno primariamente sintático (ligado especialmente à noção da obrigatoriedade de determinados termos para preencher a valência de determinados verbos) que envolve a semântica (já que há restrições semânticas nesse preenchimento) e a pragmática (já que a realização efetiva do sistema de transitividade resulta de necessidades e intenções comunicativas)".

A autora afirma que o estudo da estrutura argumental se consolida no nível pragmático. Por esse nível é que se pode entender o condicionamento discursivo da construção gramatical das orações, ou seja, entender como um determinado termo/ complemento é necessário ou não, na oração, bem como o modo de realização dos termos, seus valores e significados.

No caso da realização do sujeito, a autora afirma que há razões discursivo-textuais para ocorrer o sujeito expresso nas orações. Como por exemplo: situações em que ocorre ambiguidade entre referentes, e usa-se o sintagma nominal como elemento desambiguador.

Neves, em seu estudo de 1994 (apud NEVES 2007), examinou no *corpus* de língua falada, 163 orações de 5 inquéritos do NURC, qual das 3 formas: **sintagma nominal; pronome lexical ou zero** era mais frequente, analisando os fatores: função semântica dos argumentos e transitividade dos verbos. A autora conclui, com base nos resultados, que os esquemas mais realizados são: argumento agentivo (sujeito) de verbo transitivo direto é mais frequentemente realizado como: pronome ou zero (forma não-lexical), ex.: ***Ele** é capaz... de desenhar este animal.* (NEVES, 2007: 43) e argumento não-agentivo não-preposicionado (objeto direto) é mais frequentemente realizado como nome (forma lexical), ex.: *Ele vai tentar usar **esta criação*** (NEVES, 2007: 43).

Em sua outra pesquisa (Ilari, Franchi e Neves, 1996, apud NEVES, 2007: 44) referente à expressão dos sujeitos por meio de uma forma pronominal, Neves ressalta que, nas variedades do Português Brasileiro, ocorreu diminuição no quadro dos pronomes subjetivos, o que resultou na diminuição das formas verbais: de 6 para 4 (cf. conceito gerativista de parâmetro *pro-drop*.)[7]. Estas formas resultantes são bem discriminadas do ponto de vista fônico, de modo que a realização simultânea do sujeito pronominal e da desinência verbal pode ser interpretada discursivamente como redundante.

A autora se propõe a responder qual o papel semântico ou pragmático de um sujeito redundante em relação à flexão verbal. Para isso, ela pesquisa os casos em que o sujeito é omitido, analisando todas as pessoas do discurso, e constata que a primeira pessoa do singular é a que mais dispensa expressão de sujeito. Desse modo, enfoca o estudo da realização/ não realização do pronome de primeira pessoa do singular analisando diversos fatores linguísticos como condicionantes ou não da variação.

Um dos fatores considerados por Neves (2007: 205), em seu estudo, foi a ***polaridade***, ou seja, a escolha entre o positivo e o negativo na proposição, como em *é/ não é*.

A autora analisou um inquérito do projeto NURC e pesquisou quais os tipos de verbo em que o pronome de primeira pessoa não é expresso. Ela constatou que de 96 casos de omissão do pronome *eu*, 21 ocorreram na expressão *sei lá* e 15 na expressão *não sei*, ou seja, em formas negativas cristalizadas há o favorecimento da omissão do sujeito.

No mesmo inquérito, a autora verificou que de 230 ocorrências do pronome *eu*, menos de 20 são de enunciados negativos, e para um total de 96 casos de omissão do pronome *eu*, menos de 30 são de enunciados afirmativos. Assim, a forma negativa favorece a omissão do sujeito, enquanto a forma afirmativa favorece o *eu* expresso (fator polaridade). O pronome *eu* expresso só não ocorre em frases afirmativas quando há determinações sintáticas: coordenação de orações com o mesmo sujeito, com o sujeito expresso na primeira oração, ex.: ***Eu*** *fui... Ø* ***achei*** *um cenário uma coisa ah ótima* (NEVES, 2007: 46) ou subordinação de uma completiva a uma principal com o mesmo sujeito, este expresso na primeira oração, ex.: *Eu acho que* ***Ø vou*** *ao cinema* (NEVES, 2007: 46). Ou seja, quando o referente da primeira oração é mantido, tende a ser omitido na oração afirmativa subsequente. Quanto ao valor semântico dos verbos, a autora constatou que os **verbos** de julgamento, **opinativos** (*achar*,

[7]O português dispensa em muitas circunstâncias o uso dos pronomes sujeitos, fato que as gramáticas tradicionais explicam pela existência das desinências verbais e que a gramática gerativa reconhece enquadrando o português no parâmetro *pro-drop. (*NEVES, 2006: 45)

pensar, acreditar etc) favorecem o sujeito explícito, pois têm a função de marcar a presença do sujeito modalizador, especialmente se este for o falante.

Alguns fatores sintáticos que favorecem o uso do pronome *eu* são: **deslocamento do objeto direto para antes do verbo**, ex: *filme **eu** gosto mais de comédia* (NEVES, 2007: 47) **e o contexto de oração relativa**, ex.: *o pouco que eu::**eu** imagino* (NEVES, 2007: 47).

Outro fator que favorece o uso do pronome sujeito expresso é o **emprego do infinitivo pessoal**, em oposição ao infinitivo impessoal, com a finalidade de evitar a interpretação de referente indeterminado, a leitura genérica. Ex.: *Doc - o que precisa conter o cinema para levar a senhora até ele? Inf- **eu saber** que o filme é bom* (NEVES, 2007: 47). Podemos notar, no exemplo abaixo, que o emprego do *eu* como indeterminador do sujeito é o fator pragmático que torna categórica a expressão do pronome de primeira pessoa do singular. Isso se dá, por ser o pronome eu, na sua forma básica, determinado, assim seu uso como indeterminado torna-se o mais marcado, sendo necessária sua explicitação para indicar que a função é genérica, indeterminada. Ela apresenta o exemplo que segue.

> *Por exemplo **eu** posso saber todos os sinais de trânsito de cor, (es)tá, **eu** memorizei o meu processo, se você me trouxerem o livrinho aquele **eu** respondo todos eles e estou no nível de conhecimento: bem, mas é preciso que **eu** aplique, que **eu** utilize os sinais de trânsito na hora certa, ou que **eu** tenha a habilidade de passar meio rápido pelo guardinha, porque senão **eu** estou multada na próxima esquina. Então, quando é que **eu** com, que **eu** compreendi? Quando **eu** apliquei os sinais de trânsito na hora exata, quando **eu** passei um ano inteiro sem receber nenhuma multa* (NEVES, 2007: 48).

Como último fator analisado, temos: **determinações prosódicas** ligadas a efeitos discursivos, que condicionam a realização/ não realização do <u>*eu*</u>. A autora constata que em perífrases há maior incidência de zero, ex.: *Ø vou te contar* (NEVES, 2007: 48), em casos de clivagem há favorecimento do pronome expresso <u>*eu*</u>, ex.: *como é que eu vou dizer* (NEVES, 2007: 48) e em contextos em que o verbo é precedido de uma palavra com entonação enfática há maior ocorrência de zero, ex.: <u>*AÍ*</u> *Ø **achei** fabuloso* (NEVES, 2007: 48).

Há outras hipóteses que Neves (2007) não analisa, mas considera relevantes para estudos posteriores, como: a presença de pronome sujeito poderia anunciar uma mudança de "tópico discursivo"; haveria correlação entre presença/ ausência do sujeito pronominal e “riqueza” de flexão de pessoa e tempo do verbo que com ele concorda. Tais sugestões de

Neves foram consideradas relevantes para esta pesquisa; partindo delas será possível definir alguns fatores a serem examinados.

A autora conclui que na variação realização/ não realização do pronome sujeito de primeira pessoa não há determinações rígidas, mas construções preferenciais condicionadas por fatores prosódicos, sintáticos, semânticos e pragmáticos.

Um outro fator relevante para esta pesquisa é a referenciação.

Referenciação envolve interação e intenção do falante em se referir, no processo de comunicação, a algum indivíduo identificado como existente.

Para a autora (DIK, 1997, p. 129, apud NEVES, 2007: 75) há pelo menos dois modos de referenciar textualmente: construtivo, construção de referentes, ou seja, o falante usa um termo para que o ouvinte construa um referente para este termo e introduza esse referente em seu modelo mental; e o modo identificador, isto é, identificam-se no texto os referentes que constituem a rede referencial, ou seja, o falante usa um termo para que o ouvinte identifique um referente que já esteja disponível, o que ocorre quando há uma fonte para identificação.

Por outro lado, o processo de ***referenciação*** não se limita a construir e identificar referentes, mas trata da própria constituição do texto como uma rede em que referentes são introduzidos como objetos-de-discurso (Apothéloz e Reichler-Béguelin, 1995, apud NEVES, 2007: 76) e como tais são mantidos, segundo estratégias dependentes da formulação textual.

Assim, a progressão e manutenção referencial que abrange o texto _ representada por preservação de referentes introduzidos, introdução de novos referentes, retomada e reintrodução de referentes_ se relacionam à progressão ou manutenção tópica, que sustenta a organização informativa e dirige o fluxo de informação.

Além do conceito de ***referenciação***, há o de ***correferenciação***. ***Correferenciação*** ocorre quando um referente determinado já foi introduzido no discurso, o falante o representa em outro ponto do enunciado como elemento “dado”, e não apenas como elemento “conhecido”, assim o termo que se refere a ele, além de implicar referenciação, implica correferenciação (NEVES, 2007: 92).

Com relação à expressão do sujeito, a variação entre zero, pronome e sintagma trata-se mais de um caso de intenção de expressão de correferência do que referência independente. Assim, conforme Neves (2007: 93) “se o falante quer expressar correferência, ele preferirá, sempre que possível, zero a pronome, e pronome a sintagma nominal pleno”. O mesmo ocorre com o ouvinte ao interpretar a preferência de uso das formas reduzidas pelo falante.

No exemplo abaixo, podemos ver a intenção de expressão de correferência:

Valentim *lavou* ***a cara*** *e* ***as mãos numa fonte****, Ø examinou* ***os cavalos*** *e* ***as mulas****, Ø saltou n****o dorso*** *de* ***seu animal*** *e Ø ordenou* ***a partida.*** *Ø Viajou adiante, sempre a alguma distância d****as mulheres,*** *junto com* ***o condutor****, às vezes sumindo* ***numa curva. O negro*** *que* ***os*** *guiava, Ø sorria como um sagui e Ø falava baixo, escondendo* ***a boca*** *com* ***a mão****. Valentim conversava com* ***ele****, Ø saltava d****o cavalo*** *para observar* ***a trilha,*** *Ø afastava* ***a vegetação****, Ø previa* ***os perigos****, Ø decidia* ***a direção*** *a ser tomada.* ***Mariana*** *nem mesmo sabia* ***os nomes*** *d****as salvaguardas*** *ou d****os escravos. Sua mãe*** *sempre* ***a*** *ensinara a não se envolver com gente de casta inferior. Ø Devia tratar* ***as negras*** *com distância, despedi-****las*** *com discrição e dar* ***ordens*** *murmurando* (NEVES, 2007: 93).

No exemplo acima, retirado do texto de Neves (2007), podemos notar que ocorrem três aspectos:

a) muitos zeros na posição de sujeito, marcando a elipse de elemento correferencial, típico de casos em que a informação é dada;

b) alguns pronomes pessoais representando elemento correferencial, em posição mais marcada, como objeto indireto (com ele) e direto (os, a e las);

c) muitos sintagmas nominais plenos, que introduzem referentes novos no texto, criando objetos-de-discurso.

A proposta de Levinson (1987; 1991 apud NEVES, 2007: 93) que afirma: "quanto menor a forma da expressão referencial maior a preferência por uma leitura de correferência", tem sido criticada por Ariel (1996, apud NEVES, 2007: 94). Este considera que o falante faz a distinção entre correferência e independência referencial, mas que a função primeira das expressões referenciais é marcar diferentes graus de acessibilidade na memória. Assim o falante não pretende identificar se uma expressão deve ser interpretada correferencialmente ou independentemente, mas sim constatar o quão acessível é, para o ouvinte, uma entidade mental.

Outro aspecto importante é a ***referenciação tópica,*** ou seja, "aquela que cria e relaciona as entidades que entram na organização informativa realizada pelas proposições do texto e nascida da organização das predicações estabelecidas entre predicados e argumentos, com seus respectivos papéis" (NEVES, 2007: 99).

É o que se verifica no exemplo abaixo, em que ora é tópico discursivo *os trabalhadores,* ora é *os opressores*:

C: *Agileu dessa vez tá com a razão! O negócio é sair e passear hoje, amanhã apelar para a greve, opinião pública do nosso lado, boró pela imprensa.*

Jo: *a greve* ***<u>eles</u>[os trabalhadores]*** *num tão topando muito, não. O que o pessoal tá querendo mais é desabafa pela rua. Arriscar ordenado ninguém tá querendo.*

Al: *vida complicada, não é. João?*

J: *oh, se é! E não adianta, viu... minha vontade é berrá por aí: me esquece, gente. Eu não estou pedindo nada demais... eu quero fica sossegado, ter minha família em paz, sustenta meu filho. Mas não. É problema em cima de problema...* ***<u>Eles</u> [os opressores/repressores]*** *têm de te meter no bolso... ah, que eu tou farto!* (NEVES, 2007: 96)

No exemplo acima há sequências de referenciadores de terceira pessoa do plural (representados pela ocorrência do pronome *eles* ou por marcação morfológica). "Os referenciadores assinalados não configuram retomada correferencial de antecedente explícito (sintagma nominal), mas funcionam em anáforas que remetem a referentes de algum modo implicados no discurso, basicamente tópicos discursivos" (NEVES, 2007: 96).

Os referenciadores (no exemplo *os trabalhadores* e *os opressores*) constituem o texto e são interpretados em certos momentos como tópicos discursivos, "independentemente da existência, ou não, de sintagmas nominais antecedentes que tornem explícitas as referências" (NEVES, 2007: 96). Esses referenciadores passam, assim, a compor a teia referencial do texto, tornando-se tópicos discursivos e constituindo o conjunto temático do texto.

Muitas vezes, quando há mudança de tópico-discursivo, o pronome tende a ser explicitado. No exemplo anterior, há alternância de tópico-discursivo e os pronomes são explicitados, mas o referente do pronome "eles" não pode ser identificado. Assim, podemos perceber que em alguns casos, não é somente o pronome que determina seu referente, mas sim este é esclarecido pelo contexto discursivo.

Para esta pesquisa, também foi utilizado o *Dicionário Gramatical de Verbos do Português Contemporâneo do Brasil*, (BORBA, 1991). Nesse dicionário os verbos são analisados como sendo o ponto de partida na estruturação da frase, ou seja, como núcleo do predicado em torno do qual os demais argumentos se arranjam em graus diferentes de coesão e dependência. Assim, descreve-se a **valência** dos verbos (as relações de dependência entre o verbo e seus argumentos) como sendo responsável por determinar suas propriedades sintático-semânticas. Desse modo, os verbos são caracterizados da seguinte forma:

Verbo de ação: o primeiro Argumento é um sujeito *agente*, e não há argumento *afetado* pela ação verbal, isto é o argumento não tem nenhuma mudança de estado físico ou moral, de condição, de posicionamento no tempo ou no espaço.

Ex.: Este rapaz não bebe, não fuma, não mente. (BORBA, 1991, xiii)

Verbo de processo: o primeiro Argumento é sempre um *afetado* por algo que está fora dele, ou seja, é um *paciente*, um *experimentador*, ou um *beneficiário* do processo.

Ex.: O galo morreu. (BORBA, 1991, xiii)

Verbo de ação-processo: tem no mínimo dois argumentos: um *agente/causativo* e outro, *afetado/efetuado*, isto é, o primeiro Argumento faz que o segundo Argumento mude de estado, condição, ou então passe a existir.

Ex.: O vento derrubou o muro. (BORBA, 1991, xiii)

Verbo de estado: tem sempre um Argumento que é inativo, não é *agente*, nem *causativo*, nem *paciente*. O sujeito desses verbos funciona como suporte de propriedades (estado/ condição) ou como mero experimentador delas.

Ex.: Paulo ama Dirce. (BORBA, 1991, xiii)

2.2 Fatores sociais

Dos estudos analisados, somente no de Duarte (1996) fatores sociais são considerados relevantes.

Entre os fatores sociais considerados pela a autora, destacam-se: **faixa etária** dos informantes, utilizada para hierarquização dos condicionantes linguísticos à realização do sujeito e para evidenciar o progresso da mudança linguística; **o sexo** do informante e **o papel do indivíduo** na comunidade linguística.

A autora constata que com relação às três pessoas do discurso, o fator faixa etária é relevante. O sexo do informante mostrou ser um fator relevante apenas para a 3ª pessoa, pois os homens se mostraram mais conservadores, apagando mais o sujeito.

A autora utiliza a variável faixa etária cruzando-a com a pessoa gramatical e evidencia mudança em tempo aparente. Ela percebe que os efeitos da redução do paradigma flexional e pronominal na perda do sujeito nulo são graduais e não atuam uniformemente em todas as pessoas gramaticais.

A autora quantificou a frequência de ocorrência de sujeitos nulos com relação às pessoas discursivas e a três faixas etárias: grupo 1 (de 59 a 74 anos), grupo 2 (entre 45 e 53 anos) e grupo 3 (entre 25 e 32 anos).

Os resultados podem ser visualizados no gráfico a seguir.

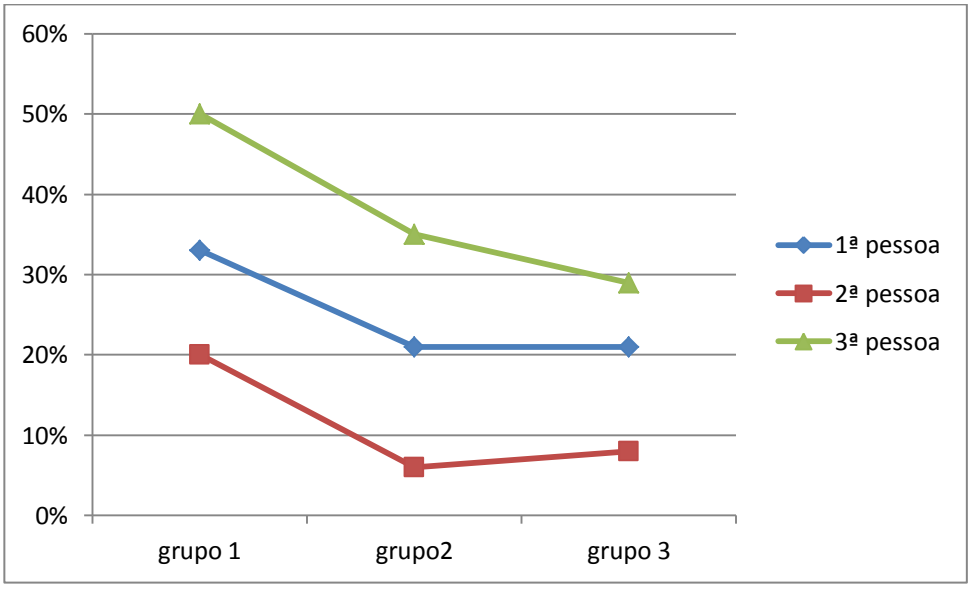

Gráfico 1: Ocorrência de sujeito nulo na fala espontânea segundo a pessoa gramatical e a faixa etária (%).
Fonte: DUARTE, 1996: 96.

No gráfico acima, conforme Duarte (1996: 96), percebe-se que os percentuais mais baixos de uso do sujeito nulo ocorrem com a segunda pessoa, em que temos 20% de sujeitos nulos no grupo 1, 6% no grupo 2 e 8% no grupo 3.

A primeira pessoa tem 33% de ocorrências para o grupo 1 e 21% para os grupos 2 e 3. Já a terceira pessoa mostrou ter os índices mais altos de sujeitos nulos: 50% (grupo 1), 35% (grupo 2) e grupo 29% (grupo 3). Estes resultados evidenciam a ocorrência de mudança em progresso, passando do maior uso de sujeitos nulos entre os idosos (grupo 1) para menor uso de sujeitos nulos entre os jovens (grupo 3), assim a curva de mudança linguística é em declive.

Duarte (1996: 96) ressalta a hierarquia mantida nos três grupos, revelando uma diferença acentuada entre os idosos, de um lado (grupo 1), e os grupos 2 e 3, de outro. Destaca-se, ademais, que o uso do pronome **nós** predomina sobre **a gente** no grupo 1, concorre em igualdade na fala do grupo 2 e é quase categórica a preferência para o uso de **a gente** no grupo 3 (os mais jovens).

Desse modo, com base nos estudos expostos acima, conclui-se que há vários fatores que condicionam a realização/ não-realização dos pronomes subjetivos. Tais trabalhos foram essencialmente relevantes como aporte para esta pesquisa, tornando possível obter uma série de fatores a serem testados em uma variedade diferenciada do Português Brasileiro, ou seja, em um *corpus* do Português Popular falado na cidade de São Paulo.

CAPÍTULO 3: O *corpus* sob análise

A presente pesquisa integra o subprojeto "*Retrato sociolinguístico da variedade popular na cidade de São Paulo no final do século XX e no começo do XXI"*, parte do projeto temático de equipe História do Português Paulista (Projeto Caipira), financiado pela FAPESP (Proc.06/55944-0), na medida em que constitui um estudo de um aspecto linguístico dessa variedade do Português Brasileiro.

3.1 O Português Popular em São Paulo

Durante a década de 1980, ainda eram poucos os estudos que abordavam a língua falada, menos ainda a língua falada por pessoas com baixa ou nenhuma escolarização, moradoras das favelas de São Paulo, ou seja, os falantes do Português Popular.

Com o intuito de desenvolver sua tese de doutoramento, a Doutora Angela Cecília de Souza Rodrigues, entre os anos de 1986 e 1987, olhou para essa variedade de língua e nela pode perceber a existência de um campo rico para o desenvolvimento de estudos Sociolinguísticos.

Por detrás da fala dessas pessoas vítimas de preconceito social e linguístico, rotuladas como "favelados", a professora Angela, com um gravador em mãos, coletou dados dessa variedade de língua, pois se tratava de uma língua viva, utilizada em situações reais de comunicação cotidiana.

Tal proposta fundamentou-se em dois pressupostos: a) a existência em São Paulo de uma variedade linguística popular ou dialeto social popular, cujas características podem ser construídas em contraste com as do dialeto culto, enquanto variedade de referência; b) a variedade de língua popular é utilizada por falantes não escolarizados ou de baixo nível de escolaridade. (Rodrigues, 1987).

Com a coleta do *corpus*, a pesquisadora percebeu como era a vida na comunidade, seus costumes, hábitos e pode, através de uma descrição minuciosa, desenvolver suas reflexões sobre a realização da regra de concordância verbal nas 1ª e 3ª pessoas e comprovar que há padrões de comportamento linguístico correlacionados com a estrutura social.

O material coletado que compreende o vernáculo do Português Popular é composto por 40 entrevistas de 30 minutos de duração cada uma, totalizando 20 horas de gravação de conversa espontânea com informantes adultos, vindos de diversas regiões brasileiras. Os informantes foram selecionados de modo aleatório para que o grau de amostragem fosse prevalecido, garantindo que todos teriam o direito de participar. Eles estão estratificados conforme os três fatores previamente estabelecidos: sexo, idade e escolaridade e estão discriminados, conforme Rodrigues (1987), da seguinte forma:

- Os dois sexos: (1) masculino (16 informantes); (2) feminino (24 informantes);
- Dois níveis de escolaridade: (0) nula (18 informantes); (1) Ensino fundamental I (de 1ª a 4ª série – 22 informantes);
- Três faixas etárias: (1) 20 a 35 anos (10 informantes); (2) 36 a 50 anos (16 informantes); (3) mais de 51 anos (14 informantes).

Com relação ao fator faixa etária, a autora considerou adultos os indivíduos a partir dos 20 anos, pela razão de já possuírem uma vida economicamente ativa e por já terem estabelecido o seu vernáculo.

Outro aspecto relevante na pesquisa foi o da procedência dos moradores das favelas. Rodrigues (1987, 2009:157) constatou que havia uma minoria de paulistanos e predominância de uma população de origem rural, oriunda de "espaços" geográficos diferentes: Noroeste do Estado de São Paulo e Norte do Paraná; Norte de Minas Gerais e Sul da Bahia, e Nordeste (Pernambuco, Alagoas, Ceará, Paraíba). Essas pessoas apresentavam as mesmas características com relação às condições de vida e problemas cotidianos na comunidade. Por esses motivos, foi possível caracterizar a comunidade linguística e consequentemente sua variedade de língua.

As entrevistas ocorreram dentro dos próprios barracos dos moradores e foram desenvolvidas de modo espontâneo, sem roteiro pré-estabelecido, sendo que o assunto surgia durante o decorrer das conversas.

Com a intenção de coletar o vernáculo_ *"o estilo em que se presta o mínimo de atenção ao monitoramento da fala"* (LABOV, 2008: 244)_ a documentadora fez uso de algumas estratégias, como o preenchimento prévio de uma ficha contendo as características sociais dos informantes, sendo possível criar uma situação de descontração, informalidade e

confiança mútua. Por meio desses instrumentos, a documentadora conseguiu romper com o *paradoxo do observador*, este definido como: o objetivo de definir como as pessoas falam quando não estão sendo observadas, no entanto, só podemos obter tais dados por meio da observação sistemática. (LABOV, 2008: 244)

Rodrigues (1987) enfrentou o problema de se gravar a fala em ambientes naturais. Em alguns inquéritos a gravação ficou comprometida, devido ao pequeno espaço dos barracos e frequente movimentação de familiares e amigos que se sentiam à vontade para participar das conversas, interferindo durante as entrevistas.

Assim, o trabalho de doutorado (*A concordância verbal no Português Popular em São Paulo)*, da referida autora, tem um grande valor no meio acadêmico por vários motivos. Entre eles, por inovar o campo das pesquisas linguísticas que eram até então (anos de 1980) focadas na análise da língua escrita ou no português culto falado na cidade de São Paulo, mudando desta forma a óptica para a análise da língua falada da variedade do português popular.

O trabalho de Rodrigues (1987) possibilitou também a coleta e o armazenamento de um *corpus* do Português Popular falado na cidade de São Paulo que serviu de subsídio para o desenvolvimento desta pesquisa e de muitos outros trabalhos já realizados, como (PEREIRA, 2007), e outros que ainda estão por vir.

3.1.1 Características do Português Popular

Os informantes entrevistados são verdadeiros falantes do Português Popular. Este se caracteriza por ser uma variedade de menor prestígio social, não é padrão de correção e adequação, pelo contrário, estigmatizada, por isso não codificada, já que não estabilizada nem normalizada, porque não padronizada. Trata-se de uma língua espontânea e natural, de cunho oral (RODRIGUES, 1987).

O Português Popular é, geralmente, falado por pessoas que vivem à margem da sociedade, com pouca escolaridade e baixo nível socioeconômico. Trata-se de significativa camada populacional de baixa renda, cujos componentes tendem a não possuir qualificação profissional e que em geral habitam os bairros periféricos dos grandes centros urbanos brasileiros, como São Paulo, principalmente favelas e núcleos de habitação popular (RODRIGUES, 2009).

O parâmetro grau de escolaridade é fundamental para identificação dessa população usuária do Português Popular em São Paulo. É o que se esclarece em:

> Outro atributo social desses indivíduos é o seu grau de escolaridade: por não terem acesso à escola, que em tese, constitui instrumento de divulgação de um padrão culto de realização linguística, tendem, em princípio, a manter os hábitos adquiridos na fase de aquisição e definição de seu vernáculo. Se já nos primeiros anos de vida aprendem a dominar os mecanismos básicos do português oral, a sociedade não lhe 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. Por isso, traços de uma variedade do português mais antiga podem ser surpreendidos em sua fala (RODRIGUES & FERREIRA NETTO, 2000: 172).

Muitas vezes, o falante do português popular não frequentou a escola, desta forma, a variedade de língua que domina é justamente a adquirida através da oralidade, aprendida e utilizada dentro de suas famílias, por pessoas próximas, enfim, é sua língua materna, sua língua natural, seu vernáculo. Se no período de desenvolvimento do vernáculo, ou seja, da *"variedade adquirida pelo falante nos anos de sua pré-adolescência"* (LABOV, 1974, apud RODRIGUES, 2009: 11) o falante não for exposto a outras variedades de língua, como a língua padrão, utilizada, regulamentada e ensinada na escola, dificilmente ele desenvolverá sua competência comunicativa, permanecendo na sua variedade de língua estigmatizada (português popular).

Outro aspecto importante que serve de parâmetro para a caracterização dessa variedade de língua é a origem do falante. Na maioria das vezes, os usuários do português popular em São Paulo, em meados da década de 1980, eram migrantes, provenientes de diferentes regiões do Brasil, principalmente do Nordeste, com predominância da zona rural. Desta forma, percebe-se que a capital paulistana se tornou um grande centro de variação linguística, devido a sua diversidade cultural, resultante desses processos migratórios.

Esses indivíduos migrantes, quando chegam à metrópole, por questões socioeconômicas acabam se concentrando nas periferias e passam a constituir uma comunidade social específica, tornando-se os estigmatizados "moradores das favelas", rotulados pela sua origem, pelo lugar onde moram e pela variedade de língua que falam. É claro que por detrás desse estrato social específico podemos encontrar uma comunidade linguística rica que merece ser descrita e analisada.

> A variedade linguística que utilizam os migrantes em seus estados de origem, deixa de representar, simbolizar sua região; tal variedade, regional na origem, torna-se variedade social, símbolo de uma posição social inferior. Os migrantes vão constituir, com a população dessas cidades e de regiões próximas a elas, pertencentes ao mesmo estrato populacional, um extenso grupo de usuários de uma variedade popular ou não-padrão, estigmatizada, que se torna ela mesma um indicador da classe socioeconômica a que pertencem (RODRIGUES, 1987).

3.1.2 Perfil Social dos Informantes

Para nosso estudo, utilizamos parte do material coletado por Rodrigues (1987) correspondente a 15 inquéritos de 30 minutos cada, gravados na segunda metade da década de 1980. São 8 diálogos com mulheres e 7 com informantes do sexo masculino, de escolaridade nula ou até 4 anos de escolaridade, distribuídos conforme as três faixas etárias: 5 informantes jovens (de 20 a 35 anos); 5 informantes adultos (de 36 a 50 anos); 5 informantes idosos (com mais de 51 anos).

Devido ao grande número de ocorrências por inquérito, selecionamos 60 ocorrências da variável dependente de cada um dos 15 inquéritos, de modo que os dados coletados totalizam 900 ocorrências. Essas ocorrências foram selecionadas de modo aleatório, umas no início, outras no meio e fim dos inquéritos, para mantermos o teor do estudo por amostragem.

Como fator social, utilizamos apenas a idade, com o intuito de fazer um estudo em *tempo aparente*, de modo que fosse possível detectarmos se está ocorrendo mudança na aplicação da regra de preenchimento da casa do sujeito por pronome subjetivo.

Apesar do *corpus* estar discriminado de modo proporcional com relação ao gênero dos informantes, este fator não foi considerado relevante, pois acreditamos que o mesmo não influencia a variável em questão, o que pudemos comprovar por meio de rodadas feitas com o programa.

Os 15 informantes podem ser caracterizados como segue:

Inf.	**Gênero**	**Idade**
1	Feminino	JOVENS (20 a 35 anos)
2	Masculino	
3	Feminino	
5	Feminino	
6	Masculino	
4	Feminino	ADULTOS (36 a 50 anos)
7	Masculino	
8	Masculino	
9	Feminino	
0	Feminino	
W	Feminino	

Y	Masculino	IDOSOS (mais de 51 anos)
Z	Feminino	
K	Masculino	
H	Masculino	

Quadro 4: informantes discriminados conforme o fatoridade.

3.1.3 Normas de transcrição

Após a coleta das entrevistas, Rodrigues (1987) desenvolveu a transcrição dos inquéritos. Para isso, optou pela utilização do sistema ortográfico vigente, utilizou sinais e regras convencionais da ortografia portuguesa, efetuou desta forma uma transcrição limpa, legível, sem o acúmulo de sinais complicadores. Os aspectos de cadencia, ritmo de fala, entonação enfática, prolongamento de sílabas ou sons isolados não foram considerados nas transcrições (Rodrigues, 1987). No âmbito do Projeto Caipira, estas entrevistas foram novamente transcritas com base nas normas de transcrição estabelecidas em RODRIGUES e FERREIRA NETO, 2000.

3.1.3.1 Normas de transcrição convencional do Projeto Caipira[8]

Essas normas são sugeridas em Rodrigues e Ferreira Netto (2000), com algumas alterações e/ou acomodações exigidas pelo material coletado.

Princípio Geral

A transcrição convencional deve levar em conta a produção real do falante, buscando representar os fonemas o mais próximo possível de como efetivamente são realizados, ainda que usando os sinais e, em parte, as normas ortográficas da Língua Portuguesa. Da seguinte forma:

a) com poucas exceções (fatos de aglutinação mais frequentes), são obedecidas as normas de separação vocabular do código escrito. Trata-se de uma transcrição ortográfica com algumas acomodações.

6- Todas as normas de transcrição são retiradas na íntegra do RODRIGUES e FERREIRA NETTO (2000).

b) são fixadas realizações peculiares dos informantes que, até certo ponto, constituirão, ou poderão constituir pistas para identificação de outros traços fonéticos ou morfossintáticos de sua fala. Ex. *médio* por *médico*, *narfabeto* por *analfabeto*. Caso se torne difícil a compreensão do vocábulo transcrito, ele é representado em rodapé segundo as normas ortográficas vigentes.

Fatos Prosódicos

1. Segmentação da cadeia falada – As pausas são representadas por reticências. Assim, não são usados os sinais de pausa típicos da língua escrita: vírgula, ponto-e-vírgula, dois pontos, ponto final. Ex. *num tavaduenu nada... travessei o reberão... fui... tinha a fazenda comprendeu?*
2. Entonação – O ponto de interrogação é usado para marcar frases interrogativas e imediatamente após marcadores conversacionais realizados com curva entonacional ascendente. Ex. *né?*
3. Letras maiúsculas só são usadas para indicar ênfase, ou seja, são grafadas com maiúsculas as letras que representam a sílaba enfatizada. Ex. *inTEIra*. Conseqüentemente, nomes próprios e siglas são grafados com minúsculas.
4. Alongamento de fonemas é indicado por**:, ::** logo após o fonema alongado.
Ex. *mininu daNA:du*

Fenômenos fonéticos

1. Supressões

1.1 Redução de ditongos. Ex. *manera, carneru*
1.2 Consoantes que travam sílaba (/s/, /r/ basicamente) só são escritas quando efetivamente realizadas. Os diacríticos ´ (agudo) e ^ (circunflexo) indicam a vogal tônica de monossílabos e palavras oxítonas das sílabas abertas resultantes da não realização das consoantes de travamento e também da semi-vogal no caso dos ditongos. Exs. *falô, vendê, partí* (infinitivo), *amá, cantá*.
Obs. As palavras paroxítonas terminadas em i e u, resultantes de apócope de /s/ e /r/ não são acentuadas graficamente, ainda que a norma ortográfica exija o diacrítico. Exs. *eli, vamu*. Cf. *táxi, júri, vôlei*.

1.3 Representações de sequências aglutinadas sem apóstrofe. Exs. *dagora* (de agora), *procê* (para você), *praqui* (para aqui), *ca* (com a), *cumé* (como é), *cuns* (com uns), *dendi* (dentro de) *casa*.

2. Inserções

Ditongação. Exs. *nóis, arroiz.*

Obs.1 - Mantém-se a grafia da palavra com a vogal não ditongada.

Obs.2 - Os ditongos nasais /ãw/ e /~ey/ são representados segundo a norma ortográfica vigente. Exs. *vendem, cantam, elis têm,* falarão.

3. Alternâncias

3.1 Palatal /lh/ é representada por *i* quando realizada como semivogal. Ex. *paia* (palha).

3.2 Consoante /l/ é grafada *r* em grupos consonantais ou travando sílaba se realizada como vibrante. Ex. *crasse* (classe), *vorta* (volta).

3.3 As vogais átonas /e/, /o/, quando realizadas [i], [u], são grafadas com as letras *i* e *u* respectivamente. Exs. *denti, dimanera, cum, mininu, discurpa, disfiava.*

OUTROS SINAIS

( )	Incompreensão de palavras ou segmentos;
(*muitumilhó*)	Hipótese do que se ouviu;
i comé/ i inicia	Truncamento de palavras (Havendo homografia, usa-se acento indicativo da tônica e/ou timbre);
i eu memu TInha	Entoação enfática das sílabas com maiúsculas;
in-te-li-gen-ti	Silabação;
((minúsculas))	Comentários descritivos do transcritor;
[	Sobreposição de vozes, quando um segundo interlocutor interrompe o que o primeiro estava falando;
“ ”	Citações literais, reproduções de discurso direto ou leituras de textos, durante a gravação;
Fáticos	ah, éh, eh, ahn, ehn,uhn;
Inf (sem ponto)	Informante;

Doc (sem ponto) Documentadora.

3.2 Coleta de Dados

Para a coleta dos dados, consideramos apenas a variável presente na fala do informante. Percebemos que a variável analisada é muito recorrente na variedade do Português Popular, assim cerca de 2 a 3 páginas de cada inquérito foram suficientes para a coleta das ocorrências.

Consideramos sentenças longas, ou seja, coletamos dados do discurso corrente, do texto por inteiro, pois analisamos fatores correlacionados com a *referência*, como: Status Informacional do Sujeito e Paralelismo.

A seguir há um fragmento do *corpus* que ilustra como ocorreu a coleta das ocorrências:

Doc *[i agora nu fim di anu? que qui cêis pretendem fazê? é diferenti di... di*

Inf *fim di anu pra mim é a mesma coisa u únicu fim di anu qui* ***eu*** *passei gostosu foi quandu* ***eu*** *tava trabalhandu im casa di familha... qui* ***eu*** *passei* ***Ø*** *vi comu ( ) Ø vi comu qui era a festa delis... mais* ***eu*** *quiria sabê nu fim di anu comu é qui é u tal di reveion qui a turma fala...* ***Ø*** *nem sei qui qui é... já* ***Ø*** *ovi falá qui qui é reveion mais* ***Ø*** *num sei u qui qui é essa festa... árvori di natal nunca* ***Ø*** *montei...* ***Ø*** *num sei comu é qui é... nunca* ***Ø*** *ganhei presenti di ninguém: di natal... im vinti i um anu nunca* ***Ø*** *ganhei... só* ***Ø*** *dei... pa minha mãe pa minha vó i pu meu pai... só qui* ***eu*** *dei...* (I. 4-4a, mulher, 21 anos)

Doc *(...) da da sua terra né? da iscola? comu é qui era mesmu?*

Inf *purque a iscola lá ...a professora tinha qui... puxá mais du alunu... du que:... aprendê qui nem aqui* ***Ø*** *aprendi aqui* ***Ø*** *vai pra recuperação... lá não existi issu ((ruídos)) qué dizê passava quem sabia.... num tinha ( ) qui nem aqui... aqui é si si* ***elis*** *sabi* ***Ø*** *passa si* ***Ø*** *não sabi* ***Ø*** *passa du memu jeitu... qui u vizinhu aí tem uma: mocinha na quinta séri ma* ***Ø*** *num sabi nada... quem* ***eu*** *perguntei pra ela quem discobriu a américa* ***ela*** *dissi qui num sabia... qué dizê qui lá* ***elis*** *puxa... qué dizê quem sabi fica quem num sabi fica fazenu u quê na iscola né?* ***elis*** *ti:ram fo:ra i* ***Ø*** *põe otrus que pricisam aprendê... qui lá* ***a gente*** *andava uma basi di:... cincu a seis quilômetru pa í pa iscola i*

*quandu **Ø** voltava **Ø** ia pra roça... puxá inxada né?... qué dizê qui **Ø** num tinha a mordomia qui tem aqui purque aqui*

Doc *[qui qui é lá?*

Inf *lá nu era nu na roça né? **Ø** (a gente) trabaiava na roça... qué dizê qui tinha um sítiu di cincu alqueris... qui ( ) café tinha pomar vaca... tudu issu **Ø** tinha qui cuidá né?... qué dizê qui **Ø** tinha qui istudava di manhã... oitu hora... **Ø** saía às onzi... dipois **Ø** trabalhava à tardi na roça... qué dizê **Ø** ficava carpindu café... otra hora trabalhanu nu pastu roçanu... qué dizê qui quandu num tinha sirviçu pra **nóis** dentru dali du sítiu... tinha pocu... **Ø** prucurava trabalhá pur dia diária fo:ra... ma memu serviçu carpindu qué dizê ( ) ocupação melhó... minha mãe num num sabi lê né? meu pai também mal e mal ( ) u nomi... qué dizê que num tinha...* (I E- 11a, homem, 31 anos)

CAPÍTULO 4: Seleção e Caracterização dos Dados

4.1 Critérios de seleção de dados

1- São consideradas somente as construções frasais com verbo explicitado, por ser a regra variável sob análise o preenchimento do sujeito por pronome. Assim, as construções consideradas podem ser duas: PRONOME + VERBO; ZERO + VERBO.

Ex.: ***eu*** *gostu di todu mundu* (pronome + verbo) (I.6-5a)

Ø *Gostu mais de música romântica* (zero + verbo) (I.6-5a)

2- Quando os verbos são repetidos em sequência, retomando o que foi dito, mas com uma outra construção, por exemplo sujeito expresso seguido de zero, consideramos as duas ocorrências como dois dados distintos.

Ex.: *daí* ***eu*** *tenho* ***Ø*** *tenho pastadu cum essis mininu...* (I.6-5a)

3- Pronomes repetidos relacionados a uma única forma verbal, numa mesma oração, foram considerados como um único dado.

Ex.: ***a gente, a gente*** *trazia da fazenda* (I.6-5a)

eu *uma veiz* ***eu*** *fui...* (I.6-5a)

4- Quando se repete a oração toda sequencialmente, tanto a primeira quanto a segunda ocorrências foram consideradas como dado dois dados distintos.

Ex.: ***eu num sô di saí muitu*** *(mãe interfere)* ***eu num sô di saí muitu****...* (I.6-5a)

5- A expressão a gente foi considerada pronome pessoal.

Ex.: ***a gente*** *fazia um monti de biscoitu... i* ***Ø*** *amarrava um nu otru...* (I.6-5a)

6- O mesmo referente explicitado pelo pronome pessoal subjetivo da primeira ocorrência foi considerado para os sujeitos zeros subsequentes, mesmo estando estes acompanhados de formas verbais não-marcadas em pessoa e número, na medida em que o verbo não concorda com o sujeito. É o que acontece no exemplo que segue com sujeito pronominal *eles.*

Ex.: Inf ***elis****...* **(informante referindo-se aos filhos)** *vêm* ***Ø*** *chega a veiz* ***Ø*** *num qué nem a a ...jantá otra hora* ***Ø*** *num qué nem armuçá purque* ***Ø*** *chega tudu cas barriguinha cheia né? (I.B- 8b)*

4.2 Critérios de exclusão de dados

Foram descartadas as ocorrências abaixo listadas.

1- Nas expressões: *vô ti falá, tá locu, qué dizê* a ausência de pronome sujeito foi considerada regra categórica.

Exs.: Ø ***vô ti falá****... era uma briga danada* (I.6-5a)

eu morei lá faiz novi anu... ( ) Ø divertia i tudu... ***qué dizê****... atravéis da minha pessoa* (I. A-23a)

2- Com formas verbais com função de marcador discursivo, como: *entendi? entendeu? compreende?,* a ausência de pronome é regra categórica.

Exs.: *im qui sintido? purque:: praticamenti quarqué coisa elisfala qui nóis somu FAveladu... Ø* ***intendeu?*** (I.B -23b)

lá já num i/ já num tem essi negóciu... tem mais várius.... é várius assim num é dizê assim.... qui tem assim di cê andá assustada Ø ***comprendi****?* (I.B-23b)

3- Sujeito composto de substantivo e pronome.

Exs.: ***eu ela i u maridu*** *dela... dançamu na casa...* (I.6-5a)

eu mais a minha colega *fomu...* (I.6-5a)

4- O primeiro de uma sequência de dois verbos quando o primeiro é desprovido de sentido pleno.

Exs.: *eu* ***pegava*** *minina* ***vistia*** (I.6-5a)

...eu ***pegava chegava*** *nu terreno* (I.6-5a)

...ela ***pegava*** *minina* ***dava*** *cada chifrada...* (I.6-5a)

5- Pronome subjetivo associado a verbo não explícito, ainda que recuperável no contexto.

Ex.: *quando* ***eli*** *saiu de casa essis mininu* ***era*** *tudu piquenu né?... i duenti e* ***eu*** *também duenti.* (I.6-5a)

6- O pronome relativo que foi considerado um conector, ou seja, não exercendo a função de sujeito. Ocorrências como a seguir não foram consideradas.

Ex.: qué dizê qui era uma vida: meia difícil ***pra genti... qui*** *num tinha: tinha trabalhu* (I.E-11a)

7- Não consideramos ocorrências com verbo no infinitivo.

Ex.: Ø tenhu... mais eu sô u mais velhu... que dizê qui num tinha ondi ***procurá*** *a sabê di vizinhu qui meu pai num consentia issu... nóis saí di casa pra* ***istudá****...* (I. E-11a)

4.3 Realização/ não-realização dos pronomes subjetivos: a variável

A variável dependente é binária, ou seja, é constituída por duas variantes: realização/ não-realização dos pronomes subjetivos de todas as pessoas, do singular e plural, utilizadas pelos falantes do Português Popular Brasileiro em São Paulo. Foi escolhida por se tratar de um fenômeno recorrente na língua falada, critério utilizado por Labov (2008: 26) *"queremos um item que seja frequente, que ocorra tão reiteradamente no curso da conversação natural espontânea que seu comportamento possa ser mapeado a partir de contextos não-estruturados e de entrevistas curtas."*

Partindo da hipótese de que o uso do pronome subjetivo não é aleatório, a sua omissão não é categórica, nem sua explicitação, mas podem ser motivadas por fatores linguísticos e sociais, utilizamos a metodologia da Sociolinguística Quantitativa para estabelecer um grupo de fatores de natureza linguística e social que condicionariam a variável dependente. Tais fatores, bem como as hipóteses desta pesquisa, só puderam ser definidos tendo como base as leituras de estudos sobre o preenchimento do sujeito, citados no capítulo 2 deste trabalho, e sobre a Gramática Funcional de Halliday (1994).

A seguir discorremos sobre os fatores linguísticos e o fator social, considerados variáveis independentes, que estão apresentadas no quadro abaixo.

4.3.1 Variáveis Independentes

Fatores Linguísticos	**Categorias**
1) Pronome pessoal subjetivo	**1**eu **2** você **3** ele, ela **4** nós **5** vocês **6** eles, elas **7** a gente
2) Forma verbal	**A** Forma verbal marcada **B** Forma verbal não-marcada
3) Paralelismo formal	**X** ocorrência precedida de zero pronominal, referência igual **Y** ocorrência precedida de pronome expresso, referência igual **Z** pronome/ zero isolado ou primeira referência
4) Polaridade	**a** sentença afirmativa **n** sentença negativa
5) Status informacional	**N** informação nova **D** informação dada
6) Indeterminação do referente	**d** referente determinado **i** referente indeterminado
7) Valor Semântico do verbo	**A** verbo de ação **P** verbo de processo **B** verbo de ação-processo **E** verbo de estado
Fator social	**Categorias**
8) Idade do Informante	**J** adultos jovens de 20 a 35 anos **M** adultos de meia idade 36 a 50 anos **V** idosos com mais de 51 anos

Quadro 5: Fatores Linguísticos e Social.

4.3.1.1 PRONOME PESSOAL SUBJETIVO

Como foi noticiado no capítulo 2, o fator pronome pessoal subjetivo foi estabelecido tendo como base o estudo de Silva (1993), em que a autora considerou a taxa de omissão dos pronomes subjetivos das três pessoas discursivas e constatou que a 1ª pessoa é a mais omitida, tendo uma taxa de 77%. Com a segunda pessoa o pronome sujeito é mais realizado, tendo uma taxa de omissão de 30% e com a terceira pessoa a taxa de omissão do pronome é de 50%.

Duarte (1993) constatou, no seu estudo diacrônico, que os pronomes sujeitos de 1ª e 2ª pessoas estão sendo cada vez mais realizados, enquanto os de 3ª pessoas, mais omitidos.

Resultados similares são encontrados em outros estudos, como o de Ilari *et al*. (1996:86), no qual é possível observar que os pronomes de primeira pessoa são os mais frequentemente realizados (75%), relacionados ao tipo de inquérito e à cidade do falante. Esse resultado vem sendo justificado pelo fato de nos inquéritos os informantes narrarem experiências pessoais e serem sempre instigados a falar sobre suas preferências, reminiscências, julgamentos, o que os levam a falar em primeira pessoa.

Assim, com este fator pretendemos obter uma visão geral sobre os pronomes subjetivos utilizados na variedade do Português Popular falada na cidade de São Paulo, de modo que seja possível identificarmos qual pronome está sendo mais realizado e qual está sendo mais omitido. Pretende-se também responder a seguinte questão: haveria influência da pessoa do discurso na realização/ não-realização do pronome?

O Fator compõe-se das seguintes categorias:

Categorias	**Exemplos**
1 eu	*Intão **eu** fui trabalhá pur pur dia im casa di família...*
2 você	*tem qui chamá di piá ((ri)) lá si **cê** falá chegá lá nu paraná...*
3 ele, ela	*Quandu **eli** saiu di casa*
4 nós	***nóis** levava biscoitu PA jogá pru spexi...*
5 vocês	***Cêis** tão namorandu i...*
6 eles, elas	*... qui **elis** tirava u leiti i Ø num lavava u peitu da vaca*
7 a gente	***A genti** tem qui fazê num dia só né?*

Quadro 6: Categorias referentes ao fator Pronome Pessoal Subjetivo.

4.3.1.2 FORMA VERBAL E AMBIGUIDADE MORFOLÓGICA

O fator forma verbal foi estipulado com base no estudo de Silva (1993), no qual a autora ressalta que no Português Culto há constante ambiguidade morfológica entre a 1ª e 3ª

pessoas do singular nos tempos: pretérito imperfeito (eu/ ele cantava) e pretérito mais-que-perfeito (eu/ele tinha cantado, cantara). No caso da 2ª e 3ª pessoas do singular, Silva atenta para a frequente ambiguidade morfológica, porque o pronome você (2ª pessoa do singular) exige a forma verbal de 3ª pessoa do singular.

Assim, podemos dizer que tais formas verbais são caracterizadas neste trabalho como **não-marcadas**, ou seja, do ponto de vista da funcionalidade não há marcas suficientes no verbo que distingam a pessoa do discurso, sendo necessário o pronome subjetivo expresso para determinar o referente.

Com uma análise prévia do *corpus*, pudemos verificar que além das formas verbais não-marcadas típicas da variedade padrão, há ainda certa simplificação no sistema flexional dos verbos na variedade Português Popular. Vejamos a seguir:

(1) ela tocava lavora (I.E- 11a)

(2) que nói sempri trabaiava na roça... (I.E- 11a)

(3) agora us otru tudu tem istudu... dificí na roça quandu nóis morava lá... a (I.E- 11a)

(4) tem elis insina trabalhu di... artesanatu di pintura di:... (I.E- 11a)

Nos exemplos acima, podemos notar que os verbos se caracterizam como **formas verbais não-marcadas**, pois elas não são suficientes para identificar o referente, ou seja, se retirarmos o pronome acarretará ambiguidade de referente, assim o pronome sujeito se faz necessário para a transmissão eficiente da mensagem.

Então temos as **formas verbais não-marcadas**:

*(5) que **nói** sempri **trabaiava** na roça...* (I.E- 11a)

*(6) agora us otru tudu tem istudu... dificí na roça quandu **nóis morava** lá...* (I.E- 11a)

As **formas verbais marcadas** trazem informação do número e da pessoa do verbo em sua desinência, não sendo morfologicamente ambíguas, de modo que se retirarmos o pronome sujeito saberemos perfeitamente qual é o sujeito da oração. Como no exemplo abaixo:

*(7) ...intão **nóis tamu** aqui lutanu... (Forma verbal marcada)* (I. 4- 4a)

Também foram consideradas **formas verbais marcadas** verbos com os morfemas de flexão número-pessoal com alterações fônicas, como nos casos em que o ditongo nasal /ãw/ se

reduz a vogal desnasalada /u/, sendo ainda possível depreender a informação de número e pessoa. Ex.: *falaram >falaru; comeram>comeru, etc*

*(8) Elas **gozaru** tanto dele...* (I.0 -2a)

Há ainda casos em que o falante do português popular utiliza formas verbais que são marcadas, mas que não estabelecem concordância com o sujeito, como no exemplo a seguir:

*(9) depois era muitu difici pagá aluguel... aí **nóis resolveu** í pa são paulu purque...* (I.4- 4a)

O exemplo acima, apesar de ser **forma verbal marcada**, o enquadraremos no fator **forma verbal não-marcada**, pois se retirarmos o pronome, apenas com a desinência verbal não ficará claro qual é o referente correto a que o falante pretende referir-se.

Com base no trabalho de Silva (1993) e análise do *corpus*, consideramos para este fator todas as pessoas do singular e plural, independentemente do tempo verbal. As categorias que compõem este fator são:

Categorias	Exemplos
A Forma verbal marcada	*...depois eu **fui** lá busquei...*
Não-ambígua morfologicamente	*... nóis **somu** FAveladu...*
B Forma verbal não-marcada	*... eu **falava**...*
Ambígua morfologicamente	*... nóis **morava**... /nóis **falô**.../ nóis r**esolveu**...*

Quadro 7: Categorias referentes ao fator Forma Verbal.

Podemos presumir então que se a forma verbal não é suficiente para indicar a pessoa do discurso, o pronome sujeito se faz mais necessário.

Foi o que Silva (1993) constatou em seu trabalho com o pronome de primeira pessoa do singular: este é mais omitido quando não há risco de interpretação equivocada de referente, ou seja, quando a **forma verbal** é **marcada**, sendo a desinência verbal suficiente para identificar o referente. A ambiguidade morfológica com a primeira pessoa existe geralmente quando não há na desinência verbal a marca desambiguadora.

Constatamos que o falante do Português Popular utiliza poucas marcas no verbo, independentemente da pessoa discursiva e do tempo verbal, e nesses casos o pronome sujeito

geralmente aparece preenchido: *"eu **brigava**, eli **tá**, a gente **sai**..., elis **dá** instrução né?..., elis **manda**... **passá** nas impresa né?..., elis **luta** cum um cum otru..."*

Assim, estabelecemos a seguinte hipótese funcional: as formas verbais marcadas em pessoa e número, por não apresentarem ambiguidade morfológica, favoreceriam a omissão do pronome sujeito; já as formas verbais não-marcadas em pessoa e número (*"eu / ele/ nóis falava*) favoreceriam a presença do pronome.

4.3.1.3 PARALELISMO FORMAL

O fator paralelismo formal foi estabelecido tendo como base os estudos de Scherre & Naro (1993), com algumas adaptações. Os autores analisam o paralelismo formal no nível oracional (marcas no sujeito) e no nível discursivo (marcas no verbo). Em nossa pesquisa, estamos analisando apenas o paralelismo no nível discursivo, adaptando as categorias para a variável realização/ não realização do pronome subjetivo.

Os autores conceituam o paralelismo formal como sendo a tendência de marcas levarem a marcas e zeros levarem a zeros (SCHERRE & NARO, 1993: 2).

De acordo com Omena & Braga (1996), uma vez estabelecida a forma, o falante tende a repeti-la, principalmente se o referente não mudar.

Com este fator, pretendemos responder a seguinte questão: o paralelismo formal poderia influenciar a variável realização/ não-realização do pronome sujeito? Ou seja, uma ocorrência de pronome subjetivo levaria a outra ocorrência de pronome subjetivo, e zeros (omissão do pronome subjetivo) levariam a zeros (omissão do pronome subjetivo)?

Para responder a essa questão, levamos em conta as três categorias, elaboradas a partir dos estudos de Scherre & Naro (1993) e Omena & Braga (1996):

X ocorrência precedida de pronome expresso, referência igual.
Y ocorrência precedida de zero pronominal, referência igual.
Z pronome/ zero isolado, primeira referência ou referência diferente.

Consideramos as categorias (X) e (Y) para a ocorrência (pronome/ zero) com a mesma referência que o pronome sujeito da construção anterior (mesma referência do discurso imediatamente anterior). Consideramos (Z) para a ocorrência de primeira referência, *"aquela que não sofre influência de outra que a antecede"* (OMENA, 1996: 72), ou quando há mudança de referência ou no caso de a ocorrência ser um pronome ou um zero isolado.

Consideramos o paralelismo formal apenas na fala do informante, mesmo sabendo que ele pode atuar entre turnos do documentador e do informante, ou quando há mudança de turno, em que o entrevistador faz a pergunta e o informante retoma o pronome na sua resposta.

Também consideramos como ocorrência a construção que não está separada da construção anterior por mais de 10 orações, conforme Scherre e Naro (1993: 9).

A seguir apresentamos fragmentos do *corpus* com exemplos das três categorias estipuladas (X), (Y) e (Z).

Doc. conta comu é qui era sua vida lá im minas

*(10) Inf. lá im minas é:... **eu (Z)** morava cum a minha irmã **eu (X)** chorava dia i noiti pa vim imbora pa são pa/ pa... aí minha irmã... minha irmã num dexava **eu (X)** ficava só choranu dibaxu dus pé di arvri ((ri)) era a maior comédia...* (I. 6 - 5a)

Em (10), a primeira ocorrência do pronome: *"**eu** (Z) morava..."*é caracterizada como primeira referência, ou seja, a primeira escolha da forma, que não sofre influência de outra que a antecede, lembrando que o paralelismo mede a repetição, no caso repetição de marcas (pronome) ou zeros. Nesta categoria (Z), não é possível medir se o paralelismo está atuando, pois a primeira referência por caracterizar o início, não está precedida de zero nem de pronome, ela é primeiro ponto em que a referência é estabelecida. A segunda ocorrência: *" **eu** (X) chorava dia..."* é precedida por pronome expresso, sendo que se trata da mesma referência do pronome anterior.

Observemos o próximo exemplo.

Doc. comu chamava?

*(11) Inf. é essa cidadi nova modica... aí... **a genti (Z)** fazia um monti di biscoitu assim uns dois treis sacu... i **Ø (X)** amarrava um nu otru **Ø (Y)** punha na na nas costa Du jegui ((risos)) u maió sarru.... depois u <u>jegui</u> ispirraçava i... **Ø (Z)** <u>jogava</u> aquilu tudu nu chão ((fala paralela))* (I. 6-5a).

Em (11), *"**a genti** (Z) fazia..."* trata-se da primeira referência, em que a forma é estabelecida. Logo, temos a ocorrência: *"...i **Ø** (X) amarrava..."* , que é precedida por pronome com a mesma referência. Em seguida temos: *"... **Ø** (Y) punha na na nas costa..."*,

uma ocorrência precedida de zero com a mesma referência. Porém, na ocorrência: *"... **Ø** (Z) jogava aquilu..."* temos uma ocorrência de zero em que há mudança de referência, pois nesse caso o verbo jogava remete a uma nova referência no discurso, o jegui, ou seja, um referente diferente.

Para este fator, **nós** e **a gente** foram considerados como correlativos, ou seja, com o mesmo referente.

*(12) **nóis** viemu di ônibus né? i... u ônibus quebrô nu caminhu... i era aqueli ônibus quandu **a genti** quiria fazê necessidade...* (I.4-4a)

4.3.1.4 POLARIDADE

Este fator foi sugerido pelo estudo de Neves (2007). Conforme a autora, estabelecemos a hipótese de que a sentença afirmativa favoreceria o uso do pronome pessoal explícito, enquanto a negativa favoreceria o uso do pronome zero (Ø).

Foram consideradas as categorias que seguem:

Categorias	Exemplos:
a sentença afirmativa	*"...i Ø gostava di fazê muitu era biscoitu..."*
n sentença negativa	*"...a genti **num** gostava di sapatu né?..."*

Quadro 8: Categorias referentes ao fator Polaridade.

Os advérbios: *nunca; num; nem,...* foram considerados como caracterizadores de sentença negativa.

(13) ***nunca** fiquei di recuperação tamém...* (I.4- 4a)

(14) ***nem** sei qui qui é... já ovi fala qui qui é reveion mais **num** sei u quiqui é essa festa...* (I.4- 4a).

4.3.1.5 STATUS INFORMACIONAL DO SUJEITO

O fator status informacional do sujeito é, neste trabalho, considerado o fator mais relevante. Em primeiro lugar, por contemplar as duas hipóteses a seguir.

1ª hipótese: *"A compensação funcional"* estaria atuando no *corpus*, ou seja, o sujeito explícito tenderia a ocorrer mais em contextos ambíguos, quando este vincula informação nova no discurso, e menos quando sua realização não se fizer necessária para a compreensão da mensagem, ou seja, quando a informação é dada;

2ª hipótese: o pronome sujeito explícito ocorreria com maior frequência quando houvesse a alternância de referente (mudança de tópico discursivo), ou seja, com menos conexão discursiva; por outro lado o pronome sujeito ocorreria com menos frequência quando o referente fosse mantido, em contextos com mais conexão discursiva (Silva, 1993).

Em segundo lugar, é por meio dele que conseguimos analisar a ambiguidade textual presente no *corpus*, ou seja, verificar os elementos referentes em competição no contexto discursivo e situacional.

De acordo com Halliday (1994: 296), a unidade de informação é a estrutura composta por dois elementos que exercem as respectivas funções: um é o *Novo*, aquilo que o ouvinte interpreta como imprevisível, inesperado, novo. O outro é o *Dado*, o que é apresentado como previsível, familiar, já conhecido pelo ouvinte.

Para o autor, o modo como a estrutura informacional é realizada é natural (não-arbitrário), em dois sentidos: (i) o Novo é marcado por proeminência tônica; (ii) o Dado tipicamente precede o Novo.

É comum que o Dado seja um elemento fórico, referindo-se a algo que já foi mencionado no contexto verbal ou não-verbal. O Dado pode sofrer elipse, ou seja, não estar presente na estrutura gramatical. Assim, Halliday afirma "que a unidade de informação consiste de um elemento Novo obrigatório, mais um elemento Dado opcional".

O ***Dado*** é o elemento recuperável por já ter sido mencionado antes, mas esta não é a única possibilidade. Pode ser também algo que está presente na situação, como *eu* e *você*; ou no ar, ou alguma coisa que não está presente, mas o falante apresenta como *Dado* por questões retóricas, por exemplo. O ouvinte tem de interpretar a informação Dada como sabendo que ela não é uma novidade. Do mesmo modo, o ***Novo*** além de ser tratado como não recuperável, algo que não foi mencionado; também pode ser algo inesperado, previamente mencionado ou não. O ouvinte tem de fazer a seguinte interpretação: preste atenção para isso, pois é algo novo.[9]

[9]Todas as informações sobre Dado e Novo foram retiradas de Halliday 1994: 296 - 298, sendo alguns parágrafos traduzidos da obra original do autor.

Halliday utiliza dois critérios para determinar o status informacional: (i) considera o contexto linguístico e situacional; (ii) considera a proeminência tônica que caracteriza o elemento novo.

Para nosso trabalho, propomos um teste, adaptado do estudo de Silva (1993), que se mostra válido para determinar o status informacional do pronome sujeito: consiste em retirar o pronome sujeito da oração e examinar a existência de outro candidato a sujeito da frase. Se a referência for mantida através de anáfora ou exófora (situacional) o pronome sujeito será Dado, por outro lado se houver ambiguidade entre candidatos a referente, o pronome sujeito será Novo.

Porém, sabemos que o referente do sujeito Novo pode ser explicitado através da flexão verbal número-pessoal e/ou do sujeito explícito, no caso pronome subjetivo (PEREIRA, 2004). Por isso, o teste proposto por Silva é restrito apenas para casos de forma verbal não-marcada, ou seja, formas em que a flexão do verbo não é suficiente para identificar o referente. Como no exemplo:

(15) ***nóis*** *marcava u nomi dipois* ***ela*** *voltava* ***nóis*** *apagava aí* ***ela*** *perguntava "quem feiz bagunça?"... **nóis** falava "ninguém tudu fico quetinhu"...* (I.4- 4a)

No exemplo acima, todas as ocorrências de pronome veiculam informação Nova, pois se retirarmos um dos pronomes acarretará uma possível ambiguidade de referentes, já que as formas verbais não são marcadas.

Em casos de redundância, em que o falante utiliza o pronome mais o verbo flexionado em número e pessoa, o teste de Silva não é válido, porque, mesmo retirando o pronome, a informação do referente pode ser depreendida através da flexão verbal. Como no exemplo:

(16) *elis também isqueci di pegá u papelzim pra mim né?... i* ***eu achu*** (Novo) *qui::* (I.B-8b)

Assim, um segundo teste, proposto por Pereira (2004: 90) para identificar se o referente veicula informação Nova/ Dada, pode ser utilizado em casos de ambiguidade. Este consiste em retirar tanto o pronome subjetivo, quanto a flexão do verbo da mesma ocorrência. Se mesmo assim for possível identificar o referente do verbo, este constitui informação dada, observando-se o discurso como um todo e ao mesmo tempo a unidade discursiva analisada.

(17) ***nóis*** (Novo) *ia PA iscola tamém* ***Ø ia*** (Dado) *vê as coisa lá i ficá cum vontadi di comê... (qui sempri tem) i* ***a genti*** (Dado) *fica com vontadi di comê intão:...* (I.4-4a)

No trecho acima, mesmo se retirarmos o pronome *a genti,* será possível identificar o referente do verbo, sendo que o verbo já não traz forma de flexão verbal desambiguadora, constituindo informação dada. Vale ressaltar que neste trabalho estamos considerando os pronomes *nós* e *a gente* como correferentes.

Assim, consideraram-se os subfatores:

D- Dado

N- Novo

Consideramos ***Dado*** quando o referente não é ambíguo, pois é mantido ou no discurso imediatamente anterior ou na situação, e quando não há alternância de tópico discursivo. Consideramos ***Novo*** quando o referente é interpretado como sendo ambíguo, pois há referentes em competição no contexto da ocorrência, ou quando há alternância de tópico discursivo, medido pelo discurso imediatamente anterior. Se há alternância de tópico será informação nova, senão será dada.

Utilizamos as ideias de Pereira (2004) e as categorias estipuladas pela autora para compor o fator Status Informacional do Sujeito. Vejamos abaixo alguns casos de ocorrência de Dado e Novo:

Dado

A informação dada ocorre nos seguintes contextos:

1) quando mencionada no discurso anterior próximo, como em (18):

(18) *eu num sô di saí muitu ((mãe interfere))* ***eu*** *num sô di saí muitu né?* (I. 6- 5a)

2) Do ponto de vista limitado ao status informacional do sujeito, consideramos *nós* como correferente de *a gente*. Assim, se *nós* estiver como antecedente de *a gente* ou vice-versa, referindo-se ao mesmo referente, será *Dado*. Utilizamos esse critério por estarmos estudando o preenchimento do sujeito por pronome, sendo que não nos atentamos à concordância verbal, para não expandir o objeto de estudo.

(19) Doc. ((pausa)) *CE falô qui... qui ia na rua... qui qui era a rua?*

Inf. *a rua qui **a genti** falava era u interi/ era...**nóis** falava cidadi... lá é rua... purque tem a fazenda né?* (I. 6- 5a)

3) A informação será considerada como *dada* quando o pronome/zero retomar o SN já mencionado, seja o verbo flexionado (20) ou não-flexionado (21).

(20) Inf ...*nóis quebrô um a cabeça duotru mais Ø num **quebramu** televisão nenhuma só qui u cara mandônóisimbora... certu?* (I.D-24B)

(21) *elis... vêm... Ø **chega**... a veiz Ø num **qué** nem a a ...jantá otra hora Ø num **qué** nem armuçá* (I.B-8b)

4) A informação será considerada como dada, quando retomar o referente do sintagma nominal mencionado no discurso imediatamente anterior:

(22) *depois u jegui ispirraçava i...Ø **jogava** aquilu tudu nu chão...* (I.6 - 5a)

(23) *qui u vizinhu aí tem uma: mocinha na quinta séri ma Ø num **sabi** nada..* (I.E-11a)

5) Ocorrências em que a informação do referente é transmitida por pronome explícito na primeira oração e mantida nas orações subsequentes, por não serem redundantes, levam, conforme Pereira (2004), a uma sequência de verbos não-marcados com zero na posição do sujeito, os quais constituem informação *dada*.

(24) *papai mudava pra lá i pra cá... uma hora **Ø tava** trabalhanu (na fibra)... ota hora Ø **tava** lá pru ladu di... rolândia...* (I. E-11a)

5.1 Casos em que há o referente expresso na primeira oração e mantido nas orações subsequentes, sendo que o falante reformula sua fala.

(25) *aí **elis** (N) veiu i Ø (D) **feiz** a Ø (D) **vieru** e Ø (D) **fizeru** a ligação... mais Ø (D) **fizeru** a ligação assim... um um uma ligação PA cincu seis pessoa sabi?* (I.D-10b)

No material estudado por Pereira (2004) foi frequente a sequência: *Nós + verbo no singular ou no plural, Zero + verbo no singular, Zero + verbo no singular.* Sequência que pode ser interpretada como resultado da atuação do paralelismo discursivo (Scherre & Naro,

1993), contudo para Pereira e Rodrigues (2008: 10), a sequência explicitada acima exprime a atuação de um princípio de natureza funcional: "um sujeito que constitui informação dada leva a uma sequência de verbos não-marcados com casa vazia do sujeito, evitando uma redundância" (PEREIRA & RODRIGUES, 2008: 10).

6) Ocorrências em que o zero/ pronome sujeito retoma um pronome (com função de complemento) anteposto próximo, sendo o referente recuperável pelo discurso imediatamente anterior, são consideradas como informação d*ada*. No exemplo abaixo, podemos verificar que o **zero** retoma o referente do pronome mim.

*(26) dia... intão eu fui trabalha pur pur dia im casa di família... dipois mamãe veiu pra cá PA ficá cum elis pra mim **Ø cumecei** trabalha im firma... Ø trabalhei dois anus i dois mesi na tricolã.* (I.6-5a)

Novo

Conforme Pereira (2004), a informação nova ocorre nos contextos que seguem.

1) Pronome subjetivo catafórico, ou seja, quando o referente é mencionado posteriormente à ocorrência.

(27) *...**elis** gostava di comê ropa as vaca...* (I.6-5a)

(28) ***a genti** foi num...na casa di um parenti deli i... foi eu i eli.... i a... i u... i a... subrinha... i u subrinhu deli qui ti/qui já era noivu, né?...* (I.2-2a)

2) Pronome subjetivo exofórico: o referente está presente no contexto situacional. O falante introduz referentes da situação discursiva, como por exemplo: ele, o documentador ou as pessoas da situação.

(29) *a genti foi só sufrimentu né?... **cê** sabi sufrimentu DA VIDA né?...* (I. B- 8b)

3) Pronome subjetivo anafórico: o pronome/ zero mais o verbo na forma marcada remetem aos dois referentes anteriores, que não tinham sido combinados em um único sujeito, ou seja, articulados anteriormente em sujeito composto.

(30) *eu minha irmã **nóis** brincava pa caramba...* (I. 6- 5a)

4) Referente implícito: o pronome/ zero mais o verbo tem um referente implícito, mas não mencionado no discurso. Classifica-se assim o referente do verbo como informação nova. No exemplo abaixo o referente de "***fizeru"*** está implícito na situação, pois sugere serem as pessoas da comunidade que fizeram o abaixo assinado.

(31) *...uns cincu meis seis meis pareci qui a genti tava aqui aí elis veiu i Ø feiz a Ø vieru e Ø fizeru a ligação... mais Ø fizeru a ligação assim... um um uma ligação pa cincu seis pessoa sabi? intão era a maior cunfusão... intão quandu chegava a conta di água uns quiria pa/ uns dava dinheru na data certa otrus num dava... intão fazia aquela bagunça sabi?... aí até qui Ø* ***fizeru*** *um abaxu assinadu aí...* (I. D- 10b)

5) Casos de ambiguidade de referentes: o pronome/ Ø + verbo flexionado tem um referente expresso anteriormente, mas o pronome ou o verbo flexionado se faz necessário para não ocorrer ambiguidade de referentes, isso ocorre nos seguintes casos:

5.1 Mudança de tópico discursivo: há mudança de referente entre uma e outra ocorrência do pronome. Essa alternância é medida pelo discurso imediatamente anterior.

(32) *eu ia juntu né? Mais eli num acompanho nóis não... tinha assim era aquela ruazinha intão* ***a genti*** *ia assim aqueli... aqueli povuadu andandu né?* (I.2-2a)

Em (32), há alternância de tópico discursivo e se retirarmos o pronome a genti, por o verbo não ter marca de flexão desambiguadora, acarretará uma possível ambiguidade na interpretação do referente.

5.1.2 Mudança de tópico discursivo na flexão: há alternância de referentes entre uma e outra ocorrência, assim o referente, com informação nova, pode ser expresso por Ø mais o verbo flexionado.

(33) *nunca Ø dancei uma festa junina... tempu di festa na iscola nóis nunca vinhemu ...* ***Ø num possu*** *nem falá comu qui era qui meu pai nunca dexô eu í... nunca mesmu* (I.4-4a)

Em (33), há alternância de tópico discursivo, expressada pela flexão.

5.2 Dois referentes para um verbo: há dois possíveis candidatos a referente de um verbo. Assim a ausência do pronome acarretaria ambiguidade.

(34) *tá locu... aí tinha uma so/ uma:... cunhada da minha irmã... ela gostava muitu di namorá né?... namorá éh nus: camim... camim di roça... ela era mocinha já eu era piquena... ih quandu **eu** chegava im casa...*(I. 6- 5a)

5.3 Casos em que a ausência do sujeito ou de marca na flexão verbal acarretará numa interpretação de sujeito indeterminado. Assim, o falante coloca o sujeito como informação nova, sendo explicitada ou pelo uso do pronome ou pela flexão verbal.

(35) Inf *... foi uma zuera... aí **nóis** num quiria falá quem foi cum medu du meu pai brigá i meu tiu tamém... aí ficamu lá quetinha...* (I. 4- 4a)

(36) Doc *ahn*
Inf *então **nóis** fazia né?... então nóis ia fazenu...* (I. 4- 4a)

5.4 Pronome que estabelece contraste com outro referente:

(37) *ela gostava muitu di namorá né?... namorá éh nus: camim... camim di roça... ela era mocinha já **eu** era piquena...* (I.6-5a)

6) Em estrutura clivada, com as seguintes possibilidades de sequências, conforme Pereira (2004, p. 93):

- *ser* SN *que* X
  (38) *é **a genti** qui pricisa né?...* (I. W-20b)

- SN é que X
  (39) ***elas** é qui tá errada...* (I. P- 17a)

- SN que X
  (40) ***elis** qui foru cheganu i fazenu né?...* (I. D-10b)

4.3.1.6 INDETERMINAÇÃO DO REFERENTE

Este fator foi proposto com base no estudo de Neves (2007: 48), no qual a autora verificou que o pronome de primeira pessoa do singular utilizado como indeterminador do sujeito, ou seja, com referenciação indeterminada, tende a ser mais explicitado.

Pretendemos verificar se o mesmo acontece no *corpus* do Português Popular em todas as pessoas do discurso, sendo possível responder a seguinte questão: haveria maior frequência de preenchimento do pronome subjetivo quando o referente fosse indeterminado?

Muitos autores como Brown e Yule (apud MICHELETTI, 2000: 67) sugerem diferentes graus de indeterminação, que compõem um contínuo gradativo do [-indeterminado] ao [+ indeterminado], em função da presença e da natureza das pistas referenciais encontradas. Porém, para essa pesquisa agrupamos todos os graus de indeterminação em um único subfator: referente indeterminado.

Com base nas sugestões de Castilho (2010: 298), Micheletti (2000: 67) e Neves (2000: 463), propomos os dois subfatores:

Categorias	Exemplos:
d referente determinado	*"...deitava u cepu nu jegui... i* ***eli ficava*** *todu inchadu..." (I.6-5a)*
i referente indeterminado	*"eu pegava um pau di cansanção qui* ***elis fala*** *ortiga..." (I.6-5a)*

Quadro 9: Categorias do fator Indeterminação do Referente.

O sujeito é considerado **determinado** quando o referente é identificável, isto é, definido no contexto discursivo.

Para Castilho (2010: 298), "o traço de definitude é definido no texto, ou seja, nenhuma classe passível de figurar como sujeito será intrinsecamente determinada ou indeterminada". Isso ocorre claramente com o pronome **eu** que é, em princípio, altamente determinado, pois remete ao falante de 1ª pessoa do singular, mas que pode em alguns contextos ser utilizado com referência genérica, remetendo ao que qualquer pessoa possa vir a fazer, a uma situação hipotética (Neves, 2000: 463). Como podemos verificar no exemplo abaixo:

Eu *vou lá,* ***fico*** *dois dias fazendo curso, eles* ***me*** *catequizam,* ***me*** *fazem comprar uma tonelada de sabão e abrir o meu negócio* (Neves, 2000: 464).

É no texto que se identifica o sujeito e sua indeterminação. Para definição de sujeito indeterminado, utilizamos os seguintes critérios estabelecidos por Neves (2000: 463) e Castilho (2010: 298):

O **sujeito indeterminado** pode ocorrer nos seguintes casos:

a) Sujeito expresso por pronomes pessoais de **referenciação genérica.**

(41) ***você*** *imagina lá* ***cê*** *vai vivê em minas vivê nu paraná em otros lugar...* ***você****... trabalha só na roça Ø* ***vai*** *tê que levá aquele estuduzinhu di antigamenti...* (I. E-11a)

No contexto do exemplo acima, o pronome *você* não remete a 2ª pessoa do singular, mas a uma pessoa qualquer, por se tratar de uma situação hipotética.

b) Pronome de **terceira pessoa do plural masculino (eles)** com referenciação genérica.

(42) *ela veiu foi na... feiz um abaxu assinado... aí veiu pego us nomi da pessô/ das pessoa qui num tinha né?... i foi na na sabesp... chegô lá pidiu a ligação...* ***Elis*** *falô qui vinham vê si da:va PA fazê a ligação né?* (I. D-10b)

c) Indeterminação com a **terceira pessoa do singular**.

(43) *Inf ...um mundu di televisão nóis quebranu u pau... u donu chego mando nóis embora ((risos))*

Doc *i quebrô alguma televisão?*

Inf ***quebrô*** *nada...* (I. D- 24b)

Neste caso, o sujeito é indeterminado, pois não é possível identificar o referente no contexto discursivo.

d) **Indeterminação prototípica**, com o sujeito elíptico e com o verbo na terceira pessoa do plural: Ø + 3ª pessoa do plural.

(44) *aí até qui Ø* ***fizeru*** *um abaxu assinadu aí Ø* ***foru*** *na... na... na* (I.D- 10b)

e) **verbos "*dicendi*"** utilizados para expressar um boato. Os mais usados são: *Ø dizem que; diz que; Ø falam que,...*

(45) *ele chego.. o rapais tava vendendo limão... descarregô o revolve em cima do rapais que tava vendendo limão* ***Ø dizem que*** *foi por causa de negócio de venda de maconha* (I. E-11a)

f) Primeira pessoa do plural como sujeito indeterminado:

(46) *fui morá (nim) otru barraquim aqui mesmu né? umas maderinha bem:... miudinha caindu... mais depois eu consigui... fazê essi aqui... era um cômudu só... ((a mãe interfere))* ***nóis dexava*** *us mininu trancadu pra mim í trabalhá pur dia...* (I. 6-5a)

Com base em Neves (2000, p. 465), podemos estabelecer que, no exemplo acima, a indeterminação é parcial, porque o pronome nós inclui ao menos um referente determinado, o falante eu. Porém neste trabalho, consideramos esses casos como referente indeterminado.

4.3.1.7 VALOR SEMÂNTICO DO VERBO

O fator valor semântico do verbo foi baseado no estudo de Neves (2007: 49). Para a autora, o verbo além de ser canonicamente entendido como núcleo do predicado, pode ser classificado levando-se em conta as relações estabelecidas entre o núcleo do predicado e a natureza dos papéis semânticos dos argumentos, que contraem relação com o predicado e que, na oração realizada, se apresentam como funções do verbo. Para a classificação do papel semântico do verbo, seguimos o modelo estabelecido por Borba (1991: ix-xvi). Apesar de outros autores, como Chafe (1970), estabelecerem igualmente uma tipologia das funções sintático-semânticas exercidas pelos verbos, a classificação de Borba se mostra mais adequada para nossa pesquisa, pois se trata de uma análise da língua portuguesa. Transcrevemos, a seguir, as definições fornecidas por Borba (1991: xii-xv), com alguns exemplos retirados do *corpus*:

A Verbo de ação: *o primeiro Argumento é um sujeito agente, e não há argumento afetado pela ação verbal, isto é, o argumento não tem nenhuma mudança (de estado físico ou moral, de condição, de posicionamento no tempo ou no espaço).*
Ex.: Este rapaz não ***bebe****, não* ***fuma****, não* ***mente****. (BORBA, 1991, xiii)*

(47) *é... pocu tempu né?... eli* ***bibia*** *dimais* (I.6-5a)

(48) *Ø* ***trabalhei*** *dois anus i dois mesi na tricolã* (I.6-5a)

(49) *eu* ***chorava*** *dia i noiti pa vim imbora pa são pa/ pa...* (I.6-5a)

P Verbo de processo: *o primeiro Argumento é sempre um afetado por algo que está fora dele, ou seja, é um paciente, um experimentador, ou um beneficiário do processo.*
Ex.: O galo ***morreu****. (BORBA, 1991, xiii)*

(50) *eu tenhu tenhu* ***pastadu*** *com essis mininu... (sentido de sofrer)* (I.6-5a)

(51) *aí eli* ***aprendeu*** *a trabalhá né?(I.2-2a)*

B Verbo de ação-processo: *tem no mínimo dois argumentos: um agente/causativo e outro, afetado/efetuado, isto é, o primeiro Argumento faz que o segundo Argumento mude de estado, condição, ou então passe a existir.*
Ex.: O vento ***derrubou*** *o muro. (BORBA, 1991, xiii)*

(52) *aí a genti* ***limpava*** *aquelis pexi lá mesmu.* (I.6-5a)

(53) *Ø* ***rasgava*** *tudu pra mim ganhá otru...* (I.6-5a)

E Verbo de estado: *tem sempre um Argumento que é inativo, não é agente, nem causativo, nem paciente. O sujeito desses verbos funciona como suporte de propriedades (estado/ condição) ou como mero experimentador delas.*
Ex.: Paulo ama Dirce. (BORBA, 1991, xiii)

(54) *eu* ***morava*** *cum a minha irmã.* (I.6-5a)

(55) *eu* ***andava*** *só di shorti mais era di shorti...* (I.6-5a)

Com este fator pretendemos verificar se há uma relação entre o preenchimento do sujeito e o tipo de verbo (de acordo com suas propriedades sintático-semânticas).

4.3.1.8 IDADE DO INFORMANTE

O fator idade do informante foi proposto tendo como base os estudos de Duarte (1996), Tarallo (1991) e Neves (2007). Os autores abordam a simplificação da morfologia verbal e sua relação com o preenchimento do sujeito.

Duarte (1996), em seu estudo diacrônico, constata uma diminuição no quadro dos pronomes subjetivos de um lado, o que resultou na simplificação do paradigma flexional de outro. Assim, de um quadro de 6 formas verbais (*canto, cantas, canta, cantamos, cantais, cantam*) ocorre a redução para 4 formas, em consequência da substituição de *tu* por *você* e *vós* por *vocês* (sendo que o pronome você [s] é de 2ª pessoa, mas a flexão verbal utilizada é de 3ª pessoa); este paradigma coexiste com um terceiro de 3 formas flexionais , com a substituição de *nós* por *a gente*.

Então se houve simplificação da morfologia verbal para 3 formas flexionais (*canto, canta e cantam*) pressupõe-se que o pronome sujeito torna-se mais necessário para transmissão efetiva da mensagem. Resultados de pesquisas anteriores como Duarte (1996) mostram que o falante do PB[10] passou, no decorrer dos anos, a realizar mais o pronome sujeito. Com isso o PB deixa de ser uma língua em que o sujeito é mais passível de omissão, tornando-se este mais explicitado, ou seja, uma língua "não-pro-drop".

Castilho (2010) constata que na variedade do Português Brasileiro preenche-se mais a posição do sujeito que a de objeto, por outro lado em Portugal ocorre exatamente o inverso.

Com este fator e levando-se em conta a noção de tempo aparente, pretendemos confirmar a seguinte hipótese: o falante do Português Popular estaria utilizando mais o pronome do que zero, sendo possível perceber que os idosos utilizariam mais o zero e os mais novos utilizariam mais o pronome. Assim, o Português do Brasil estaria deixando de ser uma língua "pro-drop", passando para "não-pro-drop".

Sendo nosso estudo sincrônico e para verificar se está ocorrendo mudança linguística, utilizamos a sugestão metodológica de Labov (2008: 194) de análise em tempo aparente, ou seja, *"ao longo das dimensões formadas por faixas etárias da população atual".* Assim, utilizamos o conceito de tempo aparente de Labov que nos propõe comparar pelo menos duas gerações sucessivas de falantes com características sociais comparáveis e pertencentes à mesma comunidade de fala (LABOV, 2008: 194).

---

[10] PB é uma forma abreviada para Português Brasileiro.

Para nosso trabalho utilizamos três gerações de falantes do Português Popular, sugeridas por Rodrigues (1987).

O fator compõe-se de três subfatores correspondentes, estratificados de acordo com as seguintes faixas etárias:

J adultos jovens de 20 a 35 anos
M adultos de meia idade 36 a 50 anos
V idosos com mais de 51 anos

É fato que para obtermos resultados mais eficientes, no sentido de verificar se realmente ocorreu mudança linguística, seria necessário conjugarmos evidencias do estudo em tempo aparente com evidências fornecidas pelo estudo em tempo real, da mesma comunidade de fala, ou seja, desenvolvermos posteriormente um estudo de painel. Essa proposta fica como sugestão para estudos posteriores.

CAPÍTULO 5: Apresentação e discussão dos resultados

5.1 Notas Introdutórias

Após os dados terem sido selecionados e codificados, conforme os fatores linguísticos e sociais estabelecidos, foram submetidos ao programa computacional GoldVarb 2001 para efetuação da análise estatística. O programa nos forneceu como resultados:

(1) a frequência geral de realização/ não-realização dos pronomes subjetivos (variável dependente), bem como as frequências associadas aos fatores estabelecidos (variáveis independentes);
(2) a seleção, por ordem de importância, dos fatores estatisticamente mais relevantes (análise Step Up) e dos fatores descartados, estatisticamente não significativos (análise Step Down);
(3) os resultados dos cálculos dos pesos relativos de realização/ não-realização dos pronomes subjetivos conforme os fatores linguísticos e sociais relevantes, sendo que os valores podem variar de 0.0 a 1.0. Considera-se favorecedor da variável o fator com

peso relativo maior que 0.5; desfavorecedor o fator com peso relativo menor que 0.5 e irrelevante o fator com peso relativo em torno de 0.5.

(4) o cruzamento dos fatores (função Crosstab).

5.2 Resultado Geral

O gráfico abaixo apresenta o resultado geral, em porcentagem, de ocorrências de realização e não-realização dos pronomes subjetivos.

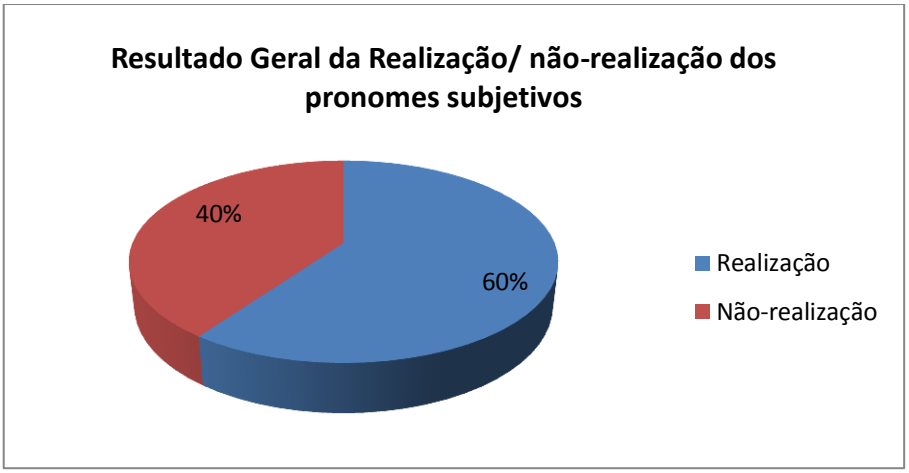

Gráfico 2: Resultado geral de realização/ não-realização dos pronomes subjetivos.

Percebemos que os falantes do Português Popular da cidade de São Paulo que nos forneceram material de análise, de modo geral tendem a realizar os pronomes subjetivos, pois, de um total de 900 ocorrências, 539 (60%) corresponderam a pronome subjetivo realizado, contra 361 (40%) de não-realização do pronome.

5.3 Fatores selecionados e descartados

Temos, a seguir, um quadro com as variáveis independentes selecionadas e descartadas pelo programa Goldvarb 2001, de acordo com sua relevância estatística. É importante ressaltar que os fatores estão representados por ordem de importância. Neste trabalho, trataremos principalmente dos fatores selecionados como mais significativos, sendo possível fazermos, ao longo do texto, observações sobre os fatores considerados descartáveis.

Fatores selecionados	Fatores descartados
5) Status Informacional do Sujeito; 3) Paralelismo Formal; ***8) Idade do Informante;***[11] 1) Pronome pessoal subjetivo; ***2) Forma verbal;*** 4) Polaridade.	***8) Idade do Informante*** 6) Indeterminação do Referente; ***2) Forma verbal*** 7) Valor Semântico do Verbo

Quadro 10: Fatores linguísticos selecionados e descartados pelo programa estatístico Goldvarb 2001.

5.3.1 Status Informacional do Sujeito

O fator Status Informacional do Sujeito foi, dentre todos os fatores, o que se mostrou mais relevante na análise estatística realizada, sendo o primeiro a ser selecionado pelo programa Goldvarb 2001. Os resultados obtidos mostram que a realização dos pronomes subjetivos está diretamente correlacionada com este fator.

Os índices de frequência de realização dos pronomes subjetivos confirmam as hipóteses formuladas: o sujeito pronominal explícito tende a ocorrer mais em contextos ambíguos, quando veicula informação nova e quando há mudança de referente (mudança de tópico discursivo), e menos quando a informação é dada e o referente é mantido no contexto discursivo ou situacional. A tabela a seguir contém os resultados encontrados.

Tabela 6: Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo da realização de acordo com o Status Informacional do Sujeito.

Status Informacional	**Realização**	**Não-realização**	**Peso Relativo**
Novo	305/357= 86%	52/357= 14%	0.88
Dado	234/543= 43%	309/543= 57%	0.21
Total	539/900 = 60%	361/900 = 40%	

Os índices de frequência de realização/ não-realização dos pronomes subjetivos correlacionados ao fator Status Informacional do Sujeito mostram que de um total de 357

[11] Os fatores destacados em negrito no quadro 10 apareceram repetidos entre os fatores selecionados e descartados pelo programa Goldvarb 2001.

ocorrências de informação nova, em 305 o pronome sujeito foi realizado, o que corresponde à frequência de 86%, contra 14% de não-realização do pronome. Com relação à informação dada, de um total de 543 ocorrências de Dado, em 309 o pronome não foi realizado, o que equivale a 57% de frequência de não-realização do pronome, contra 43% de realização.

Além disso, os pesos relativos corroboram para a comprovação das hipóteses apresentadas, ou seja, a realização do pronome subjetivo é favorecida quando o referente veicula informação nova, tendo peso relativo de 0.88, e desfavorecida quando o referente constitui informação dada, com o peso relativo de 0.21.

A título de comparação, notamos que nossos resultados se assemelham aos de Pereira (2004), quanto ao uso do pronome de primeira pessoa do plural, relacionado com o fator Status Informacional do Sujeito. A tabela a seguir ilustra os resultados comparados.

Tabela 7: Frequência de realização do pronome de 1ª pessoa do plural, de acordo com o Status Informacional do Sujeito, em nossa amostra e no *corpus* de Pereira (2004).

	Resultados de Pereira (2004)	**Nossos Resultados**
Status informacional	**Pronome subjetivo explícito**	**Pronome subjetivo explícito**
Novo	171/200= 85%	33/39=85%
Dado	33/110= 30%	8/33= 24%
Total	204/310= 66%	41/72= 57%

Conforme os resultados de Pereira (2004), nas instâncias em que o pronome sujeito de primeira pessoa do plural constitui informação nova, em 85% dos casos o pronome é explicitado, contra 30% de ocorrências de informação dada. Tais indícios condizem com nossos resultados, sendo que temos 85% de frequência de realização do pronome sujeito novo e 24% de frequência de realização do pronome cuja informação é dada.

Apresentamos, na tabela a seguir, os resultados da não-realização dos pronomes subjetivos de primeira pessoa do plural de nossa amostra e do *corpus* de Pereira (2004).

Tabela 8: Frequência de não-realização dos pronome subjetivo de primeira pessoa do plural, conforme o Status Informacional do Sujeito, em nosso *corpus* e no *corpus* de Pereira (2004).

	Resultados de Pereira (2004)	**Nossos Resultados**
Status informacional	**Não-realização do pronome**	**Não-realização do pronome**
Novo	29/200= 14%	6/39=15%
Dado	77/ 110 = 70%	25/33= 76%
Total	106/ 310= 34%	31/72= 43%

Com relação a não-realização dos pronomes subjetivos de primeira pessoa do plural, comparamos nossos resultados aos de Pereira (2004) e constatamos similaridade entre ambos. Nos resultados de Pereira (2004) há uma taxa de 14% de frequência de não-realização do pronome sujeito quando a informação é nova, sendo que, em nosso trabalho, obtivemos 15% de frequência de omissão do sujeito novo. Já no caso do sujeito dado, em nosso *corpus* temos uma taxa de 76% de omissão do pronome subjetivo, e no *corpus* de Pereira (2004), 70% de frequência de zero quando o referente veicula informação dada. Em outras palavras, o fator Dado favorece a omissão do pronome sujeito nas duas amostras de língua.

Deste modo, tanto nossos resultados quanto os de Pereira (2004), relativos à primeira pessoa do plural, nos autorizam a afirmar que o falante do Português Popular tenderia a realizar o pronome sujeito, quando este veicula informação nova, e a não realizá-lo quando a informação do referente já for conhecida, ou seja, for dada.

5.3.2 Paralelismo Formal

Este fator foi o segundo a ser selecionado em ordem de relevância. Os resultados obtidos confirmam a hipótese de o paralelismo formal estar atuando no *corpus.* Desta forma, temos tanto um fator de natureza funcional, *Status informacional do Sujeito*, considerado como mais significativo, quanto um fator de natureza antifuncional, o *Paralelismo Formal*, considerado como relevante.

Com a caracterização dos subfatores, pudemos correlacionar referência discursiva com o processamento paralelo. Assim, estipulamos que quando houvesse troca de referência, o

processamento paralelo não seria validado, pois este é identificado por meio da repetição, ou de marcas que levam a marcas ou de zeros que conduzem a zeros.

Desse modo, o subfator (Z), por marcar um zero/ pronome isolado, ou a primeira referência de uma série, ou ainda a mudança de referência, não atua no *paralelismo*, sendo que por meio dele não se torna possível medirmos a repetição, mas sim somente o gatilho desencadeador do *paralelismo*. Os resultados obtidos mostram um alto índice de frequência do subfator (Z): 75% de realização do pronome, quando este veicula uma nova referência, o que condiz com os resultados obtidos sobre o fator Status Informacional do Sujeito.

Abaixo temos resultados pertinentes sobre o paralelismo.

Tabela 9: Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo da realização, conforme o Paralelismo Formal.

Paralelismo	**Realização do pronome**	**Não realização do pronome**	**Peso Relativo**
(Z) Pronome/ zero isolado, referência diferente	250/ 335= 75%	85/335 = 25%	0.30
(Y) ocorrência precedida de pronome, ref. igual	204/339= 60%	135/339 = 40%	0.70
(X) ocorrência precedida de zero, ref. igual	85/226 = 38%	141/226= 62%	0.47
Total	539/ 900 = 60%	361/900 = 40%	

Os resultados apresentados na tabela acima, sobre a frequência de realização/ não-realização dos pronomes subjetivos relacionada ao fator *Paralelismo Formal*, mostram que de 339 ocorrências (zero/ pronome) precedidas por pronome cuja referência é mantida, em 204 o pronome subjetivo foi realizado, o que corresponde a 60%. Por outro lado, de 226 ocorrências (zero/ pronome) precedidas de zero, com a mesma referência, em 141 o pronome subjetivo não foi realizado, ou seja, 62%.

Assim, tais índices nos possibilitam afirmar que no *corpus* analisado os pronomes omitidos conduzem à não-realização dos pronomes, (subfator X com 62%); já os pronomes subjetivos explicitados conduzem à realização de pronomes explícitos (subfator Y com 60%). Este último pode ser comprovado pelos resultados dos cálculos de frequência e de peso

relativo dos pronomes explícitos que mostram uma tendência favorecedora da realização dos pronomes subjetivos quando estes são precedidos por pronomes em que a referência é mantida, tendo o peso relativo de 0.70 para o fator Y. Além do mais, o peso relativo de 0.47 do subfator X desfavorece, ainda que levemente, a realização do sujeito. Já o subfator Z não se mostrou favorecedor, com peso relativo de 0.30, pois assim como foi dito anteriormente, este não atua no paralelismo.

5.3.3 Pronome Pessoal Subjetivo / Pessoa discursiva

O fator pronome pessoal subjetivo foi estipulado com o intuito de obtermos uma visão geral dos pronomes que estão sendo mais explicitados e mais omitidos no Português Popular em São Paulo.

Podemos perceber com os resultados obtidos que o pronome correspondente à 1ª pessoa discursiva é o mais usado, sendo que de um total de 900 ocorrências, 471 são referentes à 1ª pessoa do singular, somando a realização e não-realização de *eu*. Já o pronome de segunda pessoa do plural é o menos utilizado, pois de 900 ocorrências, 11 são referentes a *vocês*, sendo este também o pronome mais omitido, com 54% de não-realização. É o que nos mostra a tabela a seguir.

Tabela 10: Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo da realização, conforme o fator Pronome Pessoal.

Pronome pessoal subjetivo	Realização do pronome	Não-realização do pronome	Peso Relativo
Eu	294/471 = 62%	177/ 471 = 38%	0.58
Você	9/17 = 52%	8/17 = 47%	0.22
Ele, ela	95/167= 56 %	72/ 167 = 43%	0.45
Nós	41/ 72 = 56%	31/72 =43%	0.37
Vocês	5/11 = 45%	6/11= 54%	0.17
Eles, elas	27/ 57 = 47%	30/57 =52%	0.33
A gente	68/ 105 = 64%	37/105 = 35%	0.47
Total	539/900 = 60%	361/900 = 40%	

Com base no estudo de Silva (1993), acreditávamos que o pronome de primeira pessoa do singular, por ser mais recorrente e também mais predizível, tenderia a ser mais omitido. Contudo, nossos resultados mostram que o pronome de primeira pessoa do singular é o mais

realizado, sendo que de um total de 471 ocorrências de *eu*, em 294 (62%) o pronome foi realizado, contra 38% de não-realização. Os resultados dos cálculos de peso relativo de realização dos pronomes subjetivos demonstram uma tendência favorecedora da realização do pronome de 1ª pessoa do singular, sendo seu peso relativo 0.58. O gráfico a seguir ilustra a frequência da realização dos pronomes subjetivos relacionados à pessoa discursiva.

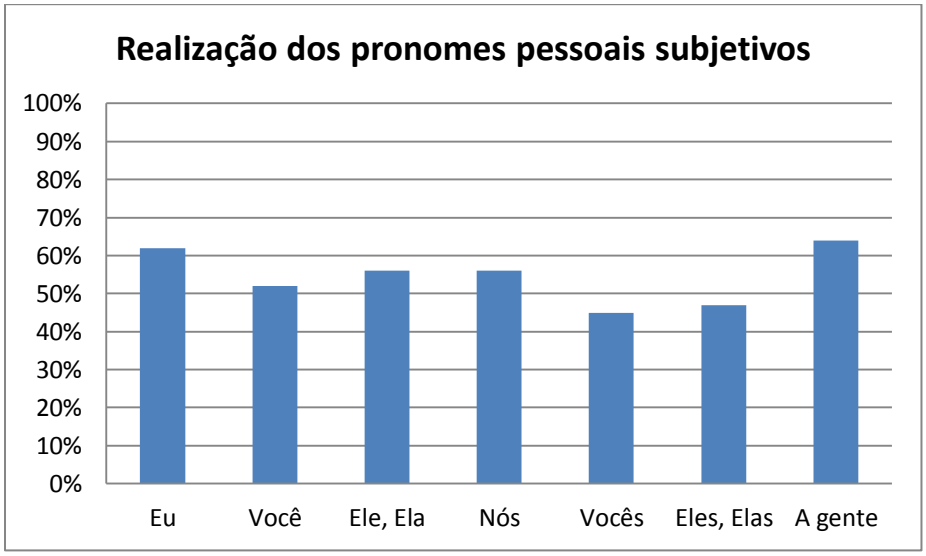

Gráfico 3: Frequência de realização dos pronomes subjetivos conforme a pessoa discursiva.

Podemos visualizar no gráfico acima uma certa regularidade com relação à realização dos pronomes, sobressaindo os pronomes *a gente* e *eu*, os quais aparecem com maior índice de realização e o pronome *vocês* com menor.

Com o cruzamento dos grupos de fatores Pronomes Pessoais e Status Informacional do Sujeito (tabela a seguir), podemos perceber que os maiores índices de realização são dos pronomes *a gente*, com 98% e do pronome *eu*, com 86%, ambos com referentes de informação nova, o que explica os resultados apresentados acima. Além disso, notamos que, em todas as pessoas do discurso, é enorme a diferença entre o índice de realização dos pronomes conforme o referente seja Novo ou Dado, ou seja, maior é a realização dos pronomes que veiculam informação nova e menor quando a informação é dada.

O resultado do cruzamento desses fatores pode ser visualizado a seguir.

Tabela 11: Frequência de realização dos pronomes subjetivos, resultante do cruzamento dos fatores Pronome Pessoal e Status Informacional do Sujeito.

	Status Informacional do Sujeito	
Pronomes pessoais	**Novo**	**Dado**
Eu	86%	52%
Você	58%	40%
Ele, ela	83%	36%
Nós	85%	24%
Vocês	67%	20%
Eles, elas	85%	16%
A gente	98%	32%

Com relação à não-realização do pronome subjetivo, pretendíamos verificar, em nossa amostra, qual o pronome que está sendo mais omitido. Assim, se compararmos os resultados da omissão dos pronomes de 1ª pessoa do singular de nosso trabalho com os de Silva (1993), podemos perceber que em nosso *corpus* o pronome *eu* é pouco omitido, tendo um percentual de 38% de não-realização, o que diverge dos resultados de Silva (1993), em que o mesmo pronome apresenta uma taxa de 77% de omissão. Como podemos verificar abaixo.

Tabela 12: Comparação da frequência de não-realização dos Pronomes de nosso *corpus* e do *corpus* de Silva (1993).

Pronomes	**Não-realização dos pronomes. Nossos Resultados.**	**Não-realização dos pronomes (SILVA, 1993)**
Eu	38%	77%
Você	47%	30%
Ele, ela	43%	50%

Podemos perceber mesma tendência nos resultados do trabalho diacrônico de Duarte (1993), em que os pronomes de 1ª e 2ª pessoas estão sendo, no decorrer dos anos, mais explicitados. Além disso, notamos que no trabalho de Ilari (1996), os pronomes de 1ª pessoa são os mais frequentemente realizados (75%). Tais resultados ocorrem, possivelmente, pelo tipo de discurso ser entrevistas de narrativas pessoais, em que o falante geralmente se coloca

como foco da conversa, utilizando o pronome *eu*, para expressar geralmente suas experiências pessoais, ou *a gente*, para falar de seus familiares e/ou amigos, ou até mesmo referindo-se a primeira do singular. Pensamos que é por conta disso que o pronome *a gente* se mostrou o mais frequente no *corpus*, tendo 64% de realização.

Outro indicio que nos possibilita fazer tais afirmações são os pesos relativos da realização dos pronomes de segunda pessoa do singular e plural: *você* com peso relativo de 0.22 e *vocês* com peso relativo de 0.17, sendo que neste tipo de entrevista raramente há interação entre documentador e informante, sendo a fala do informante a que de fato prevalece nesse tipo de interação.

Dessa forma, há correlação entre a frequência da pessoa do discurso e a realização/ não-realização do pronome subjetivo, pois se o falante expressa mais suas experiências de vida dentro de sua comunidade através de narrativas pessoais, ele tenderá a utilizar mais a primeira pessoa, assim o tipo de discurso interfere na realização/ não-realização de um determinado pronome.

5.3.4 Polaridade

O fator polaridade se mostrou relevante na análise estatística, sendo o último a ser selecionado pelo programa. Os resultados gerais obtidos estão expostos na tabela a seguir.

Tabela13: Frequência de realização/ não-realização do pronome subjetivo e peso relativo da realização, conforme fator Polaridade.

Polaridade	**Realização do pronome**	**Não-realização do pronome**	**Peso Relativo**
Afirmativa	456/ 741= 61%	285/ 741= 39%	0.52
Negativa	83/159= 52%	76/159= 48%	0.41
Total	539/900= 60%	361/900= 40%	

Os resultados apresentados confirmam em parte (somente pelo peso relativo) a hipótese formulada: sentenças afirmativas favorecem a realização dos pronomes subjetivos, pois de 741 ocorrências de sentenças afirmativas, em 456 o pronome subjetivo foi explicitado, o que equivale a uma frequência de 61%. O peso relativo de 0.52 demonstra favorecer levemente a realização do pronome no contexto afirmativo.

Porém, se de um lado temos a afirmação favorecendo a realização do pronome, de outro lado, as sentenças negativas favoreceriam o zero.

Com relação às sentenças negativas, percebemos que os índices de frequência de não-realização do pronome não se mostraram significativos, pois de 159 ocorrências de sentenças negativas, em 76 o pronome foi omitido, o que equivale a uma frequência de 48% contra 52% de realização do pronome nestas instâncias. Dessa forma, pelos índices de frequência não podemos afirmar que as sentenças negativas favorecem a omissão do pronome subjetivo.

Por outro lado, o peso relativo de realização do pronome negativo é de 0.41, ou seja, desfavorecedor, o que confirma nossa hipótese, nos possibilitando dizer que as sentenças negativas inibem a realização do pronome subjetivo.

Nossos resultados correspondem aos de Neves (2007: 205), com relação às sentenças afirmativas. A autora analisou apenas os pronomes de primeira pessoa do singular, por este motivo cruzamos os fatores *Pronome pessoal subjetivo* com *a Polaridade* para selecionar os índices de realização e não-realização do pronome de primeira pessoa do singular, sendo possível obtermos resultados para serem comparados aos de Neves.

Na tabela a seguir, apresentamos os resultados obtidos com base nos dados de nosso *corpus*, comparando-os aos resultados de Neves (2007).

Tabela14: Frequência de realização/ não-realização do pronome de primeira pessoa do singular, conforme a Polaridade, no *corpus* sob análise e no *corpus* de Neves (2006).

	Nossos resultados		**Resultados de Neves (2007)**	
Polaridade	**Realização**	**Não-realização**	**Realização**	**Não-realização**
Afirmativa	235/364= 65%	129/364= 35%	66/96= 69%	30/96= 31%
Negativa	59/107= 55%	48/107= 45%	20/230= 8%	210/230= 92%

Neves (2007) constatou, em seu *corpus*, que a sentença afirmativa favorece a realização do pronome, tendo 69% de frequência, o que condiz com nossos resultados, em que temos 65% de frequência de realização do pronome de primeira pessoa do singular. Já para as sentenças negativas, verifica-se, no trabalho de Neves (2007), uma taxa de 92% de omissão do pronome de primeira pessoa quando inserido em enunciados negativos. Neste ponto nossos resultados não se mostraram semelhantes, tendo a taxa de 45% de frequência de não-realização do pronome quando inserido em enunciado negativo.

Assim, podemos afirmar, com relação às frequências comparadas, que a polaridade influencia em parte a variável sob análise, sendo que sentenças afirmativas favorecem a realização do pronome, enquanto as negativas não se mostraram relevantes.

5.3.5 Idade do Informante

O fator *Idade do informante* foi o terceiro a ser selecionado quanto ao grau de relevância. Porém, o que nos chamou atenção foi o fato de os fatores *Idade do informante* e *Forma verbal* estarem presentes na análise Step up (fatores mais significativos) e se repetirem na análise Step down (fatores menos significativos), o que a princípio não deveria acontecer.

Desta forma, passamos a pensar que os resultados possivelmente estivessem sendo enviesados. Por esse motivo, fizemos várias rodadas no programa eliminando um fator por vez para verificarmos se a repetição dos fatores nas duas análises (Step up e Step down) continuaria ocorrendo. Como resultado desse teste, notamos que quando retiramos o fator *Idade do informante* e rodamos o programa, o fator *Indeterminação do Referente* que antes só estava na análise Step down, assume o lugar do fator excluído, passando a se repetir na análise Step up, juntamente com a repetição do fator *Formal verbal*.

Na terceira rodada, retiramos o fator *Forma verbal* e rodamos o programa, desta vez o fator *Idade do Informante* se repetiu nas duas análises (Step up e Step down). Percebemos assim que um enviesamento de resultados poderia estar ocorrendo de fato.

No quadro abaixo temos os resultados das três rodadas.

Rodada 1: Todos os fatores		**Rodada 2: sem fator Idade do Informante**		**Rodada 3: sem o fator Forma Verbal**	
Step up	**Step down**	**Step up**	**Step down**	**Step up**	**Step down**
Status informacional	***Idade***	Status Informacional	***Indeterminação do referente***	Status informacional	***Idade***
Paralelismo	Indeterminação do referente	Paralelismo	***Forma Verbal***	Paralelismo	Indeterminação do referente
Idade	***Forma verbal***	***Indeterminação do referente***		***Idade***	
Pronome pessoal		***Forma verbal***		Pronome pessoal	
Forma Verbal		Pronome pessoal			

Polaridade					

Quadro 11: Análises Step Up e Step Down referentes a três rodadas: com todos os fatores; sem o fator Idade; sem o fator Forma Verbal, respectivamente.

Os resultados obtidos com o grupo de fatores *Idade do Informante* não confirmaram a hipótese de que o Português Popular estaria caminhando em direção a uma língua "não-pro-drop", em que os pronomes subjetivos são mais explicitados. A seguir temos os resultados encontrados.

Tabela 15: Frequência de realização/ não-realização dos Pronomes Subjetivos e peso relativo de realização, conforme a Faixa Etária.

Idade	**Realização do pronome**	**Não-realização do pronome**	**Peso Relativo**
Idosos	177/301= 59%	124/301= 41%	0.46
Adultos	212/299= 71%	87/ 299 = 29%	0.61
Jovens	150/300= 50%	150/300= 50%	0.42
	539/900= 60%	361/900= 40%	

Observando os resultados, podemos perceber que os informantes idosos têm uma frequência de 59% de realização dos pronomes subjetivos contra 41% de omissão, ou seja, nos nossos dados o falante desta faixa etária utilizou mais o pronome do que o omitiu, justamente o resultado contrário ao qual esperávamos, por outro lado o resultado do peso relativo de aplicação do pronome não demostrou uma tendência favorável a sua realização, sendo o peso relativo de 0.46.

Os adultos apresentaram maior frequência de realização do pronome, 71%, contra 29% de omissão, e um peso relativo significativo de 0.61, ou seja, a faixa etária adultos favorece a realização do pronome. Analisando os resultados desta faixa etária (adultos) e comparando-os aos da faixa etária anterior (idosos), ou seja, um estudo em tempo aparente, tornaria possível afirmarmos que os falantes estão caminhando em direção a uma tendência de maior preenchimento do pronome sujeito. Porém, ao analisarmos os resultados da faixa etária jovens ficamos impossibilitados de fazermos tal afirmação, pois seus índices foram de 50% de frequência de realização dos pronomes subjetivos, ocorre assim uma diminuição da frequência de realização do pronome das faixas etárias adultos para jovens, sendo neste

último o peso relativo de 0.42, ou seja, não favorecedor da realização do pronome; a mesma frequência (50%) temos para a não-realização dos pronomes.

Podemos ilustrar os resultados obtidos, no gráfico a seguir.

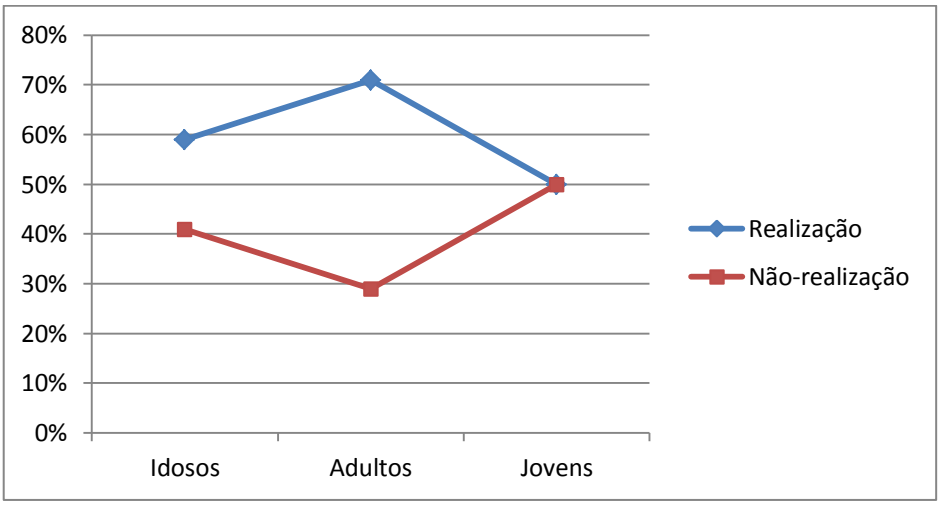

Gráfico 4: Frequência de realização/ não-realização dos pronomes subjetivos, de acordo com as Faixas Etárias dos informantes.

Analisando a linha azul, que corresponde à realização do pronome, podemos verificar o pico de 71% na faixa etária adultos e depois uma queda na faixa etária jovens, ou seja, trata-se de um padrão curvilíneo, o que indica não estar ocorrendo mudança linguística em progresso. Vale ressaltar que o índice de realização do pronome na faixa etária adultos foi o que se mostrou mais relevante, o que provavelmente pode ter causado o enviesamento de resultados com outro grupo de fatores, já que foi o fator mais significativo dentro deste fator Idade.

À medida que a realização do pronome aumenta entre as faixas etárias idosos e adultos, diminui sua não-realização (linha vermelha), passando de 41% para 29% de omissão do pronome, e aumenta entre as faixas etárias adultos e jovens: de 29% para 50%.

Assim, de modo geral, podemos afirmar que nossos dados não nos possibilitam fazer previsões de tendência de mudança linguística em direção à maior realização ou não-realização dos pronomes subjetivos, sendo que os resultados obtidos no grupo dos jovens são de 50%, tanto para a realização quanto para a não-realização dos pronomes subjetivos. E uma nova análise de material coletado posteriormente, num estudo de painel, poderia desvendar o caminho pelo qual a variável está seguindo.

5.3.6 Forma Verbal

O fator Forma Verbal foi, dentre os fatores analisado, um dos que apareceram nas análises Step up e Step down, o que sugere ter ocorrido um enviesamento de fatores.

Os resultados obtidos confirmam, em parte, a hipótese formulada: com a forma verbal marcada, o pronome tenderia a ser omitido, já com a forma verbal não-marcada tenderia a ser explicitado. Os índices de formas verbais não-marcadas em pessoa e número demonstraram-se favorecedores da realização do pronome subjetivo, sendo que de 441 ocorrências de formas verbais não-marcadas, em 278 o pronome subjetivo foi realizado, o que corresponde a 63% de frequência. Desta forma, notamos uma tendência favorecedora da realização do pronome nestas instâncias, com peso relativo de 0.55.

Já para as formas verbais marcadas em pessoa e número, por não apresentarem ambiguidade morfológica, esperávamos que favorecessem a omissão do pronome sujeito. Porém, os índices demonstraram o contrário: baixa frequência de omissão do pronome (43%) quando a forma verbal é marcada, e maior frequência de realização do pronome (57%) para este fator. Analisando apenas os resultados da frequência, notamos que as formas verbais marcadas não favorecem a omissão do pronome subjetivo. Por outro lado, o peso relativo de 0.45 desfavorece levemente a realização do pronome em contextos de forma verbal marcada.

Assim, o uso do pronome pode ser interpretado, a princípio, como redundante, ou seja, antifuncional, pois na forma verbal marcada já há informações de número e pessoa discursivas nas desinências verbais. A tabela a seguir contém os resultados encontrados.

Tabela 16: Frequência de realização/ não-realização dos pronomes subjetivos e peso relativo de realização, conforme a Forma Verbal.

Forma verbal	**Realização do pronome**	**Não-realização do pronome**	**Peso Relativo**
Não-Marcada	278/441= 63%	163/ 441= 37%	0.55
Marcada	261/ 459= 57%	198/459= 43%	0.45
Total	539/900= 60%	361/900= 40%	

Decidimos, então, cruzar os fatores Forma Verbal e Status Informacional do Sujeito para verificarmos se estava ocorrendo redundância de fato. Na tabela 17, podemos observar que o subfator Novo prevalece tanto para as formas verbais não-marcadas com 93% de

frequência, quanto para as formas verbais não marcadas, tendo esta 76% de frequência de realização do pronome.

Desse modo, na tabela 16, a frequência de 57% do pronome com forma verbal marcada não pode ser interpretada como redundante, muito menos como um índice de um resultado antifuncional. Essa taxa é altamente justificada através dos resultados do cruzamento dos fatores, que indicam não ser a desinência verbal fator prioritário que favorece a realização do pronome subjetivo, mas sim o seu Status Informacional, sendo o subfator Novo definidor da realização do pronome, de modo que a realização do pronome está extremamente relacionada ao discurso. Podemos perceber isso, na tabela 17, pela alta taxa de frequência de realização do pronome sujeito novo com forma verbal marcada (76%).

Tabela 17: Frequência de realização dos pronomes de acordo com o cruzamento dos fatores Forma verbal e Status informacional do sujeito.

	Forma verbal	
Status Informacional	**Não-marcada**	**Marcada**
Novo	182/195=93%	123/162= 76%
Dado	96/246= 39%	138/297= 46%
Total	278/441= 63%	261/459=57%

5.3.7 Indeterminação do Referente

O fator Indeterminação do Referente foi descartado pelo programa, sendo que em todas as rodadas feitas este foi considerado como não significativo. Assim, nossa hipótese de que haveria maior frequência de preenchimento do sujeito quando o referente fosse indeterminado não foi confirmada.

Os índices abaixo correspondem aos resultados encontrados.

Tabela 18: Frequência de realização/ não-realização dos pronomes subjetivos conforme a Indeterminação do Referente.

Indeterminação do Referente	**Realização do pronome**	**Não-realização do pronome**
Determinado	422/687= 61%	265/ 687= 38%
Indeterminado	117/213 = 54 %	96/213= 45%
Total	539/900= 60%	361/900= 40%

.

Os resultados de frequência de realização do pronome subjetivo demonstram que de 687 casos de referente determinado, em 422 o pronome subjetivo foi realizado, ou seja, 61% de frequência. Já no caso de referente indeterminado, em 213 casos de referente indeterminado, em 117 o pronome foi realizado, 54% de frequência, o que não possibilita confirmar nossa hipótese, sendo que o maior índice de realização do pronome está em contextos de referente determinado, como já foi dito anteriormente.

5.3.8 Valor Semântico do verbo

Este fator não apareceu em nenhuma das análises, Step Up e Step Down, apesar de estar, neste trabalho, discriminado no conjunto de fatores descartados pelo programa. Desse modo, não conseguimos confirmar a hipótese de que há relação entre a realização/ não-realização do pronome sujeito e as propriedades sintático-semânticas dos verbos. A tabela abaixo apresenta os resultados obtidos.

Tabela 19: Frequência de realização/ não realização dos pronomes subjetivos, conforme o Valor Semântico do Verbo.

Valor Semântico do Verbo	**Realização do pronome**	**Não-realização do pronome**
Ação	235/377= 62%	142/377 = 37%
Processo	156/77= 49%	156/79 = 50%
Estado	236/154= 65%	236/82= 34%
Ação-processo	131/ 73= 55%	131/58= 44%
Total	539/900= 60%	361/900=40%

Os índices mostram que os verbos de Estado favorecem a realização dos pronomes subjetivos, tendo 65% de frequência, enquanto os verbos de Processo obtiveram o maior índice (50%) de omissão do pronome. Percebemos também que o verbo de Ação foi a segunda categoria com maior índice de realização do pronome, 62%, contra 37% de omissão.

Tais resultados sugerem que os verbos de Estado e Ação favorecem a realização dos pronomes subjetivos. Os verbos de Estado apresentam o sujeito inativo que funciona como um suporte de propriedades, ou como um experimentador dela (BORBA, 1991: xiii). A falta de atividade do sujeito sugere sua maior omissão, porém nossos índices demonstraram que os pronomes subjetivos inativos apresentam a maior frequência de realização de pronomes com 65%.

Os verbos de Ação ocorrem sempre quando o sujeito é *agente* da atividade, assim, ele de certa forma, se destaca, pois é ele que faz a ação. O verbo de ação compõe sempre uma frase ativa que indica um *fazer* por parte do sujeito. (BORBA, 1991: xii)

Sendo o sujeito agente, acreditávamos que os pronomes com verbo de ação seriam os mais explicitados, pois é o sujeito agente que, em geral, constitui o foco da oração. Os índices de frequência de realização dos pronomes sujeitos indicaram que o pronome sujeito *agente* tende a ser um dos mais realizados.

Fizemos o cruzamento dos fatores Valor Semântico do verbo e Status Informacional, para verificar se este influencia aquele. Os resultados na tabela a seguir, apontam decisivamente para os índices da informação nova que favorecem mais a realização do pronome subjetivo do que propriamente o fator valor semântico.

Tabela 20: Frequência da realização dos pronomes subjetivos, considerando-se o cruzamento dos fatores Valor Semântico do verbo e Status Informacional.

Valor Semântico do verbo	**Status Informacional**	
	Novo	**Dado**
Ação	145/172= 84%	90/205= 44%
Processo	43/49= 88%	34/107= 32%
Estado	81/93 = 87%	73/143= 51%
Ação-processo	36/43= 84%	37/88 = 42%
Total	305/357 = 85%	234/ 543= 43%

Os índices acima demonstraram que há maior frequência de realização dos pronomes subjetivos de todos os verbos quando seu primeiro argumento constitui informação nova, sendo a maior frequência com o verbo de Processo, com 88%, e os menores índices com os verbos de Ação e Ação-Processo. Assim, o sujeito novo tem maior peso que o fator Valor Semântico do Verbo, sendo o foco da informação discursiva preponderante, e que de fato favorece a realização do pronome.

5.3.9 Análise da Realização dos pronomes por indivíduo

O gráfico abaixo demonstra os resultados da realização dos pronomes subjetivos analisados por indivíduo.

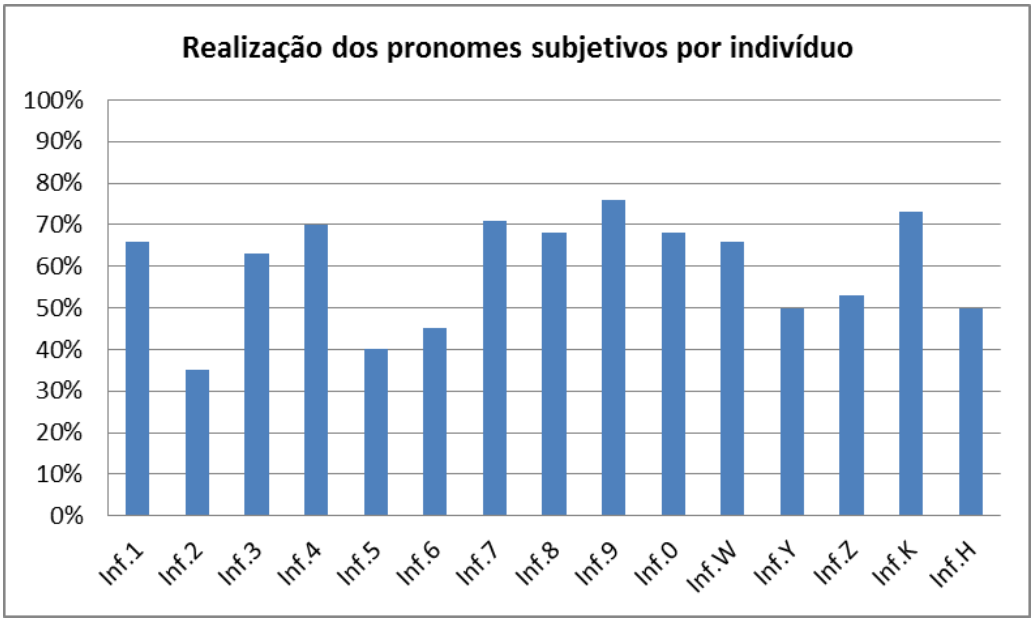

Gráfico 5: Frequência de realização dos pronomes subjetivos por indivíduo.

No gráfico acima, podemos notar, de modo geral, que ainda há uma tendência, mesmo não muito alta, para a realização dos pronomes subjetivos, sendo que de 15 informantes 10 apresentaram taxa maior de 50% de realização.

Podemos perceber que o informante 2 foi, dentre os 15 informantes, o que menos realizou o pronome, tendo uma taxa de 35 % de realização do pronome. O informante 9, por outro lado, foi o que mais realizou o pronome com uma frequência de 76% de realização. O informante 2, que menos realizou os pronomes, é um homem de faixa etária jovem, que cursou o Ensino Fundamental I (até 4ª série) e possui a profissão de biscateiro. Já o informante 9 é uma mulher, dona de casa, adulta, que nunca foi à escola. Apesar de não considerarmos o fator sexo, comparando esses dois informantes, percebemos que a mulher realiza mais o pronome enquanto o homem o omite mais. A tabela abaixo comprova esta afirmação.

Tabela 21: Frequência de realização/ não-realização dos pronomes subjetivos, conforme o Sexo dos informantes.

Sexo	**Realização**	**Não-realização**
Mulher	303/480= 63%	177/480= 36%
Homem	236/420= 56%	184/420 = 43%
Total	539/900= 60%	361/900= 40%

Podemos perceber, pelos resultados acima, que as mulheres realizam mais os pronomes subjetivos, com 63% de frequência, do que os homens, estes com 56%, ainda que seja pouca a diferença entre ambos.

Com relação à ocupação dos informantes 2 e 9, a mulher é dona de casa, possuindo um ambiente mais limitado de convívio social e consequentemente de comunicação, sendo que sempre se comunica com as pessoas que mantém vínculos; o fato de nunca ter ido à escola corrobora essa afirmação. Já o informante 2 é homem e possui a ocupação de ser biscateiro, ou seja, viver de serviços autônomos, arranjados, o que requer uma maior comunicabilidade com várias pessoas fora de seu ambiente familiar, abrangendo mais suas possibilidades de estabelecer relações com diversão pessoas e de se comunicar com falantes de outras variedades de língua, além de ter frequentado a escola, sendo exposto a várias situações de interação linguística.

Apesar dessas observações, é lícito afirmar que fatores sociais como: ocupação, escolaridade e maior exposição a diferentes variedades linguísticas, parecem não influenciar significativamente a realização/não-realização do pronome subjetivo.

Considerações Finais

Por meio desta pesquisa pudemos analisar, de um ponto de vista sincrônico e quantitativo, o perfil do comportamento linguístico dos falantes do Português Popular da cidade de São Paulo em relação à variável realização/ não-realização dos pronomes subjetivos de todas as pessoas discursivas.

Inicialmente ressaltamos o objetivo geral e as hipóteses norteadoras desta pesquisa.

No primeiro capítulo foram apresentados os pressupostos teórico-metodológicos utilizados, fornecidos pela Sociolinguística Variacionista e pela Gramática Funcional de Halliday. Nele tratamos dos conceitos teórico-metodológicos postulados por Labov (1972) para o estudo da variável linguística, e dos conceitos da linguística funcionalista de Halliday (1994), os quais serviram de suporte para o desenvolvimento da presente dissertação.

O capítulo 2 compreendeu uma resenha dos estudos que abordam o fenômeno linguístico do preenchimento do sujeito por pronome em diversas variedades do Português do Brasil: variedade falada não-culta; variedade falada culta; variedade escrita culta. Foram apresentados os fatores linguísticos e sociais que condicionam a realização/ não-realização do pronome subjetivo, como também as hipóteses formuladas que explicam a variável. Esses trabalhos serviram de subsídio tanto para o desenvolvimento desta pesquisa quanto para explicação dos resultados obtidos.

No terceiro capítulo apresentamos esclarecimentos sobre o *corpus*, ou seja, discorremos sobre o Português Popular falado em São Paulo e suas características; o perfil social dos informantes selecionados; as normas de transcrição e coleta de dados.

No quarto capítulo, abordamos: os critérios de seleção e exclusão de dados, a apresentação e explicação da variável dependente e a caracterização das variáveis independentes, demonstrando os contextos de variação da realização/ não-realização do pronome subjetivo.

No quinto capítulo discorremos sobre a análise e interpretação dos resultados obtidos a partir dos cálculos fornecidos pelo programa Goldvarb 2001. Apresentamos o resultado geral de realização/ não-realização dos pronomes em nossa amostra; os fatores considerados significativos e descartados pelo programa, bem como os fatores que se mostraram condicionantes da realização dos pronomes subjetivos.

Os resultados obtidos confirmaram algumas das hipóteses formuladas, contestaram outras, e há ainda algumas hipóteses que foram parcialmente confirmadas. Com o resultado geral de realização/ não-realização do pronome subjetivo pudemos perceber que os falantes

do Português Popular de São Paulo ainda preenchem mais o pronome subjetivo do que o omitem, sendo que obtivemos uma taxa de 60% de frequência de realização dos pronomes.

Com relação aos nossos resultados, confirmamos a hipótese de que o uso do pronome subjetivo não é aleatório, mas condicionado por fatores linguísticos, principalmente pelo fator funcional Status Informacional de natureza pragmático-discursiva.

Dentre os fatores analisados foram selecionados como significativamente relevantes os que seguem.

a) Status Informacional do Sujeito

Com os resultados obtidos dos cálculos do programa, pudemos perceber que, dentre todos os fatores analisados, o fator Status Informacional do Sujeito foi o que se mostrou mais significativo, sendo o principal fator que condiciona de fato a variável dependente, prevalecendo com relação aos demais fatores.

Os índices confirmaram a hipótese funcional estabelecida, ou seja, a informação nova (em contextos ambíguos) favorece decisivamente a realização dos pronomes subjetivos, enquanto a informação dada a inibe. Pudemos perceber isso, devido à maior frequência (86%) e peso relativo (0.88) de realização do pronome sujeito quando o referente constitui informação nova, e frequência maior de omissão (57%) do sujeito dado.

Nossos resultados demonstraram ser semelhantes aos do trabalho de Pereira (2004), no que se refere à primeira pessoa do plural, tendo índices similares de 85%, em ambos trabalhos, de frequência de realização de pronome sujeito novo. Já para as instâncias de Dado, obtivemos uma taxa de 76% de frequência de omissão do pronome subjetivo, similar aos resultados de Pereira (2004), este com 70% de omissão de sujeito pronominal dado.

b) Paralelismo Formal

O fator Paralelismo Formal se mostrou significativo no *corpus*. Os resultados obtidos apontaram para uma tendência de realização do pronome quando a ocorrência é antecedida por pronome, com 60% de frequência de realização do pronome, e realização de zero quando a ocorrência é antecedida por zero, 62% de omissão de pronome, sendo a referência mantida nos dois casos.

É interessante salientarmos que o Paralelismo Formal é considerado um fator antifuncional, na medida em que se utilizam marcas consideradas desnecessárias ou redundantes.

Com base nos resultados dos cálculos, pudemos perceber que, em nossa amostra de língua, tanto um fator de natureza funcional, o Status Informacional, quanto um fator de natureza antifuncional, Paralelismo, estão atuando.

Verificamos que o alto índice de frequência do subfator Z (75%) é justificado por condizer com o subfator Novo, pois, quando a referência é trocada, o sujeito veicula informação nova, o que favorece a sua explicitação.

c) Pronome Pessoal Subjetivo/ Pessoa Discursiva

Os resultados evidenciaram a existência de correlação entre realização/ não-realização do pronome subjetivo e a de pessoa discursiva. A primeira pessoa do singular se mostrou mais recorrente nos inquéritos (471 ocorrências de realização/ não-realização de *eu),* sendo o segundo pronome mais realizado, o que condiz com os resultados de Duarte (1993) e Ilari (1996), e diverge dos resultados do trabalho de Silva (1993) em que o pronome de primeira pessoa do singular foi o mais omitido por ser o mais predizível.

O tipo de discurso também influencia a frequência de uso da pessoa discursiva e a realização do pronome, sendo que em entrevistas o falante tende a expressar suas experiências de vida, o que o leva a utilizar a primeira pessoa do singular ou a forma *a gente* referindo-se ao seu grupo social.

O pronome *a gente* foi o mais realizado no *corpus,* o que se justifica por este pronome assumir vários valores semânticos de acordo com o contexto discursivo, sendo usado como indeterminado e muitas vezes com o valor de *eu*. Percebemos também que tanto o pronome *a gente* quanto o pronome *eu* demonstraram tendência de realização quando veiculam informação nova.

Assim, percebemos a necessidade de desenvolver pesquisas que analisem os vários significados de *a gente.*

d) Polaridade

Os dados confirmaram em parte a hipótese de que sentenças afirmativas favorecem a realização do pronome sujeito, enquanto as sentenças negativas favorecem o zero. Os

resultados obtidos indicam 65% de realização do pronome quando a oração é afirmativa; contudo os pronomes não tendem a ser omitidos em sentenças negativas, pois apresentaram índice de 48% de omissão do pronome nessas instâncias, razão pela qual a hipótese foi parcialmente comprovada. Nossos resultados são semelhantes aos de Neves somente com relação às sentenças afirmativas, pois ambos apresentam taxa relevante (65% e 69%, respectivamente) de realização dos pronomes de primeira pessoa do singular nesses contextos. Apesar do considerável índice de frequência de realização do pronome na oração afirmativa, o peso relativo não se mostrou muito favorecedor, 0.52.

Assim, podemos afirmar que a polaridade influencia em parte (somente pelo peso relativo) a variável sob análise, sendo que sentenças afirmativas favorecem a realização do pronome, enquanto as negativas a inibem.

e) Idade

Este fator foi o único fator extralinguístico utilizado no trabalho, com a finalidade de efetuarmos um estudo em tempo aparente para identificarmos uma possível mudança linguística em andamento. Os dados obtidos não confirmaram a hipótese de o Português Popular falado na cidade de São Paulo estar caminhando na direção de uma língua "não-pro-drop", em que os pronomes subjetivos deixam de ser omitidos para serem mais explicitados. Apesar de este fator ter sido possivelmente enviesado, percebemos que as frequências de realização nas faixas etárias idosos (59%) e jovens (50%) não nos autorizam a sugerir uma mudança linguística em progresso. Na medida em que na faixa etária jovens, os índices de realização e não-realização dos pronomes subjetivos estão equiparados em 50%.

f) Forma Verbal

Os índices obtidos deste fator confirmaram em parte a hipótese de que o pronome sujeito tende a ser explicitado com as formas verbais não-marcadas, as quais mantém ambiguidade morfológica, com 63% de frequência de realização do pronome, sendo 0.55 o peso relativo; assim há tendência de realização do pronome nestes casos em que não é possível depreender o número e a pessoa discursiva. Já as formas verbais marcadas não favoreceram a omissão dos pronomes subjetivos, com frequência de 43%, o que não confirma a nossa hipótese, sendo este resultado interpretado, a princípio, como antifuncional. Porém

através do cruzamento de fatores, percebemos que o subfator Novo predomina na realização do pronome tanto para as formas verbais marcadas (76%) quanto não-marcadas (93%).

g) Indeterminação do Referente

Os dados obtidos não confirmaram a hipótese de que o sujeito indeterminado tende a ser mais explicitado.

h) Valor semântico do Verbo

Este fator não foi selecionado pelo programa como estatisticamente significativo. Então, a hipótese de que há relação entre a realização/ não-realização do pronome sujeito e as propriedades sintático-semânticas dos verbos não pôde ser confirmada. Os resultados apontaram favorecimento da realização do pronome com os verbos de Estado e Ação. Com o cruzamento deste fator com o Status Informacional, notamos que novamente o subfator Novo é predominante e que de fato condiciona a realização do pronome subjetivo, sendo alta a frequência de realização do pronome novo (84% a 88%) em todas categorias do Valor Semântico do Verbo.

Com a realização desta pesquisa, pudemos identificar, de modo geral, o padrão de variação de realização/ não-realização dos pronomes subjetivos, sendo que o falante do Português Popular tende a utilizar mais o pronome subjetivo explicitado, com 60% de realização. Apesar deste resultado, nosso estudo em tempo aparente não nos possibilitou confirmar tendências de mudança linguística, de modo que um estudo de painel provavelmente nos fornecesse respostas mais consistentes.

Com relação aos fatores linguísticos, podemos afirmar que o fator Status Informacional é o que realmente atua de maneira decisiva. Além do mais, é nítido nos casos em que há mudança de referente, sendo que este veicula informação nova e o contexto é ambíguo, o pronome tende a ser realizado, já quando a informação é dada, em contextos não-ambíguos, o pronome sujeito tende a ser omitido.

Deste modo, Novo/ Dado são definidos pela ***referência*** e a ***ambiguidade textual***, comprovando que a "há uma tendência para a informação semanticamente relevante ser mantida na estrutura superficial" Kiparsky (1972, apud SILVA, 2003: 123), ou ainda que a perda de informação tende a não ocorrer em contextos ambíguos, Poplack (1980, apud SILVA, 2003: 123).

Assim, acreditamos que esta pesquisa tenha contribuído de alguma forma para uma melhor compreensão da realização/ não realização dos pronomes subjetivos no Português Popular do Brasil, somando algum valor ao que foi produzido até então. Além dos estudos que estão sendo realizados sobre a variável em questão, vale salientarmos que este trabalho não se dá por concluído, pois sempre novas leituras, discussões e interpretações podem emergir do mesmo.

Partindo dos resultados obtidos fornecidos pelo programa Goldvarb 2001, podemos sugerir algumas questões para serem analisadas em estudos futuros, como: um estudo de painel e um estudo diacrônico da mesma variedade de língua, obtendo condição de afirmar que a Língua Portuguesa está caminhando numa tendência de língua "não pro-drop"; os vários valores semânticos que o pronome *a gente* adquire; um estudo da mesma variedade linguística, alterando o tipo de texto, ou seja, de entrevistas para diálogos entre os "iguais linguísticos", possibilitando verificar se o falante utiliza outras pessoas discursivas; estudo mais aprofundados dos fatores que foram descartados pelo programa; estudo estilístico para detectar se o falante tende a preencher mais o pronome em situações formais. Além do mais, todos os nossos fatores poderão ser utilizados em estudos futuros com outros *corpora*, verificando em que medida os resultados se assemelham, contribuindo para uma efetiva descrição e caracterização das diversas variedades que compõem o Português do Brasil.

REFERÊNCIAS

BORBA, F. da Silva (coord.). **Dicionário Gramatical de Verbos do Português Contemporâneo.** São Paulo, Ed. da UNESP, 1991. 1373 p.

CAMARA Jr., J. M. **Estrutura da língua portuguesa.** 2. ed. Petrópolis: Vozes, 1970.

CASTILHO, A. T. A elipse do sujeito no português culto falado em São Paulo. In: Seminários do Grupo de Estudos Linguísticos, 1987, Campinas. **Estudos Linguísticos 14**, 1987, p. 32-40.

_____ **Nova Gramática do Português Brasileiro.** 1 ed. São Paulo: Contexto, 2010.

CHAFE, W. L. ***Meaning and the structure of language.*** Chicago/ London: The University of Chicago Press, 1970.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no Português do Brasil. In: ROBERTS, I. e KATO, M. **Português Brasileiro.** Campinas: Editora da UNICAMP, 1993, p. 107-128.

_____. **A perda do princípio "evite pronome" no Português Brasileiro**. Campinas: UNICAMP, 1996. Tese de Doutorado.

GALVES, C. O enfraquecimento da concordância no português brasileiro. In: ROBERTS, I. e KATO, M. (orgs.) **Português brasileiro: uma viagem diacrônica.** Campinas, Editora da UNICAMP, 1993, p. 387-408.

HALLIDAY, M. A. K. "*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, 1974.

_____. Estrutura e função da linguagem. In: Lyons, John (org.). **Novos Horizontes em Linguística.** São Paulo, Ed. Cultrix, 1976, p.134-160.

_____. ***Language as social semiotic – the social interpretation of language and meaning.*** Great Britain: Edward Arnold, 1987.

_____. ***An Introduction to functional grammar***. Great Britain: Edward Arnold, 1994, 434 p.

ILARI, R., FRANCHI, C. , NEVES, M. H. M. Os pronomes pessoais do português falado: roteiro para a análise. In: CASTILHO, A. T., BASÍLIO, M. ***Gramática do Português Falado***. Vol. IV: Estudos Descritivos. Campinas: Unicamp/ FAPESP, 1996, p.79-166.

LABOV, W. Estágios na aquisição do Inglês Standard. In: FONSECA, M. S., e NEVES, M. F. (orgs) **Sociolinguística.** Rio de Janeiro: Eldorado, 1974.

_____. ***Sociolinguistic Patterns*.** Philadelphia, University of Pennsylvania Press, 11 ed. 1972.

_____. **Modelos Sociolinguísticos.** Madrid: Edicione Cátedra, 1983.

_____. ***Principles of Linguistic Change: Internal Factors***. Cambrigde/ Amsterdam/ Philadelphia: Blackwell, 1994.

_____. **Padrões Sociolinguísticos**. São Paulo: Parábola, 2008.

MARTELOTTA, M. E. (org*)* **Manual de Linguística**. São Paulo: Contexto, 2009.

MICHELETTI, H. **A Indeterminação do sujeito: um estudo do pronome de terceira peoa do plural**. São Paulo: USP, 2000. Dissertação de Mestrado.

MOLLICA, M. C. e BRAGA, M. L. **Introdução à Sociolinguística, o tratamento da variação**. São Paulo: Contexto, 2003.

MUSSALIM, F. e BENTES, A. C. **Introdução à Linguística**. Vol. 1. 6 ed. São Paulo: Cortez, 2006.

NARO, A. J. ***Morphological constraints on subject diletion.*** In: SANKPFF, G. & CEDERGREN, H. (eds) *Variaion Omnibus*. Canada: Linguistic Research Inc., 1981, p.351-358.

NARO, A.J. & SCHERRE, M. M. P. **Sobre as origens do português popular do Brasil**. *Delta*, 9, nº especial, 1993, p.437-454.

_____. **Origens do Português Brasileiro**. São Paulo: Parábola Editorial, 2007, p. 49-69.

NEVES, M.H.M. **Uma visão geral da gramática funcional**. Alfa 38: 109-127, 1994.

_____. **Gramática de usos do Português**. São Paulo: Editora Unesp, 2000.

_____. **Texto e gramática**. São Paulo: Contexto, 2007.

OMENA, N. P. e BRAGA, M. L. A gente está se gramaticalizando? In.: MACEDO, A. T.; Roncarati, Cláudia e Mollica, M. C. (Orgs). **Variação e discurso.** Rio de Janeiro: brasileiro, 1996, p. 63-80.

PAIVA, e DUARTE, M. E. L. **Mudança linguística em tempo real**. Rio de Janeiro: FAPERJ/ Contracapa, 2003.

PAIVA, M. C. e SCHERRE, M. **Retrospectiva Sociolinguística: Contribuições do PEUL**. In: D.E.L.T.A., Vol. 15, N.º ESPECIAL, 1999 (201-232) `

PEREIRA, D. C. **Concordância verbal na língua falada nas trilhas das bandeiras paulistas**. São Paulo: USP, 2004. Dissertação de Mestrado.

_____. **Variação e mudança no uso dos pronomes reflexivos no português popular da capital paulista: uma abordagem funcionalista e cognitiva**. São Paulo: USP, 2006. Tese de Doutorado.

PEREIRA, D. C. e RODRIGUES, A.C.S. **O peso do fator discursivo Dado X Novo na variação da concordância verbal de 1ª pessoa do plural no português brasileiro**. *XV Congresso Internacional da ALFAL* (Associação de Linguística e Filologia da América Latina), 18 a 21 de agosto de 2008, Montevidéu, Uruguai.

RAMOS, J. **O uso das formas *você, ocê* e *cê no dialeto mineiro***. In: D'HORA, (org.) *Diversidade linguística no Brasil.* João Pessoa: Ideia Editora, 1997.

RODRIGUES, A.C. S. **A concordância verbal no português popular em São Paulo**. São Paulo: USP, 1987. Tese de Doutoramento.

_____. Fotografia sociolinguística do português do Brasil: o português popular em São Paulo. In:: CASTILHO, A. T. (org). **História do português paulista**. Série Estudos, vol. 1. Campinas: Instituto de Estudos da Linguagem, Unicamp, 2009, p. 151-158. ISBN 978-85-062641-00-8.

RODRIGUES, A. C. S. e CAMPOS, O. G. L. A. S. **Reflexões sobre fatos de não-concordância verbal no português culto brasileiro.** Congresso da ALFAL, Costa Rica, 2002.

RODRIGUES, A. C. S & FERREIRA NETTO, W. Transcrição de Inquéritos: problemas e sugestões. In: MEGALE, H. (org.) **Filologia Bandeirante: Estudos 1**. São Paulo, Humanitas, FFLCH/USP, 2000. p.171-193.

SAID ALI, M. **Gramática secundária e gramática histórica da língua portuguesa.** 3. ed. Brasília: Editora da Universidade de Brasília, 1964.

SILVA, V. L. P. **Cartas cariocas. A variação do sujeito na escrita informal**. Rio de Janeiro: UFRJ, 1988. Tese de Doutorado.

_____. Subject omission and functional compensation: evidence from written Brazilian Portuguese. In: ***Language variation and change***, 5, USA, Cambridge University Press, 1993, p.35-49.

_____.Variação no uso de pronomes de segunda pessoa do singular. In: **Revista de Estudos Linguísticos**. Belo Horizonte, vol. 7, n. 2, p. 122-138, 2003.

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.

TARALLO, F. **Diagnosticando uma gramática Brasileira: O português d´aquém e d´além-mar ao final do século XIX**. Tradução do artigo original, apresentado no colóquio sobre *"La Citoyenneté au XIX e Scièle au Brèsil et em France"*, Paris, Universidade de Paris VII, 1991.

TRASK, R. L. trad. ILARI, R. **Dicionário de Linguagem e Linguística**. São Paulo: Contexto, 2006.

WEINREICH, U., LABOV, W. & HERZOG, M. **Fundamentos empíricos para uma teoria da mudança lingüística.** Trad. BAGNO, M. São Paulo: Parábola Editorial, 2006 [1968].