



Instituto de Letras
Programa de Pós-Graduação em Letras e Linguística
Rua Barão de Jeremoabo, nº147 - CEP: 40170-290 - Campus Universitário Ondina Salvador-BA
Tel.: (71) 3263 - 6256 – Site: <http://www.ppgll.ufba.br> - E-mail: pgletba@ufba.br



LANUZA LIMA SANTOS

A ORDEM VERBO-SUJEITO:

Uma análise sociolinguística da fala popular do interior do estado da
Bahia.

Salvador
2010

LANUZA LIMA SANTOS

A ORDEM VERBO-SUJEITO:

Uma análise sociolinguística da fala popular do interior do estado da
Bahia.

Dissertação apresentada ao programa de Pós-Graduação em Letras e Linguística do Instituto de Letras da Universidade Federal da Bahia como parte dos requisitos para obtenção do grau de Mestre em Letras.

Orientador: Prof. Dr. Dante Lucchesi

Salvador
2010

Sistema de Bibliotecas - UFBA

Santos, Lanuza Lima.

A ordem verbo-sujeito: uma análise sociolinguística da fala popular do interior do estado da Bahia / Lanuza Lima Santos. - 2010.

130 f. : il.

Orientador : Prof. Dr. Dante Lucchesi.

Dissertação (mestrado) - Universidade Federal da Bahia, Instituto de Letras, Salvador, 2010.

1. Língua portuguesa - Ordem das palavras. 2. Contato linguístico. 3. Sociolinguística. I. Lucchesi, Dante. II. Universidade Federal da Bahia. Instituto de Letras. III. Título.

CDD - 469
CDU - 811.134.3

Lanuza Lima Santos

A ORDEM VERBO SUJEITO:

Uma análise sociolinguística da fala popular do interior do estado da
Bahia.

Dissertação apresentada ao programa de Pós-Graduação em Letras e Linguística do Instituto de Letras da Universidade Federal da Bahia como parte dos requisitos para obtenção do grau de Mestre em Letras.

APROVADO EM: 17/05/2010

BANCA EXAMINADORA:

Prof. Dr. Jacyra Andrade Mota - Universidade Federal da Bahia (UFBA)

Prof. Dr. Cláudia Roncarati – Universidade Federal Fluminense (UFF)

Prof. Dr. Dante Lucchesi – Universidade Federal da Bahia (UFBA)

À minha família, meu pai Antonio, minha mãe Genisia e minha irmã Laiane, motivo da minha dedicação e eterna inspiração.

Agradecimentos

A concretude desta dissertação é resultado não somente de dedicação e trabalho, mas do empenho e ajuda de muitos que docemente me cederam seus olhos, seu cérebro, seu sorriso, seu tempo, ou simplesmente sua companhia. A essas pessoas devo não só carinho e admiração, mas a eterna e sincera gratidão.

A Deus as primícias, portanto rendo graças ao meu Senhor que me trouxe e me susteve até aqui, mantendo meu equilíbrio mental e espiritual e minha saúde física.

Agradeço, em especial, aos meus pais, pela imensurável compreensão e pelo amor devotados durante a vida e, sobretudo, nesses dois anos.

A minha irmã, amor de minha vida, pela confiança e compreensão.

Agradeço também aos amigos, próximos e distantes:

A Darliane, pela fé que me transmitiu durante esses anos, fazendo-me crer que, mesmo nas horas mais difíceis, “Deus proveria!”

A Luanda, amiga de todas as horas, por acreditar em mim, mesmo quando nem eu acreditava.

A Elisângela, pelos doces momentos de descontração e, sobretudo, por cumprir brilhantemente o papel de “chata”.

A Vívian, pela amizade e companhia nos congressos e pelas sempre presentes palavras de estímulo e confiança, desde os primeiros momentos de contato com a pesquisa.

Às meninas do Vertentes, Renata, Camila, Shirley e Manuele pelo carinho e pelo apoio com a teoria gerativa.

A Lidiane, pela preocupação com a manutenção do meu equilíbrio e da minha vida social.

As minhas companheiras de congressos, Ana Paula, Daiane, Vanessa Pessanha, Viviane e Gilce por estarem presente nos melhores momentos do mestrado.

A Fábio, pela ajuda geográfica e pela preocupação sempre dispensada.

A todos os amigos que me dedicaram horas de terapias presenciais e, sobretudo, “virtuais”.

Pela valiosa leitura e revisão dos capítulos e referências, agradeço a Elisângela, Gilce, Vivian, Luanda, Cecília Guimarães, Manuele Bandeira, Manoela Ribeiro.

Aos bolsista e voluntários do projeto Vertentes, que laboriosamente coletaram, transcreveram e revisaram o *corpus* aqui investigado.

Por último, mas não menos importante, agradeço à CAPES, pelo apoio financeiro, e ao meu orientador, Dante Lucchesi, que me iniciou nos caminhos da sociolinguística, por acreditar em mim há cinco anos atrás e por aceitar a orientação desta dissertação.

O português perdeu o dono, quer dizer, ficou sem dono. Felizmente! E namorou, e namorou no chão, e namorou na poeira do Brasil, e namorou também aqui, na poeira de Moçambique. (...) Sujou-se, nesse sentido em que é capaz de casar-se com o chão.

Mia Couto (no documentário *Língua: Vidas em português*)

A base do conhecimento intersubjetivo na lingüística tem de ser encontrada na fala – a língua tal como usada na vida diária por membros da ordem social, este veículo de comunicação com que as pessoas discutem com seus cônjuges, brincam com seus amigos e ludibriam seus inimigos.

Weinreich, Labov e Herzog (2006 [1968], p. 13)

RESUMO

O Português Brasileiro (PB) apresenta como padrão mais frequente a ordem sujeito-verbo (SV), apresentando fortes restrições à inversão verbo-sujeito (VS), sobretudo com verbos transitivos. Esse comportamento é o resultado de uma tendência de crescente fixação da ordem SV que caracterizou a diacronia do PB do século XVIII ao XX. Partindo da observação empírica em sentenças interrogativas e declarativas, o presente trabalho busca, sob o enfoque teórico-metodológico da sociolinguística variacionista, identificar os condicionamentos e restrições para as ordenações VS/SV em duas variedades do PB situadas no interior do Estado da Bahia: o português popular dos municípios de Santo Antonio de Jesus e Poções; e o português afro-brasileiro, de duas comunidades rurais remanescentes de quilombo, denominadas Cinzento e Sapé e situadas, respectivamente, nos municípios de Planalto e Valença. A análise busca identificar o padrão linguístico dessas comunidades, observando em que medida se aproxima ou se distancia da tendência geral do português brasileiro. Além disso, a comparação dos dois tipos de comunidades, marcadas e não marcadas etnicamente, visou identificar possíveis marcas da influência do contato entre línguas na sua formação histórica. Para este trabalho, foram utilizados 72 inquéritos do *corpus* Projeto Vertentes do Português Popular do Estado da Bahia: 24 inquéritos do *corpus* do português afro-brasileiro (das comunidades de Cinzento e Sapé) e 48 do *corpus* do português popular (dos municípios de Santo Antônio de Jesus e Poções). As ocorrências levantadas e codificadas foram submetidas ao tratamento estatístico do programa de regras variáveis VARBRUL. Os resultados encontrados apontaram uma forte restrição à ordem VS, em média 5% de ocorrências, contra a média de 20% encontrada na variedade culta. Os resultados, embora convergentes no que tange a preferência pela ordenação SV, demonstraram um contraste entre o português popular e as variedades cultas. Quanto aos fatores condicionadores, fatores de natureza sintática, semântica e discursiva foram apontados como fortes favorecedores da ordem VS, tais como, os verbos inacusativos, o traço indefinido do SN sujeito e o *status* informacional [+novo], entre outros. Entre os fatores sociais, apenas a escolaridade foi apontada como relevante pelo VARBRUL.

Palavras chave: Ordem verbo-sujeito. Sociolinguística. Português popular do Brasil. Contato entre línguas.

ABSTRACT

Brazilian Portuguese (BP) shows more frequently standard direct order subject-verb (SV), showing several restrictions on inversion verb-subject (VS), especially with transitive verbs. This behavior derives from a growing tendency of SV order's fixation that characterized the diachronic course of BP from XVIII century until XX. From empirical observation in declarative and interrogative sentences, the present work investigates, from the standpoint of theoretical and methodological variational sociolinguistics, identifies conditions and variation restrictions for the VS/SV orders in two BP's varieties located in the interior of Bahia: the popular portuguese from Santo Antonio de Jesus and Poçoões; and afro-brazilian portuguese of two rural quilombola communities, Cinzento and Sapé, located in Planalto and Valença respectively. The analysis investigates the linguistics pattern these communities, observing how much they grow or apart from general Brazilian portuguese tendency. Moreover, the comparison of two kinds of communities, ethnically marked and unmarked, aimed in identifying possible marks of contact among languages in its historical formation. For this work, it was used 72 surveys of the Projeto Vertentes do Português Popular do Estado da Bahia's *corpus*: 24 surveys from *corpus* of Afro-brazilian-portuguese (Cinzento and Sapé) and 48 from *corpus* of popular portuguese (Santo Antônio de Jesus and Poçoões). The gathered and encoded occurrences were submitted to statistical treatment with the software of variable rules VARBRUL. The results pointed a strong restriction on the VS order, on average of 5%, for the occurrences, against an average of 20% found in cults varieties. Although convergentes about the preference of SV order, the results shows a contrast between popular portuguese and cults varieties. Syntactic, semantical and discursive factors were appointed as strong favoring factors of VS order, as the inacusatives verbs, indefiniteness trace of SN subject and the informational status [+new], among others. Among social factors, only scholasticity was appointed as important for VARBRUL.

Key words: verb-subject order. Popular portuguese. Sociolinguistics. Linguistic contact.

LISTA DE QUADROS

QUADRO 1: Estruturalismo X sociolinguística variacionista	22
QUADRO 2: Variação e mudança: relação indivíduo/ comunidade.....	32
QUADRO 3: Visão de conjunto da operacionalização do VARBRUL (adaptado de SCHERRE, 1993, p. 6).....	40
QUADRO 4: A relação dos fatores condicionadores com a ordem VS no português afro-brasileiro e no português popular.	100
QUADRO 5: Fatores condicionantes da ordem VS no português afro-brasileiro e no português popular.....	118

LISTA DE TABELAS

TABELA 1: Realização pós-verbal do sujeito na comunidade de Poções, segundo a variável <i>animacidade do sujeito</i>	38
TABELA 2: População do Brasil por etnia do século XVI ao XIX.....	70
TABELA 3: <i>Corpus</i> delimitado para o estudo.....	79
TABELA 4: Distribuição da população brasileira – dados dos censos brasileiros.....	84
TABELA 5: A ordem VS no português afro-brasileiro, segundo a variável <i>constituição do SN sujeito</i>	98
TABELA 6: A ordem VS no português popular, segundo a variável <i>constituição do SN sujeito</i>	99
TABELA 7: A ordem VS no português afro-brasileiro, segundo a variável <i>status informacional do SN sujeito</i>	105
TABELA 8: A ordem VS no português popular, segundo a variável <i>status informacional do SN sujeito</i>	105
TABELA 9: Ordem VS no português afro-brasileiro, segundo a variável <i>tipo de verbo</i>	107
TABELA 10: Ordem VS no português popular, segundo a variável <i>tipo de verbo</i>	108
TABELA 11: A ordem VS no português afro-brasileiro, segundo a variável <i>tipo de frase</i>	110
TABELA 12: A ordem VS no português popular, segundo a variável <i>tipo de frase</i>	110
TABELA 13: A ordem VS no português popular, segundo a variável <i>presença de constituinte à esquerda</i>	111
TABELA 14: A ordem VS no português popular, segundo a variável <i>escolaridade</i>	113

LISTA DE FIGURAS

FIGURA 1: A polarização sociolinguística do PB.....	74
FIGURA 2: Transmissão Linguística Irregular.....	76
FIGURA 3: Disposição geográfica das comunidades.....	79
FIGURA 4: Distribuição das variedades linguísticas do PB, conforme o <i>continuum</i> de urbanização. FONTE: Bortoni-Ricardo (2006, p. 52).....	85
FIGURA 5: Disposição das comunidades rurais e afro-brasileiras no <i>continuum</i> de urbanização.	86

LISTA DE GRÁFICOS

GRÁFICO 1: Ordem VS na diacronia do PB, adaptado de Berlinck (1989, p. 97).	53
GRÁFICO 2: correlação de frequência de SV, retenção pronominal na posição de sujeito e objeto direto nulo, segundo Tarallo (1993, p. 84) e Berlinck (1989, p. 97). FONTE: Coelho (2006, p. 97)	55
GRÁFICO 3: A ordem VS segundo alguns estudos sociolinguísticos.	57
GRÁFICO 4: Ocorrência da ordem SV segundo o tipo de interrogativa <i>qu-</i> e o período de tempo. FONTE: Duarte (1992, p. 41).....	62
GRÁFICO 5: Resultado geral da variável dependente no português popular e no português afro-brasileiro.	94
GRÁFICO 6: A ordem VS nas declarativas em variedades cultas e populares do PB.....	95
GRÁFICO 7: A ordem VS no português afro-brasileiro e no português popular, segundo a variável <i>animacidade do SN sujeito</i>	101
GRÁFICO 8: A ordem VS no português popular e no português afro-brasileiro, segundo a variável <i>definitude do SN sujeito</i>	103
GRÁFICO 9: A ordem VS no português popular e no português afro-brasileiro, segundo a variável <i>faixa etária</i>	114
GRÁFICO 10: A ordem VS no português popular e no português afro-brasileiro, segundo a variável <i>estada fora da comunidade</i>	115

SUMÁRIO

INTRODUÇÃO	16
1 O ENQUADRAMENTO TEÓRICO	19
1.1 A SOCIOLINGUÍSTICA VARIACIONISTA.....	19
1.2 A PROBLEMATIZAÇÃO DA MUDANÇA LINGUÍSTICA	22
1.3 O ESTUDO DA VARIAÇÃO E DA MUDANÇA.....	28
1.3.1 Os efeitos das variáveis sexo e escolaridade no tratamento da variação e mudança.	33
1.3.2 O suporte estatístico-computacional.....	36
2 A INVERSÃO VERBO SUJEITO NO PORTUGUÊS BRASILEIRO.....	42
2.1 A INVERSÃO VS NO PORTUGUÊS BRASILEIRO: CARACTERIZAÇÃO GERAL	42
2.1.1 Os verbos inacusativos e a ordem VS	44
2.2 A ORDEM DOS CONSTITUINTES NO PORTUGUÊS: BREVE CONTEXTUALIZAÇÃO HISTÓRICA.....	49
2.2.1 A evolução da ordem VS no Português Brasileiro	52
2.3 A INVERSÃO VS NO PORTUGUÊS BRASILEIRO CONTEMPORÂNEO: CONDICIONAMENTOS ESTRUTURAIS	56
2.3.1 A inversão VS nas frases declarativas	56
2.3.2 A inversão VS nas frases interrogativas.....	60
2.4 A ORDEM VS/SV: CONDICIONAMENTOS SOCIAIS.....	64
2.5 SUMARIZANDO.....	66
3 O UNIVERSO DE OBSERVAÇÃO	67
3.1 A FORMAÇÃO DO PORTUGUÊS BRASILEIRO: CONTEXTO SÓCIO-HISTÓRICO.....	68
3.2 A REALIDADE LINGUÍSTICA DO PB COMO RESULTADO DE UM PROCESSO DE POLARIZAÇÃO SOCIOLINGUÍSTICA	72
3.2.1 A transmissão linguística irregular.....	74
3.3 O RECORTE DE OBSERVAÇÃO.....	77
3.3.1 O Projeto Vertentes	77
3.3.2 A constituição do corpus	78

3.3.3 O português afro-brasileiro: as comunidades remanescentes de quilombos.	80
3.3.4 O português popular: as comunidades rurais do interior do estado da Bahia.	83
3.4 A COLETA E A PREPARAÇÃO DOS DADOS	88
3.5 A VARIÁVEL DEPENDENTE	90
3.5.1 Dados descartados	91
3.6 A CHAVE DE CODIFICAÇÃO: ALGUMAS CONSIDERAÇÕES	92
4 ANÁLISE DOS DADOS	94
4.1 OS CONDICIONAMENTOS ESTRUTURAIS	96
4.1.1 Condicionamentos relacionados ao sujeito	96
4.1.2 Condicionamentos relacionados ao verbo	106
4.1.3 Condicionamentos no nível da oração	109
4.2 OS CONDICIONAMENTOS EXTRALINGUISTICOS	112
4.2.1 A variável escolaridade	112
4.2.2 As demais variáveis sociais	114
CONSIDERAÇÕES FINAIS	117
REFERÊNCIAS	120

INTRODUÇÃO

A história da formação do Brasil é marcada por um processo massivo de contato cultural e linguístico, sobretudo entre africanos e portugueses. Durante o período da colonização, deu-se o convívio entre as línguas africanas, ameríndias e portuguesa, gerando, a princípio, uma situação de multilinguismo generalizado. Os negros, graças à política do tráfico negreiro que vigorou entre os séculos XVI e XIX, obtiveram um papel de destaque, atuando como principais difusores do português vernáculo (ou popular) brasileiro (MATTOS E SILVA, 2004). Os escravos africanos, transplantados de sua terra e falando diversas línguas entre si, foram levados, pelas circunstâncias sócio-culturais que marcaram o cenário da colonização brasileira, a adquirir a língua do dominador, o português. Deu-se então uma aprendizagem de segunda língua de maneira precária. Conforme Lucchesi (2001), é esse modelo defectivo do português que servirá, em muitos casos, de base para aquisição da língua portuguesa dos descendentes dos escravos, dando início ao processo de *transmissão linguística irregular* que teria sido a principal motivação da constituição polarizada do português brasileiro (PB) (LUCCHESI, 2003; 2006) ¹.

Desse processo, teriam resultado algumas características que diferenciam o português brasileiro do português europeu, tais como a perda da morfologia flexional que, de modo indireto teriam motivado outras mudanças, como, por exemplo, a forte restrição à inversão verbo-sujeito (VS).

O português brasileiro (PB) apresenta um caráter bastante rígido no que tange a ordenação dos constituintes sentenciais *sujeito (S) e verbo (V)*. A forte restrição à ordem verbo sujeito (VS) nas sentenças declarativas seria, segundo Menuzzi (2003), um dos pontos de diferenciação entre o português europeu (PE) e a variedade nacional idioma. Isso também se mantém nas sentenças interrogativas QU-, nas quais, em PE, processa-se uma regra obrigatória de inversão, ausente no PB (DUARTE 1992; ROSSI, 1993; KATO, 2005).

O atual panorama linguístico reflete um quadro de mudança, já que, como aponta Berlinck (1988, 1989), a diacronia do português registra uma mudança considerável nas frequências de aplicação da ordem inversa VS. A observação de alguns estudos nos revela que este quadro de mudança encontra-se associado a outras alterações em itens linguísticos, tais

¹ Vale destacar que, no âmbito da linguística moderna, são admitidas outras hipóteses explicativas para a variação no PB. Acreditamos, no entanto, que considerar a expressiva influência do contato é hipótese mais coerente.

como a perda do sujeito nulo e a proeminência de objetos não realizados no PB (TARALLO, 1993).

Nessa perspectiva, nosso trabalho tem como objetivo descrever e analisar a ordem dos constituintes sentenciais, especificamente a posição do sujeito em relação ao verbo, no português afro-brasileiro e no português popular do interior do Estado da Bahia. Estas variedades do português encontram-se representadas em nossa pesquisa, respectivamente, pelos falantes das comunidades remanescentes de quilombo, Cinzento e Sapé, e por falantes das comunidades rurais do interior do estado da Bahia – Santo Antônio de Jesus e Poções.

Esta pesquisa busca responder à seguinte questão: quais seriam os condicionamentos linguísticos e extralinguísticos para a inversão *VS* (*verbo-sujeito*) em comunidades rurais e afro-brasileiras do interior do estado da Bahia, e qual a possível relação com o contato linguístico presente na gênese dessas variedades?

Levanta-se a hipótese de que o fenômeno da inversão, como fenômeno variável e inserido no sistema, é passível de ser descrito e sistematizado, sendo condicionado por fatores de natureza sintática (tipo de verbo, presença de constituinte à esquerda da sentença) semântico-discursiva (animacidade e *status* informacional do SN sujeito) e extralinguística (sexo dos informantes, faixa etária, escolaridade).

Para a análise do nosso objeto de estudo, adotamos o referencial teórico-metodológico da sociolinguística variacionista (LABOV, 2008 [1972]), no entanto, a natureza notadamente sintática do nosso objeto não nos permite ignorar algumas considerações extremamente importantes da gramática gerativa acerca da ordem dos constituintes e da estrutura sintática.

Nossa concepção de língua e a metodologia de estudo, no entanto, estão centrados no modelo Laboviano e constituem a discussão inicial do nosso trabalho. No capítulo 1, são apresentados os principais pressupostos teórico-metodológicos da sociolinguística variacionista, corrente teórica da linguística moderna à qual este trabalho está vinculado, bem como a concepção de língua do modelo e os conceitos basilares desta teoria, os quais serão recorrentemente tratados ao longo do trabalho.

De posse do enquadramento teórico do trabalho, o segundo capítulo – A inversão verbo sujeito no português brasileiro – apresenta um panorama do fenômeno em estudo, a ordem VS no PB, destacando, brevemente, a contextualização histórica e as principais concepções teóricas sobre o tema.

No capítulo 3, apresenta-se o universo de observação ao qual esta dissertação se debruça, a saber, o português popular do Brasil, mais especificamente as variedades rurais do

interior do estado da Bahia. Além de uma prévia apresentação dessa variedade do idioma e de sua constituição histórica, apresentamos e caracterizamos o recorte amostral aqui tomado para observação. O *corpus*, constituído no âmbito do Projeto Vertentes, é formado por quatro comunidades (Helvécia, Cinzento, Santo Antônio e Poções), devidamente apresentadas no capítulo. Também são dispostas as decisões metodológicas adotadas durante a elaboração do trabalho, desde a seleção dos informantes, até a delimitação do fenômeno de estudo.

O último capítulo, mas não menos importante, apresenta a análise dos dados. Para tanto, utilizamos o recurso estatístico computacional, VARBRUL. Nesse momento, apresentamos as variáveis selecionadas como estatisticamente relevantes pelo pacote de programas. Optamos por apresentar as variáveis estudadas, à medida que apresentamos os resultados, pois acreditamos que isto facilita o trabalho de interpretação dos resultados. Assim, cada tabela de resultados é precedida pela descrição da variável e seguida pela descrição e análise dos resultados quantitativos.

Em face à carência analítica da abordagem das GT, os estudos linguísticos têm se dedicado amplamente ao estudo do fenômeno da ordem, sobretudo no âmbito da gramática gerativa e do funcionalismo, buscando lançar luz aos questionamentos e dúvidas a respeito do fenômeno da inversão no PB. Embora seja um tema bastante discutido pelos estudos linguísticos (cf. BERLINCK, 1988; ALBERTON, 2001; COELHO, 2006; SPANO, 2008), a inversão VS carece de estudos empíricos, como atesta Menuzzi (2003). Assim, a realização deste trabalho justifica-se no enriquecimento da discussão sobre a natureza do tema proposto que se diferencia por sua abordagem empírica. Ademais, pode vir a contribuir como instrumento de descrição e análise do português popular falado no Brasil, na caracterização do português brasileiro.

1 O ENQUADRAMENTO TEÓRICO

Este capítulo tem o objetivo de apresentar os pressupostos e orientações metodológicas aos quais se vincula este trabalho, a saber, o modelo da **Teoria da Variação**, ou Sociolinguística Variacionista, proposto por Labov (2008 [1972]) e Weinreich, Labov e Herzog (2006 [1968]).

1.1 A SOCIOLINGUÍSTICA VARIACIONISTA

O termo *sociolinguística* surge no debate sobre a relação entre *status* social e comportamento linguístico. Em face ao amplo domínio da sociolinguística e à diversidade de estudos que tem sido tomados como sociolinguísticos, distinguem-se, segundo Monteiro (2002, p. 26), duas abordagens: uma macrosociolinguística, voltada para o estudo das relações entre sociedade e língua como um todo, sobretudo no que concerne às investigações sociológicas e políticas; e a microsociolinguística, alicerçada na análise das interferências sociais sobre a estrutura linguística. Com base em Romaine (1994), o autor argumenta que a diferença está no foco de observação das abordagens: a primeira teria como ponto de partida a sociedade, tomando a língua como um elemento de organização das comunidades; a segunda estaria centrada na língua, fazendo uso dos elementos do entorno social como fatores condicionantes na determinação das estruturas linguísticas. Nosso estudo está, portanto, vinculado à segunda abordagem, mais especificamente, ao modelo da sociolinguística variacionista.

William Bright foi um dos primeiros a tentar uma definição e delimitação do campo de atuação dessa nova ciência, através da organização da *Sociolinguistics Conference*, na Universidade da Califórnia em Los Angeles, em 1964, o qual contou, dentre outros, com a participação de William Labov com o tema “a hipercorreção como fator de variação”. Na introdução das atas do congresso, intitulado *As dimensões da sociolinguística*, Bright (1974 [1966]) sintetiza, como função da nova área de estudo, demonstrar que a variação não acontece de forma livre, mas antes encontra correlação com as diferenças sociais.

Dessa forma, a sociolinguística estaria preocupada com os empregos linguísticos concretos inseridos e influenciados pelo contexto social, do qual faz parte, de forma

sistemática e funcional, a variação linguística. A preocupação em enfatizar a contraparte social da língua nos estudos linguísticos fica evidente na resistência de Labov (2008 [1972], p. 216) ao termo: “Se não houvesse necessidade de contrastar este trabalho com o estudo da língua fora de seu contexto social, eu preferiria dizer que se trata simplesmente de linguística”.

A consolidação desse modelo se dá principalmente a partir dos estudos de William Labov sobre o inglês falado na Ilha de Martha’s Vineyard (1963) e a estratificação social do inglês falado na cidade de Nova York (1966). Esses estudos, bem como as reflexões desenvolvidas no *Empirical foundations for a theory of language change* (WEINREICH; LABOV; HERZOG, 2006 [1968]), inauguram uma nova mentalidade e método para a linguística moderna ao propor não somente uma nova concepção de língua, mas também a possibilidade de estudar a mudança em processo.

Ao propor a língua como um sistema diferenciado, provido de heterogeneidade e sistematicidade, o modelo sociolinguístico rompe com axiomas e fundamentos vigentes até a primeira metade do século XX. Até o modelo estruturalista, representado principalmente pelos estudos de Saussure (2006 [1916]), vigorava uma visão da língua como um sistema homogêneo, unitário e autônomo. De acordo com essa orientação teórica, o enfoque linguístico detinha-se sobre a *langue* (língua), compreendida como a abstração do sistema linguístico. Para o linguista genebrino, a *parole* (fala), a manifestação externa e concreta do sistema, não poderia ser concebida como objeto de estudo científico por ser considerada caótica e encontrar-se associada a fatores externos à língua.

A língua, portanto, concebida pelo estruturalismo como um sistema encerrado em sua lógica interna, não comportava as questões de variação, que eram prescindidas desse sistema abstrato. Ademais, não era considerada como objeto de discussão a contraparte histórica da língua, muito menos a influência do meio social. De acordo com Lucchesi (2004a, p. 157):

[...] o modelo teórico estruturalista não podia incorporar como objetos de sua reflexão sistemática nem a prática linguística concreta, nem o processo sócio-histórico de constituição da língua. Baniam-se, desse modo, todas as questões relativas ao caráter social e histórico da língua dos principais cenários teóricos da ciência da linguagem.

A posição estruturalista, em parte, é fruto do desprezo à influência do contexto social, conforme ressaltam Weinreich, Labov e Herzog (2006 [1968], p. 89), “a correlação negativa entre estruturalidade e comunicabilidade foi uma extrapolação perfeitamente natural para uma teoria linguística socialmente agnóstica”.

Ao propor a língua como um sistema homogêneo e encerrado em si mesmo, os estruturalistas se deparam com um paradoxo essencial: “se uma língua tem de ser estruturada, a fim de funcionar eficientemente, como ela funciona enquanto sua estrutura muda?” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 35).

Este paradoxo, de acordo com os autores supracitados, constitui a questão central a ser enfrentada por uma teoria da mudança linguística e sua superação encontra-se na quebra da identificação entre estruturalidade e homogeneidade. A heterogeneidade linguística seria parte da própria competência linguística do falante, de forma que, “(...) numa língua que serve a uma comunidade complexa (i.e., real), a ausência de heterogeneidade estruturada é que seria disfuncional” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 36). Perde-se, dessa forma, a noção de variação livre e se propõe uma abordagem da variação linguística dentro do próprio sistema, de modo que se torne possível a descrição e a identificação de contextos linguísticos e extralinguísticos (como *sexo, idade, origem étnica*, etc) que a condicionam.

Esta nova concepção desenvolve-se ancorada em numa nova definição do objeto de estudo da linguística. Com a teoria da variação, o olhar linguístico desloca-se da *langue* para a *parole*. Conforme destacam Weinreich, Labov e Herzog (2006 [1968], p. 13):

A base do conhecimento intersubjetivo na linguística tem de ser encontrada na fala – a língua tal como usada na vida diária por membros da ordem social, este veículo de comunicação com que as pessoas discutem com seus cônjuges, brincam com seus amigos e ludibriam seus inimigos.

O material de estudo da sociolinguística constitui, portanto, o vernáculo dos falantes de determinada comunidade. Em termos metodológicos, o que se sugere é um estudo baseado na observação empírica do desempenho linguístico de uma comunidade de fala². A partir desse tipo de observação, seria possível deslindar como os processos linguísticos operam de modo a conferir à língua estados de variação e mudança.

Tomando como pressuposto o princípio geológico do uniformitarismo, apresentado inicialmente por James Hutton (século XVIII), assume-se que, assim como os processos geológicos que modificaram a superfície terrestre hoje são os mesmos que atuaram no passado, no âmbito linguístico, “os mesmos mecanismos que operaram para produzir as mudanças em larga escala do passado podem ser observados em ação nas mudanças que

² Conceito retomado na seção 2.3.

presentemente ocorrem à nossa volta.” (LABOV, 2008 [1972], p. 192). Rompe-se, dessa forma, a estanque dicotomia saussuriana entre sincronia e diacronia.

As principais diferenças entre o estruturalismo e o modelo sociolinguístico encontram-se sintetizadas no Quadro 1, abaixo:

	Estruturalismo	Sociolinguista variacionista
<i>Objeto de estudo</i>	Língua	Fala (comunidade)
<i>Concepção de Língua</i>	Sistema autônomo, atemporal e homogêneo	Produto do processo histórico heterogeneamente estruturado
<i>A heterogeneidade na língua</i>	Variação livre, abstraída, externa ao sistema.	Funcional, condicionada por fatores internos e externos, interna ao sistema.
<i>Falante</i>	Passivo	Ativo

QUADRO 1: Estruturalismo X sociolinguística variacionista.

Vale salientar que não é propósito da sociolinguística desconsiderar a contraparte estrutural da língua, muito menos é motivação deste estudo ignorar ou rechaçar a pertinência dos estudos estruturalistas. Busca-se apenas destacar como a corrente teórica que orienta este trabalho se insere no cenário linguístico do século XX. Embora consistam em diferentes concepções da língua, recentemente tem crescido uma tendência de estudos que buscam associar a orientação metodológica sociolinguística, com uma abordagem estruturalista mais contemporânea, o gerativismo. Isto sugere que essas concepções teóricas não são de todo excludentes e têm, cada uma a seu modo, muito a contribuir para a compreensão do fenômeno da linguagem.

1.2 A PROBLEMATIZAÇÃO DA MUDANÇA LINGUÍSTICA

A mudança linguística sempre esteve na esteira do pensamento linguístico, no entanto, é na sociolinguística que se introduz uma metodologia suficientemente capaz de estudá-la enquanto processo. A aceção da variação como parte do sistema e o estudo da mudança em curso são as inovações basilares dessa nova teoria.

O estudo da mudança na linguística histórica restringia-se à descrição das mudanças já concluídas. Com essa abordagem, supriam-se os questionamentos a respeito da direção da mudança e seus condicionamentos universais. No entanto, não era possível identificar as

causas da mudança, muito menos definir seus mecanismos e sua “função adaptativa” (LABOV, 2008 [1972], p. 192).

De acordo com Coseriu (1979 [1921]), questionar o porquê de as línguas mudarem é um problema injustificável para a linguística. Para o autor, o equacionamento da mudança linguística em termos causais provém de um raciocínio das ciências naturais, não aplicável a um organismo social complexo como a língua.

As mudanças lingüísticas (...) só podem ser explicadas (motivadas) em termos funcionais e culturais. Mas as explicações culturais e funcionais das mudanças não são de maneira alguma “causais”. A própria idéia de “causalidade” na chamada “evolução” idiomática é um resíduo da velha concepção das línguas como “organismos naturais”, assim como do sonho positivista de descobrir as supostas “leis” da linguagem (ou das línguas) e de transformar a lingüística numa “ciência de leis” análoga às ciências físicas. (COSERIU, 1979 [1921], p. 155)

Essa questão, que persegue o pensamento linguístico já há algum tempo, encontrou diferentes posicionamentos, os quais, de acordo com Coseriu (1979 [1921], p. 163-164), podem ser classificados em três atitudes causais:

- (i) atitude “corajosa”: logram encontrar as causas externas da mudança linguística, ou mesmo sua causa fundamental ou universal. Para o autor, as condições das mudanças não podem ser tratadas em termos genéricos, uma vez que as evidências empíricas demonstram que elas são múltiplas “e não são em si mesmas operantes”;
- (ii) atitude “prudente”: Toma as causas da mudança como provisoriamente desconhecidas. Incorrem no erro de continuar postulando a existência de causas, que se ainda são desconhecidas, o são por limitações circunstanciais da teoria analítica. Dessa forma, as causas não seriam preditivas, mas sim dedutivas, de modo que o estudo empírico precedesse a teoria;
- (iii) atitude “conciliadora”: Defende que as causas da mudança seriam parcialmente conhecidas e que a porção desconhecida poderia vir a ser descoberta com base em futuras investigações. Para Coseriu, a junção das análises parciais não contemplaria a solução da questão universal da mudança, mas apenas as modalidades genéricas das mudanças.

Frente a essas três atitudes, surge a sociolinguística variacionista. A teoria retoma a questão da mudança em termos de relações de determinação, no entanto reconhece que, para

se compreender o fenômeno da mudança e seus mecanismos, é necessário responder a pelo menos cinco questões, os conhecidos cinco problemas da mudança linguística.

O primeiro deles, o **problema das restrições** (*constraints problem*), também conhecido como **problema dos fatores condicionantes**, diz respeito às limitações da mudança e busca identificar quais seriam as mudanças e condições possíveis e impossíveis (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 121). Esse problema pauta-se em uma visão da competência linguística universal e possui a implicação de conduzir a generalizações (LUCCHESI, 2004a, p. 173). A tendência de consoantes em contexto intervocálico tornarem-se vozeadas, e nunca desvozeadas, seria um exemplo desse tipo de restrição. No plano morfológico, podemos pensar, por exemplo, na generalização do tipo: quando ocorre uma mudança na marcação de número no sintagma nominal do PB, a marca sempre será mantida no determinante, perdendo-se nos demais constituintes: *As meninas bonitas => as menina∅ bonita∅*. Este problema, de acordo com Lucchesi (2004a), incorre numa busca por universais, visão incompatível com a concepção de língua do modelo. Ainda segundo o autor, esta teria sido a motivação para que, mais tarde, o próprio Labov descartasse esse problema, fundindo-o com o problema do encaixamento.

O segundo, o **problema do encaixamento** (*embedding problem*), sustenta-se na noção de que nenhuma mudança acontece no vácuo, mas antes se encontra encaixada nas estruturas linguística e social circundantes. Conforme Labov (2008 [1972], p. 193), o problema do encaixamento está associado à noção de covariação, seja esta com aspectos intra ou extralinguísticos. De acordo com Lucchesi (2004a, p. 175), a proposição de um encaixamento linguístico sustenta-se parcialmente na noção estruturalista de que “a mudança linguística só poderá ser compreendida considerando-se a sua inserção no sistema linguístico que ela afeta”. O encaixamento na estrutura social tem origem na própria concepção de língua do novo modelo. A esse respeito, Weinreich, Labov e Herzog (2006 [1968], p. 123) afirmam que “a estrutura linguística mutante está ela mesma encaixada no contexto mais amplo da comunidade de fala, de tal modo que variações sociais e geográficas são elementos intrínsecos da estrutura”. Embora tradicionalmente apresentado bifurcado, o próprio Labov (2008 [1972], p. 193) assinala que o encaixamento constitui um único problema:

Note-se que o problema do encaixamento é apresentado aqui como um problema único, apesar do fato de haver nele dois aspectos distintos: correlações dentro do sistema linguístico e com elementos de fora do sistema.

Isto pode ser exemplificado através do estudo realizado por Labov sobre a mudança nos ditongos do inglês falado na ilha de Martha's Vineyard, na costa oeste dos Estados Unidos. Conforme apresentado pelo autor, a centralização do ditongo (*aw*) em Martha's Vineyard estaria encaixada num processo mais amplo de mudança no sistema dos ditongos e teria sido precedida pela centralização do ditongo (*ay*), que por sua vez revela seu encaixe social ao apresentar-se como um traço de orientação cultural. A centralização, em certo plano, estaria ligada a uma atitude de resistência cultural por parte de seus habitantes em face da penetração dos veranistas na ilha.

Como o modelo baseia-se na análise da mudança em processo, é possível buscar resolver um terceiro questionamento: qual o percurso transcorrido por uma mudança? Denominado **problema da transição** (*transition problem*), essa questão busca identificar os estágios intervenientes de um processo de mudança de uma estrutura *A* para uma estrutura *A'* (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 122). A mudança que opera no português brasileiro no uso pronominal de 3ª pessoa plural (*nós* → *a gente*), por exemplo, teria se processado em estágios intermediários. Conforme destacam alguns estudos (OMENA, 1986), o uso do *a gente* na norma urbana teria sido iniciado nas funções sintáticas de sujeito, devido ao seu caráter nominal; e de adjunto adverbial, motivado pela perda da flexão de caso pronominal que afetou a forma *conosco*³. Só *a posteriori* teriam sido atingidas as demais funções de complementos verbais e, por último, adjuntos adnominais⁴. O processo de transição da forma *nós* para *a gente* apresentaria, portanto, alguns estágios intermediários de mudança que foram motivados por alguns condicionamentos linguísticos e sociais específicos.

Do ponto de vista da estrutura social, esse percurso pode ser deslindado a partir de um estudo sistemático das gerações sucessivas que compõem uma comunidade de fala. Isto porque “a transferência parece ocorrer entre grupos de pares de faixas etárias levemente diferentes” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 122).

A preocupação da teoria com o papel do indivíduo no processo de mudança está na base de uma quarta questão. O chamado **problema da avaliação** (*evaluation problem*) identifica o nível de consciência social como uma importante propriedade para a compreensão e determinação da mudança. Segundo Labov (2008 [1972]), há dois tipos de abordagens possíveis para este problema: uma direta e outra mais indireta. A primeira concentra-se na

³ Sobre a perda da flexão pronominal, ver Mendes (2009).

⁴ Entretanto, estudo sobre variedades rurais (BANDEIRA, 2008) tem apontado para um quadro diferenciado de mudança, com *a gente* sendo mais frequente nas funções de complemento (tanto quanto ou mais que a função de sujeito) e pouco frequente na função de adjunto adverbial.

correlação entre atitudes e desejos e o comportamento linguístico do informante. O desejo de *status* ou aceitação em uma comunidade de fala, por exemplo, pode levar o indivíduo a privilegiar ou rejeitar determinada forma em seu uso linguístico. A outra, mais indireta, encerra os mecanismos metodológicos de “medição” das reações subjetivas dos falantes frente ao emprego de determinada variante. Para este fim, podem ser aplicados alguns testes de percepção, que avaliam a reação do falante frente ao emprego de determinadas formas, ou de produção, que conduzem o falante à produção das formas. A partir desses tipos de teste, pode ser analisada a reação subjetiva dos falantes frente ao fenômeno linguístico e depreendidas situações de prestígio ou estigmatização das formas dentro de uma comunidade de fala. (TARALLO, 1985, p. 55).

É importante destacar que a avaliação do falante pode definir o rumo de uma mudança linguística. No PB, o problema da avaliação pode ser exemplificado com a realização do objeto direto anafórico. Diante das três formas possíveis de realização, a categoria vazia (1a) tem sido preferida como uma estratégia de neutralidade social, frente ao uso do pronome lexical *ele* (1b), socialmente estigmatizado, e do clítico acusativo (1c), tomando como alternativa pedante.

(1) Você viu o João?

- a. *Eu vi* Ø
- b. *Eu vi* *ele*
- c. *Eu o* *vi*

É necessário, no entanto, destacar que nem todo fenômeno linguístico está afeto ao juízo do indivíduo. Mesmo porque algumas mudanças se processam abaixo do nível de consciência do falante. Dessa forma, distinguem-se os fenômenos linguísticos como:

- I. *indicadores*: fenômenos que variam quanto à posição social dos indivíduos, mas não sofrem influência do nível de formalidade, de modo que se encontram abaixo do nível de consciência do falante. Ex.: Uso do pronome sujeito (*eu viajei/ viajei*);
- II. *marcadores*: variam quanto ao estrato social e estão mais ou menos abaixo do nível de consciência do falante, de modo que podem ser observados em testes de avaliação subjetiva. Ex.: emprego dos pronomes *nós* e *a gente* e a regência dos verbos de movimento (*vou na praia/ vou à praia*);

- III. *estereótipos*: formas linguísticas fortemente marcadas do ponto de vista social, sendo estigmatizadas pelos falantes/ ouvintes. Ex.: Concordância verbal (*Nós vai/ nós vamos*) na norma culta brasileira⁵.

O quinto problema proposto, o **problema da implementação** (*actuation problem*), inquire sobre o porquê de uma mudança linguística se instaurar em determinado espaço de tempo e lugar. É latente a dificuldade de se resolver essa questão: “A dificuldade do enigma da implementação é evidente no número de fatores que influenciam a mudança: é provável que todas as explicações a serem propostas no futuro próximo serão a *posteriori*” (WEINREICH; LABOV; HERZOG, 2006 [1968] p. 124). No entanto, os autores sugerem que, a partir do estudo de casos particulares de mudança, essas explicações possam ser acuradas e reunidas numa abordagem mais abrangente do processo de mudança. A partir desse tipo de observação, sugere-se que a mudança linguística implementa-se da seguinte forma:

- 1º) Origem: uma forma alternativa é aprendida e passa a ser empregada por um grupo restrito de pessoas;
- 2º) Propagação: duas formas, a antiga e a inovadora, passam a contrastar durante o processo de interação social;
- 3º) Término: por motivações de natureza estrutural e social, uma das formas é suplantada pela outra. (WEINREICH; LABOV; HERZOG, 2006 [1968] p. 124; LABOV, 2008 [1972] p. 152).

Paiva e Duarte (2006) defendem que a questão da implementação está vinculada à questão dos fatores condicionantes, haja vista que “na medida em que identificamos os condicionamentos que agem sobre a mudança, adiantamos uma explicação da forma como a mudança vai se expandindo por diferentes contextos estruturais”. A fim de sustentar esta argumentação, exemplificam com o caso da perda de sujeito nulo no PB, o qual teria o entendimento da sua implementação associado à compreensão da relação com a acessibilidade e à hierarquia de animacidade.

⁵ Acreditamos que na norma popular este fenômeno encontra-se não marcado estando mais abaixo do nível de consciência dos falantes, de forma que pode ser caracterizado como um *marcador*.

No que se refere à hierarquia de referencialidade, por exemplo, podemos supor que a mudança progride numa direção que vai de sujeito [+referencial, +animado] para sujeito [-referencial, -animado] até atingir finalmente sujeitos com traço [-referencial]. Essa hierarquia vai explicar porque os sujeitos com o traço [-referencial], como é o caso de sujeitos expletivos nulos, são os últimos a se tornar lexicalmente preenchidos. (PAIVA; DUARTE, 2006, p. 143)

Como bem podemos observar no exemplo de Paiva e Duarte, que nos parece mais um exemplo para o problema da transição, o que se tem é mais uma descrição dos mecanismos de mudança do que uma explicação de como esta se implementa.

A propósito dessa questão, vale destacar o posicionamento de Lucchesi (2004a). Segundo o autor, o modelo sociolinguístico “identifica a explicação da mudança com a descrição dos seus mecanismos de implementação, ou seja, a descrição dos mecanismos de causa e efeito que constituem o processo de mudança” (LUCCHESI, 2004a, p. 179). Dessa forma, a questão da implementação no modelo sociolinguístico sugere uma discussão sobre o que seria de fato produzir uma explicação. Além disso, o autor aponta um conflito nessa questão e nos remonta à discussão de Coseriu (1979 [1921]), destacada no início desta seção, a respeito da contradição entre perspectiva causalista e a concepção de língua da sociolinguística como um objeto “social-histórico e cultural”.

Estes questionamentos só podem ser respondidos através de um estudo que contemple a mudança enquanto ela ainda está acontecendo e que seja capaz de sistematizar o aparente caos da variação.

1.3 O ESTUDO DA VARIAÇÃO E DA MUDANÇA.

A fim proceder a uma análise capaz de contemplar os domínios da variação e da mudança, é preciso, primeiramente, definir o que é a variação linguística e como ela se organiza, ou seja, como são delimitados os conceitos de *variável* e *variante linguística*.

Os falantes de uma língua têm a seu dispor diferentes formas linguísticas para expressar um referente. Dessa forma, podemos dizer que um fenômeno linguístico encontra-se em variação quando apresenta duas ou mais formas de ser expresso no mesmo contexto, conservando seu significado referencial.

Tomando o exemplo do português brasileiro, temos a referência pronominal de 2ª pessoa, a qual pode ser expressa pelos pronomes *tu* ou *você*. Obviamente, as duas são controladas por regras variáveis, obedecendo a condicionamentos específicos. No entanto, carregam o mesmo significado referencial, ou seja, indicam a pessoa a quem o falante se dirige: ***Tu foi na festa? / Você foi na festa?***

Nesse sentido, Labov define como mesmo significado referencial a condição que formas linguísticas possuem de expressar o mesmo estado de coisas (*state of affairs*). Nas palavras do autor, “(...) duas sentenças que se referem ao mesmo estado de coisas possuem o mesmo valor de verdade⁶” (LABOV, 1978, p. 2, tradução nossa).

Tomando de empréstimo o jargão da estatística, designa-se o fenômeno em variação como *variável dependente* e suas diferentes formas de realização como *variantes*. No exemplo a seguir, *a ordem do sujeito em relação ao verbo* é a variável dependente e tem como variantes as ordens direta SV (i) ou inversa VS (ii).

- Ordem do sujeito em relação ao verbo:
 - (i) Ordem direta (SV): ... ***a mãe de Ana Maria morreu*** (POR-12)
 - (ii) Ordem inversa (VS): ...***há dois anos atrás, morreu meu sogro.*** (POS- 06)

Vale ressaltar que o termo *variável* pode ainda designar os grupos de fatores (linguísticos e extralinguísticos) que condicionam a realização das variantes (MOLLICA; BRAGA, 2004). Nesse caso, são denominadas *variáveis independentes* ou *explanatórias*. As variáveis linguísticas *independentes* ou *explanatórias*, por exemplo, são grupos de fatores de natureza estrutural que podem estar relacionados à seleção de uma das formas em questão. Tem-se, ainda, como *variáveis independentes* ou *explanatórias*, as variáveis discursivas e sociais.

Tradicionalmente, a noção de variação linguística tem sido aplicada com maior aceitabilidade aos níveis fonológico e morfológico, apresentando alguma resistência para fenômenos dos níveis sintático e semântico/pragmático (LAVANDERA, 1977).

Quando um falante diz *chegaram as cartas* ou *as cartas chegaram*, refere-se ao fato de existirem cartas e que estas chegaram em algum lugar. Do ponto de vista cognitivo, temos a mesma informação.

⁶ “two utterances that refer to the same state of affairs have the same truth-value (...)”

Não desprezamos que as variantes de uma variável, sobretudo as sintáticas, podem apresentar diferenciações quanto à avaliação social e estilística ou mesmo pragmáticas. A este propósito, o próprio Labov afirma que “[...] as variantes são idênticas em seu valor de verdade ou referencial, mas se opõem em sua significação social e/ou estilística” (LABOV, 2008 [1972], p. 313). Isto, aliás, já está previsto pela teoria, no problema do encaixamento (cf. seção 1.2). Dessa forma, acreditamos que a noção de variação pode sim ser aplicada ao nível sintático e que as diferentes nuances discursivas podem ser tratadas como possíveis condicionamentos para a variável dependente em estudo⁷.

De posse dessas explicações preliminares, partamos para os pressupostos metodológicos do modelo de análise da variação e da mudança. Concebendo a variação e a mudança como processos imanentes às línguas, e tendo o problema delimitado, como seria possível, então, estudar a variação e a mudança quando estão correndo?

Baseado numa orientação essencialmente empírica, o estudo sociolinguístico se desenvolve a partir da observação dos usos linguísticos de uma comunidade de fala. Assim, para a sociolinguística, é a comunidade, e não o indivíduo, o universo de observação do linguista. A noção de comunidade de fala então adotada perpassa pela assunção de juízos linguísticos partilhados por uma comunidade.

Nesse sentido, comunidade de fala define-se não como um grupo de falantes que partilha o mesmo código, ou utiliza as mesmas formas, mas sim como um grupo que se identifica pela mesma reação subjetiva frente a uma variedade lingüística (LABOV, 2008 [1972] p. 188).

Mais do que a comunidade de fala, é a amostra de fala vernácula desse universo que servirá de material para análise linguística. Esta se define como o registro mais espontâneo e informal do falante, ou seja, a fala mais irrefletida e menos monitorada, na qual os indivíduos estão mais atentos ao conteúdo emitido do que à forma. Salvo as limitações de se coletar uma fala essencialmente vernácula, dificuldade consagrada no âmbito da sociolinguística como o *paradoxo do observador*⁸, cabe ao linguista a constituição de uma amostra representativa do universo linguístico em estudo. A partir dessa amostra, serão depreendidos os resultados linguísticos generalizáveis à comunidade de fala.

⁷ As variáveis dependentes do fenômeno em estudo encontram-se desenvolvidas no capítulo 3 desta dissertação.

⁸ Diz respeito à situação paradoxal de o documentador ter que coletar a fala vernácula numa situação em que sua inevitável presença pode ativar o monitoramento do falante e inibir a captação de uma fala verdadeiramente espontânea e irrefletida.

De acordo com Almeida e Campoy (2005, p. 60), não há consenso sobre o tamanho da amostra, pois do mesmo modo que uma amostra muito grande pode ser redundante e tornar o estudo lento, fazer inferências com base em dados de poucos informantes também apresenta seus prejuízos para a análise. De um modo geral, devem ser consideradas, além das características socioculturais e a dimensão da comunidade, a influência dos fatores sociais no fenômeno em variação.

A partir da recolha de dados na amostra, podem ser identificadas as situações de variação e mudança. O que se propõe é que, a partir do estudo de um recorte sincrônico, podem ser depreendidos os desdobramentos de uma situação de variação. Isto porque, embora uma mudança linguística pressuponha um estado pretérito de variação, nem toda situação de variação desencadeará um processo de mudança.

Nesse sentido, Labov (2008 [1972], p. 194) afirma que:

Os dados mais simples para se estabelecer a existência de uma mudança linguística são um conjunto de observações de duas gerações sucessivas de falantes – gerações de características sociais comparáveis que representam estágios na evolução da mesma comunidade de fala.

Essa metodologia, conhecida no modelo como o construto do **tempo aparente**⁹ baseia-se na **hipótese clássica**, segundo a qual, um falante encontra sua gramática estabilizada na fase da puberdade. Assim, um indivíduo de 25 anos apresentaria um estado de língua de 10 anos atrás, enquanto um indivíduo de 70 anos, representa a fala de 55 anos atrás. Essa estabilidade gramatical do indivíduo tem sido questionada por alguns estudos, no entanto defende-se que possíveis mudanças no falante seriam esporádicas (NARO, 2004), ou atingiriam estruturas da língua mais flexíveis. Acredita-se, portanto, que uma amostra organizada em níveis geracionais refletiria estados diferentes de língua, correspondentes a sincronias sucessivas. O estudo em tempo aparente é uma projeção de estágios de mudanças através da correlação com o comportamento linguístico das faixas etárias e pode indicar dois diagnósticos. Se não há entre a faixa mais nova e as faixa mais velhas uma diferenciação quanto ao uso privilegiado da forma inovadora, podemos inferir que as variantes convivem paralelamente e podem se manter dessa forma sem que haja uma mudança, tem-se um caso de *variação estável*. Por outro lado, se a faixa mais jovem apresenta um uso predominante da

⁹ No âmbito da sociolinguística, podemos também desenvolver estudos a partir da comparação de duas sincronias separadas por um lapso de tempo. Esta metodologia, designada estudo em *tempo real*, pode comparar sincronias distantes (tempo real de longa duração) ou sincronia próximas (tempo real de curta duração).

forma inovadora, decrescendo este uso conforme eleva-se a faixa etária, diz-se que estamos diante de um processo de *mudança em curso*. Este trabalho, bem como os demais desenvolvidos no âmbito do Projeto *Vertentes*, desenvolvem-se em tempo aparente.

Reconhecemos, no entanto, que este constructo apresenta alguns pontos que merecem alguma discussão, haja vista sua excessiva generalidade e a questionável estabilidade do falante. Além disso, o que pode aparentemente ser uma projeção de mudança pode, na verdade, traduzir-se em caso de gradação etária, como revelam alguns estudos em tempo real de curta duração.

No âmbito da sociolinguística, os estudos em tempo real desenvolvem-se a partir a partir da comparação de duas sincronias separadas por um lapso de tempo. Esta metodologia pode comparar sincronias distantes (tempo real de longa duração) ou sincronia próximas (tempo real de curta duração). Este último compara duas sincronias separadas por um período discreto de tempo e pode ocorrer no modelo painel, no qual são comparados os mesmos falantes em dois momentos da vida (junventude e velhice, por exemplo), ou tendência, no qual são utilizados indivíduos da mesma comunidade em diferentes estágios de vida.

Nesse sentido, Labov (2001, p. 76) destaca que do estudo variacionista podem ser identificados, na verdade, quatro diagnósticos:

(1) Estabilidade (<i>stability</i>)	(2) Gradação etária (<i>age-grading</i>)	(3) Mudança geracional (<i>generation change</i>)	(4) Mudança comum (<i>communal change</i>)
O indivíduo e a comunidade se mantêm constantes.	O indivíduo muda, mas a comunidade se mantém constante.	O indivíduo se mantém constante, mas a comunidade, como um todo, muda.	O indivíduo e a comunidade mudam juntos.

QUADRO 2: Variação e mudança: relação indivíduo/ comunidade.

Sobre mudança no indivíduo, entende-se o grupo de falantes de determinada faixa etária que podem ter seu comportamento linguístico modificado ou não em relação ao uso das variantes. Quanto à comunidade, entende-se como os padrões de comportamento observáveis em conjunto, tendo em vista que “na medida que um falante vai mudando de faixa etária, muda a distribuição das variantes na comunidade” (NARO, 2004), de forma a alterar a configuração desse conjunto.

Na situação (1), as variantes se mantêm sem que haja superação de uma das formas, numa relação de contemporização (TARALLO, 1985). Nesse caso, tanto a comunidade, quanto os indivíduos mantêm seu comportamento, não apresentando tendências de privilégio

de nenhuma das formas; temos, portanto, um quadro de variação estável. A situação (2) também constitui um quadro de variação estável, pois o uso das variantes na comunidade se mantém, no entanto, há uma mudança de comportamento no nível do indivíduo: a medida em que os falantes envelhecem, assumem o comportamento linguístico inerente a determinada faixa etária, seja pelo aumento ou diminuição das pressões sociais. Em (3) e (4), tem-se situações de mudança, no entanto sutilmente diferenciadas. Quando um indivíduo, no transcurso das gerações, sofre mudança no seu comportamento, passando a privilegiar uma forma em detrimento de outra, ao passo que a comunidade como um todo segue esta mesma tendência, como em (4), dizemos estar diante de um processo de mudança acelerado pela ação do falante. Mesmo que o falante permaneça estável, frente à mudança da comunidade, como em (3), temos também um processo de mudança em curso, porém mais lento em relação à anterior.

Por esse motivo, uma observação simplista dos padrões de comportamento das faixas etárias nem sempre pode conferir um resultado claro quanto ao panorama de mudança. Nesse caso, quando possível, cabe conjugar os dois tipos de pesquisa, o estudo em tempo real e em tempo aparente, a fim de mitigar os possíveis equívocos de interpretação. Quando isso não se torna possível, podemos ainda recorrer aos demais fatores sociais, a fim de definir a caracterização dos papéis e pressões sociais exercidos no seio da comunidade e, a partir deste ponto, tentar promover uma análise mais acurada dos resultados. Esta tem sido a perspectiva adotada nas pesquisas desenvolvidas no Projeto *Vertentes*, ao qual este trabalho encontra-se vinculado.

1.3.1 Os efeitos das variáveis sexo e escolaridade no tratamento da variação e mudança.

Além da variável faixa etária, a teoria variacionista levanta importantes questões sobre o papel do sexo e da escolaridade do informante, além de outros fatores sociais no estudo da variação e da mudança. Os efeitos dessas variáveis sociais podem indicar orientações de mudança a depender das características sócio-culturais da comunidade de fala. Os níveis fonético-fonológico e morfológico têm se mostrado bem mais sensíveis aos condicionamentos extralinguísticos do que o nível sintático, o qual, aparentemente, está mais associado aos condicionamentos estruturais. Vejamos como algumas dessas variáveis podem contribuir no diagnóstico da mudança em curso.

1.3.1.1 A variável sexo

Os estudos variacionistas têm destacado a pertinência da variável sexo na configuração das variáveis. Embora haja divergências quanto à tendência de cada sexo em uma situação de variação, tem sido atestado um comportamento sensivelmente diferenciado entre homens e mulheres.

Fernández (1998), apontando os argumentos de Chambers e Trudgil, demonstra que as mulheres teriam uma tendência ao uso de formas de prestígio, que coincidem com as formas da *norma padrão* nas sociedades que tem uma normatização linguística mais explícita. Isso, no geral, é o resultado de uma situação de opressão a que as mulheres estão submetidas na grande maioria dos agrupamentos humanos e de uma busca das mulheres por *status* social. O autor destaca, ainda, como justificativa para este comportamento, o fato de as mulheres receberem uma educação mais coercitiva, voltada ao cumprimento e adequação às normas sociais. Assim, nesse contexto, é assinalado um comportamento conservador das mulheres.

Labov (2001), por sua vez, discute o comportamento linguístico feminino em termos relativos, apontando a existência do *princípio do conformismo linguístico das mulheres*, segundo o qual elas, nas situações de variação estável, tenderiam a utilizar mais as variantes de prestígio e evitar as formas estigmatizadas. Nas mudanças de cima para baixo, orientadas em função do padrão linguístico das classes mais altas, as mulheres tenderiam a utilizar as variantes de prestígio; nas mudanças de baixo para cima, cuja orientação da mudança parte dos grupos centrais ou baixos, tenderiam a usar a forma inovadora, que, nesse caso, geralmente corresponde à forma não padrão.

De acordo com Fernández (1998, p. 35 [tradução nossa]), “o sexo pode mostrar-se, portanto, mais como um fator de segunda ordem, como algo que costuma subordinar-se a dimensões sociais diferentes e com maior poder de determinação¹⁰”. Corrobora com esse raciocínio a afirmação de Lucchesi (2004, p. 192):

Generalizações do tipo ‘as mulheres são mais inovadoras do que os homens’ (...) tem a meu ver, um valor heurístico bastante questionável, pois a ação de um determinado fator social sobre um processo particular de mudança é determinada pela maneira específica através da qual esse fator se integra no

¹⁰ *El sexo puede mostrar-se, por tanto, mas como um factor de segundo orden como algo que suele subordinar-se a dimensiones sociales diferentes y con mayor poder de determinación.*

conjunto complexo de interações que constitui o processo social em que a mudança acontece.

Isto fica claro quando identificamos um sensível contraste no comportamento linguístico de homens e mulheres em zonas urbanas e rurais do Brasil. De um modo geral, os estudos sociolinguísticos de centros urbanos apontam as mulheres como inovadoras, no sentido de seguirem uma orientação para o prestígio social. Por outro lado, quando observamos os estudos em áreas rurais, como, por exemplo, nas comunidades afro-brasileiras do Estado da Bahia, estudadas pelo Projeto Vertentes, percebemos que os homens têm se mostrado mais inovadores, liderando os processos de mudança em direção ao padrão urbano culto (LUCCHESI, 2009). As mulheres, por sua vez, apresentam um comportamento mais conservador, utilizando mais as variantes locais, as quais, em geral, são não-padrão e estigmatizadas ou neutras.

Este comportamento justifica-se pela configuração social dessas comunidades. Em geral, os homens desses locais, em virtude da busca por atividade laboral, tendem a sair mais da comunidade, entrando em contato com falantes de outras variedades do português. As mulheres, por sua vez, mantêm-se mais ligadas às atividades domésticas e pouco saem da comunidade, o que justificaria o comportamento mais conservador do falar local.

Nas comunidades de Santo Antônio de Jesus e Poções, algumas mulheres entrevistadas costumam sair da comunidade, em geral em busca de tratamento médico, no caso das mais velhas, ou de emprego, sobretudo como empregadas domésticas, no caso das mais novas. Isto revela um comportamento linguístico particular das mulheres dessas comunidades.

1.3.1.2 A variável escolaridade

A relação do fator escolaridade com a variação e a mudança linguística tem sido alvo de inúmeras pesquisas sociolinguísticas. A proeminência concedida ao fator nos estudos está relacionada à identificação do seu papel na aquisição das variantes de prestígio. Tal interferência sustenta-se no fato de que “a escola incute gostos, normas, padrões estéticos em conformidade de dizer e escrever” (VOTRE, 2004, p. 51). Assim, a exposição à escolarização, na medida em que atua em favor das variantes de prestígio, opera também de modo coercitivo

sobre as variantes socialmente estigmatizadas. Em paralelo ao processo de variação e mudança, pode-se inferir que falantes mais afetados pela escolarização tenderiam a privilegiar mudanças em favor das variantes de prestígio, mais próximas do padrão normativo.

Obviamente, a ação escolar se faz mais direta nos fenômenos que constituem objeto de ensino ou sofrem coerção social, e, por tanto, mais susceptíveis à ação normatizadora, como, por exemplo, a aplicação das regras de concordância verbal e nominal. Fatores não estigmatizados, como o uso do modo imperativo; ou neutros do ponto de vista da avaliação social, como a realização nula do pronome sujeito tendem a sofrer pouca ou nenhuma influência da variável em questão.

Ademais, como os outros fatores, a análise da escolaridade deve levar em conta as determinações sócio-culturais do contexto. Isto porque as condições de acesso e os níveis de escolaridade dos falantes do meio urbano são bem maiores do que no meio rural. Mesmo os níveis de escolaridade apresentam valorações diferenciadas, a depender do contexto e do período analisado. A influência exercida pela escola no sentido de estimular as variantes da norma padrão não pode ser avaliada em função apenas do tempo, mas também da qualidade do ensino e do contexto sócio-histórico em que atua.

Indiscutivelmente, o contato com a escolaridade, ou em sentido mais amplo com o letramento, constitui ferramenta de acesso às formas padrão, haja vista a veiculação de textos escritos, além da ação normatizadora, seja pela ação dos professores e do material didático, seja pela atuação dos pares. Assim, a escolaridade pode atuar como difusor das variantes linguísticas prestigiadas e interferir no sentido da mudança em curso.

1.3.2 O suporte estatístico-computacional

A sociolinguística laboviana tem se desenvolvido com uma forte vinculação à ferramenta estatística, fato que levou o modelo a ser conhecido como *sociolinguística quantitativa*. Na busca da resposta à questão de como a variação e a mudança operam, o suporte estatístico tem servido de base para o estudo de diversos fenômenos linguísticos. No entanto, vale ressaltar que o recurso estatístico não pode ser usado como lei para a compreensão de qualquer fenômeno linguístico, mas sim como um auxílio, conforme assinalam Guy e Zilles (2008, p. 32),

Deve-se observar que toda estatística inferencial deve ser vista como ferramenta para ajudar no processo de descoberta do pesquisador, na testagem de hipóteses e assim por diante, em vez de ser considerada como “prova” (ou refutação) definitiva de uma hipótese.

Dessa forma, é o linguista o principal responsável pela análise e interpretação dos dados gerados pelo suporte estatístico computacional. Dentro dessa perspectiva, o programa VARBRUL (PINTZUK, 1988) tem sido usado eficientemente nas análises linguísticas do português brasileiro e serviu de suporte para a nossa pesquisa.

O pacote de programas das regras variáveis – VARBRUL – tem servido de suporte metodológico para o tratamento e análise de processos de variação e mudança na sociolinguística. Esse programa funciona com base no modelo matemático misto ou logístico. Embora esse não seja o único modelo disponível – antes desse, figuram outros como o modelo aditivo (LABOV, 1969) e o modelo multiplicativo de aplicação (CEDERGREN; SANKOFF, 1974) –, tem-se mostrado o mais eficaz para análises linguísticas e é o mais utilizado e indicado pelos linguistas.

O VARBRUL funciona a partir da formulação de regras variáveis, “um tipo de análise multivariada” que tem como objetivo “separar, quantificar e testar a significância dos efeitos de fatores contextuais em uma variável linguística.” (GUY; ZILLES, 2007, p. 33-34).

O uso do método multivariado na análise linguística justifica-se pela natureza desse tipo de problema/fenômeno. A análise de fenômenos linguísticos revela a atuação concomitante de vários fatores de ordem linguística e extralinguística. Dessa forma, uma análise univariada, que trate separadamente essas influências, não comporta a complexidade e ação em cadeia dos fatores sobre o fenômeno em questão. A adoção de uma análise multivariada, por sua vez, é capaz de “mensurar” a influência mútua dos fatores e “dará resultados mais precisos, porque ao mesmo tempo em que computa o efeito de uma variável independente, ela controla explicitamente o efeito de todas as outras variáveis independentes conhecidas” (GUY; ZILLES, 2007, p. 34).

Os dados gerados pelo pacote fornecem, em termos probabilísticos, as inferências para a identificação dos prováveis condicionamentos estabelecidos pelo linguista. Para tanto, são usadas medidas de confiabilidade¹¹, tais como *input*, *log-likelihood*, nível de significância (*threshold*) e peso relativo. Cada uma dessas medidas desempenha um papel específico na formulação dos resultados finais. O *input* corresponde à média global de uso de um valor da

¹¹ Para mais informações, consultar: Guy e Zilles, 2007 e Scherre, 1993.

variável dependente e “funciona como um ponto de referência para o fenômeno variável” (NARO e SCHERRE, 2004, p. 165). É feita a comparação do *input* geral com os gerados em cada um dos fatores, de modo que um desvio em referência à taxa geral pode ser interpretado como um desequilíbrio dos dados em função dos fatores (GUY e ZILLES, 2007, p. 238). Os valores dos *log-likelihood* e dos níveis de significância são os responsáveis pela indicação de determinado fator como influente no condicionamento da variável dependente. Convencionalmente, define-se uma margem de no máximo 0,05 para o nível de significância, de forma que os fatores que apresentam ao longo da rodada índices superiores a este valor (>0,05) são indicados como irrelevantes no condicionamento do fenômeno em estudo. Em síntese, um nível de significância <0,05 equivale à possibilidade de se negar a *hipótese nula*¹², de modo que essa hipótese teria menos de 5 chances em 100 de ser verdadeira, o que indicaria o grupo de fatores como relevante no condicionamento do fenômeno. O peso relativo refere-se especificamente aos valores das variáveis independentes. Na análise dos fatores que condicionam a ordem pós-verbal do sujeito, por exemplo, tomemos o grupo de fatores animacidade do sujeito (tendo como valores os traços [-] ou [+] animado) como um fator selecionado como significativo pelo VARBRUL. O peso relativo de cada valor indicará qual fator favorece e qual desfavorece a aplicação da ordem pós-verbal. Em análise de variáveis dependentes binárias define-se o ponto neutro como 0,50 (GUY; ZILLES, 2007, p. 239). Os valores acima ou abaixo, respectivamente, favorecem ou desfavorecem a aplicação da regra em questão. Assim, na Tabela 1 abaixo, o traço [+animado] favorece a colocação pós-verbal do sujeito com peso relativo 0,80, bem acima do ponto neutro. Enquanto o traço [-animado] desfavorece a aplicação da regra, com peso 0,43. Valores iguais ou muito próximos a 0,50 indicam que o fator não exerce nenhum tipo de influência sobre a variável.

TABELA 1: Realização pós-verbal do sujeito na comunidade de Poções, segundo a variável *animacidade do sujeito*.

Animacidade	Oco./total	Frequência	Peso relativo
[+ animado]	28/1068	3%	0,43
[- animado]	29/132	22%	0,80
Total	57/1200	5%	

Os dados numéricos fornecidos pelo programa não têm em si valor linguístico algum. Cumpre destacar que “os número selecionados ou eliminados só tem valor estatísticos: o

¹² “sempre afirma que nada está acontecendo, que a fonte da distribuição é normal, que as variáveis independentes não influenciam as variáveis dependentes” (GUY; ZILLES, 2007, p. 86).

significado lingüístico é de total responsabilidade do pesquisador”. (SCHERRE, 1993, p. 27). O suporte consiste em um auxílio e o êxito da análise probabilística depende em grande parte da qualidade dos dados e da percepção e conhecimento aprofundados do linguista. Além dos procedimentos elementares de definição da variável e suas variantes, bem como a identificação de hipotéticos contextos condicionadores, cabe ao pesquisador garantir a composição de uma base de dados consistente para a submissão ao VARBRUL. Para tanto, devem ser cuidadosamente observados algumas critérios na formulação do modelo analítico, tais como a ortogonalidade, exaustão e exclusividade.

Ao definir os fatores condicionadores a serem investigados, deve-se estar atento à independência entre os grupos de fatores. Em outras palavras, os grupos de fatores devem ser ortogonais, podendo coocorrer livremente, de modo que não constituam subcategorias uns de outros (GUY; ZILLES, 2007, p. 52). Quando isso não ocorre, o VARBRUL pode gerar resultados imprecisos ou não-convergentes, o que prejudicará a análise do fenômeno.

Além disso, os fatores de um grupo devem abarcar, tanto quanto possível, todas as possibilidades de cada contexto, de modo que seja alcançada a exaustão. Vale ressaltar que a formulação de um grupo de fatores nem sempre chega à exaustão na primeira observação. Ao longo do estudo, podem ser identificados e inseridos novos fatores na análise. A aplicação dos fatores a cada ocorrência tem de obedecer, também, ao critério da exclusividade, de maneira que cada ocorrência deve ser identificada com apenas um fator de um grupo.

Observando essas orientações, cabe ao linguista a formulação de uma chave de codificação com os prováveis fatores condicionantes. Os dados coletados na amostra devem ser submetidos a uma rigorosa codificação, para posterior submissão ao programa.

O VARBRUL, na verdade, constitui um pacote de programas que, desempenhando funções diferenciadas, operam conjuntamente para obtenção das regras variáveis. Alguns aplicativos desempenham funções básicas de preparação de dados para a obtenção dos valores probabilísticos (*CHECKTOK*, *READTOK*, *MAKECEL*). Outros, como o *CROSTAB*, *TEXTSORT*, *TSORT*, oferecem funções opcionais, como cruzamento de variáveis e procura e reordenação de dados. A realização de uma rodada obedece a um fluxo contínuo de preparação dos dados e segue a ordem esquematizada no quadro a seguir:

ENTRADA	PROGRAMA	SAÍDA
ARQ.ESP + ARQ.TXT Dados codificados	<i>CHECKTOK</i>	ARQ.COR dados corrigidos
ARQ.COR Dados corrigidos	<i>READTOK</i>	ARQ.OCO dados agrupados

ARQ.OCO + ARQ.CON Dados agrupados	MAKECEL	ARQ.CEL Percentagens
ARQ.CEL Percentagens	IVARB/ TVARB/ MVARB/VARB2000	ARQ.VAR probabilidades (pesos relativos)

QUADRO 3: Visão de conjunto da operacionalização do VARBRUL (adaptado de SCHERRE, 1993, p. 6).

Como é possível observar no Quadro 3, a operação da ferramenta estatística segue uma ordem linear de execução, na qual cada programa desempenha uma tarefa a fim de preparar os dados para a etapa seguinte. Vejamos uma descrição sumária de cada etapa:

1º) *CHECKTOK*: Verifica erros mecânicos de codificação nos arquivos de dados, comparando o conteúdo das cadeias de codificação do arquivo de dados com a lista de fatores definidos para cada grupo de fatores (arq.esp). Os arquivos corrigidos gerados pelo *checktok* serão usados como entrada para a etapa seguinte.

2º) *READTOK*: A partir da leitura dos arquivos de dados corrigidos, cria o arquivo de ocorrências. Neste arquivo, constam apenas as cadeias de codificação agrupadas.

3º) *MAKECEL/MAKE3000*: Cria o arquivo de células para a entrada do *IVARB*, *TVARB*, *MVARB*, *VARB200*. Arquivos gerados pelo programa que apresentem *KNOCKOUT FACTORS* (comportamento categórico das variantes) ou *SINGLETON GROUPS* (ocorrência de um único fator dentro de um grupo de fatores) não são aptos para entrada nos programas de regras variáveis. Esses problemas podem ser resolvidos através da rodada com um novo arquivo de condições modificado, com a amálgama ou anulação de fatores ou eliminação do grupo problemático. No arquivo de células gerado pelo programa, são apresentados os resultados percentuais de cada variável, bem como os parâmetros para a submissão ao programa das regras variáveis.

4º) *IVARB / TVARB/ MVARB/ VARB2000* – programas de regras variáveis que fornecem as probabilidades (pesos relativos). O *IVARB* é utilizado para variáveis dependentes binárias, as demais versões são usadas para variáveis enérgicas e não apresentam processos de seleção de variáveis independentes (PINTZUK, 1988).

É importante destacar mais uma vez que a ferramenta estatística constitui apenas um suporte para o trabalho analítico da linguagem. Os dados numéricos obtidos pelo VARBRUL não constituem verdades ou juízos categóricos, servem apenas como base para o trabalho de inferência do pesquisador. As apropriadas análise e manipulação do modelo dependem do

domínio do pesquisador sobre seu objeto de estudo e o universo de observação, seja na recolha adequada da amostra, na definição dos grupos de fatores ou no conhecimento necessário para a interpretação adequada dos resultados.

2 A INVERSÃO VERBO SUJEITO NO PORTUGUÊS BRASILEIRO

O fenômeno da ordem dos constituintes na sentença tem sido objeto de diversos estudos, em muitas línguas, sob diferentes enfoques teóricos. As análises para o português perpassam dados de língua falada e de língua escrita, desenvolvendo-se sob as perspectivas sincrônica e diacrônica. O presente capítulo traz uma revisão geral de algumas dessas análises que orientaram este trabalho e possibilitam uma compreensão mais global do fenômeno. O capítulo inicia-se com a revisão do tratamento dado pela tradição gramatical à questão da ordem sujeito-predicado, seguindo-se de uma discussão acerca da relação entre os chamados *verbos inacusativos* e a inversão VS. Posteriormente, apresentamos uma breve contextualização histórica do fenômeno e um retrato geral da mudança pela qual teria passado o português brasileiro (PB), até chegarmos às análises variacionistas do português brasileiro contemporâneo, apresentando sua configuração e condicionamentos nas sentenças declarativas e interrogativas. Terminamos o capítulo com a apresentação dos possíveis condicionamentos sociais do fenômeno.

2.1 A INVERSÃO VS NO PORTUGUÊS BRASILEIRO: CARACTERIZAÇÃO GERAL

Conforme Neves (1990), os gramáticos tradicionais concordam em apontar a disposição S(sujeito) V(verbo) como a ordem normal/típica/usual/fundamental para os constituintes do sintagma no português, paradigma comum às línguas românicas. Tal ordem baseia-se no padrão mais frequente na língua portuguesa, o qual se caracteriza pela: i) colocação do sujeito e seus “atributos” na primeira posição; ii) o verbo em segunda; iii) os complementos em terceira. Essa ordem, embora básica, não é exclusiva, havendo até situações em que a inversão desse padrão é a ordem preferencial, como destacam alguns gramáticos. Neves (1990) destaca que alguns autores mais antigos chegam até mesmo a alegar uma suposta flexibilidade em relação à ordenação no nosso idioma (por exemplo, COELHO, 1891; BARBOSA, 1871; VASCONCELOS, 1898). Mesmo em autores mais modernos, ainda é possível encontrar alguma consideração sobre uma relativa liberdade na colocação dos constituintes no português. Segundo Cunha e Cintra (1985, p. 157), “as inversões não repugnam o nosso idioma, pelo contrário, com muito mais facilidade do que

outras línguas (...) ele nos permite alterar a ordem normal dos termos da oração”. Em geral, são aludidos como princípios detonadores da ordem inversa: clareza, ênfase, organização lógica do pensamento, eufonia e efeito estilístico.

Embora haja consenso entre os autores no que concerne à existência de um padrão ocasional ou divergente ao lado da disposição habitual dos constituintes, há diferenças quanto ao tratamento dessa suposta variação, conforme destaca Neves (1990). Para Bechara (2001), por exemplo, a ordem inversa constitui um “afastamento da norma”, com efeito estilístico. Cunha e Cintra (1985), por sua vez, destacam que além da motivação estilística, a inversão, em contextos “consagrados pelo uso”, constitui uma exigência gramatical, como, por exemplo, nas orações interrogativas (a) e imperativas (b) ou com verbos intransitivos (c):

- a. *Onde estão as garotas?/ Quanto custa esse carro?*
- b. *Faça você o trabalho.*
- c. *Chegaram as crianças.*

Sobre o efeito estilístico, destaca-se o uso da inversão como recurso literário, conforme aponta Neves (1990, p. 188): “É freqüente a ligação da inversão dos padrões usuais da ordem com a obtenção de efeitos literários”. Esse recurso, corrente na prosa e, sobretudo, na poesia, é justificado, segundo os gramáticos analisados pela autora, em face de imposições como a métrica e a rima, bem como a conferência de ênfase a determinado tema e costuma ser reconhecido pela designação de *hipérbato*¹³.

Para Berlinck *et al.* (2001), as descrições e análises propostas pelas Gramáticas Tradicionais (GT) são imprecisas e não levam a um fator ou princípio geral que explique satisfatoriamente todos os casos. A ênfase, por exemplo, é criticada como explicação vaga, pois é um recurso restrito a determinados contextos e pouco frequente. Observa-se, assim, uma série de desacordos em relação ao tema da inversão da ordem sujeito-verbo entre a abordagem tradicional e análise linguística moderna: “Se há um tema que as obras tradicionais tenham tratado de modo diferente em relação aos ‘modernos’ estudos baseados nos avanços da linguística, esse tema é a ordem das palavras.” (NEVES, 1990, p. 187).

¹³ Embora o termo hipérbato, em sentido estrito, refira-se a separação de termos de um mesmo sintagma, pela interposição de um outro termo, destaca-se uma aplicação mais corrente, designativo de “toda inversão da ordem normal das palavras na oração, ou da ordem das orações no período, com finalidade expressiva.” (CUNHA; CINTRA, 1985, p. 610)

Porém, apesar das divergências entre as duas abordagens, podem-se identificar algumas convergências na identificação de alguns dos fatores que condicionam a ordem sujeito-verbo na sentença. Uma dessas convergências seria a correlação da inversão verbo-sujeito (VS) e uma classe específica de verbos.

2.1.1 Os verbos inacusativos e a ordem VS

Já na tradição gramatical, é reconhecido um comportamento particular entre alguns verbos ditos intransitivos, no que diz respeito à ordem sujeito-verbo. Cunha e Cintra (1985) inserem no grupo das inversões verbo-sujeito de natureza gramatical, as construções com verbos unipessoais¹⁴. Rocha Lima (1972, p. 237), por sua vez, salienta, sob o título *Uma inversão que requer cuidado*, a dificuldade na identificação de alguns sujeitos pospostos de verbos intransitivos, os quais, por vezes, são interpretados pelos leitores como objetos, como no caso da seguinte frase:

(2) *Apareceu, enfim, o cortejo real* (ROCHA LIMA, 1972, p. 238)

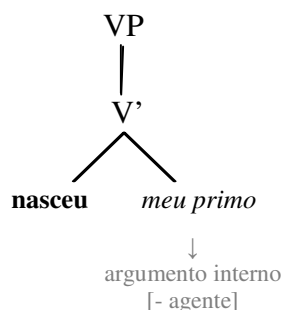
Como se pode observar, embora não sejam claramente diferenciados pelas gramáticas tradicionais subgrupos para os chamados intransitivos, vislumbra-se um comportamento diferenciado entre alguns desses verbos. Essas diferenças foram escrutinadas e sistematizadas a partir da década de 70 e encontram hoje uma considerável concordância no campo da linguística moderna.

No âmbito dos estudos linguísticos, as considerações sobre a grade argumental dos verbos revelaram importantes contribuições para questão da ordem. A análise dos verbos monoargumentais demonstrou características idiossincráticas em um subgrupo desses verbos. Designados *inacusativos*, esse subgrupo, além de propriedades estruturais específicas, agrega em seu único argumento selecionado características que o identificam com a função de objeto. Dessa forma, nos verbos inacusativos (a), o sujeito sintático é gerado na posição de argumento interno, apresentando características semânticas de objeto, tais como um papel

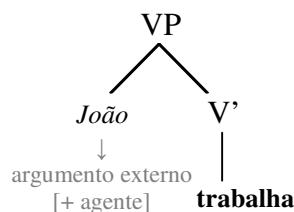
¹⁴ São unipessoais “os verbos que, pelo sentido, só admitem um sujeito na 3ª pessoa do singular ou do plural”, tais como *acontecer, constar, parecer*, e verbos que exprimem estado ou ação animal, como *ladrar, rosnar, trotar*. (CUNHA e CINTRA, 1985, p. 433).

temático [-agentivo]. Os demais verbos intransitivos (b), por sua vez, selecionam o argumento externo, em geral com traço mais animado e papel temático [+agentivo].

(a) Inacusativo



(b) Intransitivos



Esta diferenciação é proposta inicialmente por Perlmutter (1978) sob a designação *hipótese inacusativa*. Posteriormente, Burzio (1986) introduz o que ficou conhecido no âmbito da teoria como *generalização de Burzio*, a qual postula que o verbo inacusativo não é capaz de atribuir caso acusativo e não seleciona argumento externo.¹⁵

Assim, as propriedades básicas do fenômeno da inacusatividade podem ser sintetizadas da seguinte forma (CARVALHO, 2006, p. 15):

- a) O verbo não é capaz de atribuir caso acusativo ao seu argumento interno;
- b) O verbo não é capaz de atribuir papel temático externo à posição de sujeito.

Em outras palavras, embora verbos como *morrer*, *aparecer* e *nascer* apresentem como complemento, do ponto de vista estrutural, apenas o objeto, não são capazes de atribuir a este o caso acusativo, o que o leva a ser movido para uma posição anterior ao verbo, na qual recebe caso nominativo, identificando-se, do ponto de vista sintático, como sujeito da sentença. Sobre a manutenção do sujeito dos inacusativos *in situ*, ou seja, posposto ao verbo, são propostas algumas soluções teóricas para explicar a atribuição de caso, a exemplo da atribuição de *caso partitivo* (BELLETI, 1988) e *atribuição de caso por cadeia* (BURZIO, 1986). Essas questões, no entanto, não serão aqui exploradas, haja vista não constituírem diretamente matéria desta dissertação.

¹⁵ Para o conhecimento de outras propostas, bem como uma abordagem mais aprofundada do assunto, consulte Carvalho (2006).

A respeito da relação entre esses verbos e a ordem dos constituintes, os verbos inacusativos, por apresentarem sujeitos selecionados originalmente na posição de objeto, licenciariam a ordem VS, pois, na verdade, apresentariam a manutenção do argumento interno na posição em que foi gerado. Isto explica a ampla frequência de sujeitos pospostos com este tipo de verbo.

Sobre a designação desse grupo verbal, Carvalho (2006, p. 19), salienta que

[...] para designar os diferentes tipos de verbos que manifestam o fenômeno da inacusatividade, alguns autores preferem a designação de “inacusativos”, como Perlmutter (1978) e Duarte (2003), por exemplo; outros, a de ergativos, como Burzio (1981/1986) e Eliseu (1984); outros, ainda, distinguem inacusativos de ergativos, conforme Haegeman (1996).

Cumprе ressaltar que, numa abordagem mais ampla, o rótulo construções inacusativas engloba não somente os verbos monoargumentais, delimitados pelas propriedades acima descritas, mas também certas construções que partilham dessas propriedades, consagradas como *fenômeno da inacusatividade*, a saber, as construções passivas (3), copulativas (4), ergativas (5) e médias (6), apresentadas a seguir:

(3) *Os presentes foram dados a José*

(4) *João é inteligente*

(5) *O pneu do carro furou*

(6) *Estas flores encontram-se facilmente*

Nesse sentido, Carvalho (2006) distingue a inacusatividade lexical da inacusatividade estrutural. O primeiro caso corresponde à classe dos verbos inacusativos, os quais apresentam intrinsecamente as características acima apontadas, constituindo uma classe delimitada na tipologia verbal. Ainda segundo a autora, podem ser subdivididos em: de existência (existir, viver, constar), de movimento (ir, vir cair), de aparição (aparecer, desaparecer, sumir) e de causa interna (nascer, morrer, crescer). O segundo caso diz respeito às construções que participam do fenômeno da inacusatividade por processos sintáticos ou morfológicos. As construções ergativas caracterizam-se por alternar uma variante transitiva/causativa e outra ergativa/inacusativa, como exemplo (13), abaixo:

- (7) a. *O prego furou o pneu do carro.* (transitiva/causativa)
 b. *O pneu do carro furou.* (ergativa/inacusativa) olhar o espaçamento entre a e b

Tomando como objeto os casos da inacusatividade lexical, buscamos identificar critérios para a distinção entre verbos inacusativos e os intransitivos. A fim de distinguir estes subgrupos dos verbos monoargumentais, Mateus *et al.* (2003) apresentam alguns testes, os quais nortearam a análise empreendida neste trabalho. Assim, classificaram-se como intransitivos verbos como *correr* e como inacusativos verbos como *morrer*, com base nas seguintes propriedades atribuídas aos inacusativos:

i) admitem construções com particípio absoluto:

- (8) *O paciente morreu* → *morto o paciente*
 (9) *O atleta correu* → **corrido o atleta.*

ii) podem ocorrer em posição predicativa ou atributiva:

- (10) *O paciente [está morto]* PRED / *[O paciente morto]* ATRIB. *é meu amigo*
 (11) **o atleta [está corrido]* PRED./ **[O atleta corrido]* ATRIB. *é meu amigo.*

iii) não permitem nominalização em –or:

- (12) *morrer* → **morredor*
 (13) *correr* → *corredor*

iv) selecionam verbo *ser* como auxiliar nos tempos compostos, propriedade observável no português arcaico e em algumas línguas românicas como o francês e o italiano (MIOTO *et al.* 2004, p. 162):

- (14) *O atleta correu* → *L'athète a couru* (*passé composé* com o verbo *avoir*)
 (15) *O paciente morreu* → *Le patient est mort* (*passé composé* com o verbo *est*)

O exame dessas propriedades revela que, em alguns aspectos, os inacusativos assemelham-se aos verbos transitivos, mais especificamente, os “sujeitos” dos inacusativos

partilham de características comuns aos objetos dos verbos transitivos, tais como i) *aditem construções com participio absoluto* e ii) *ocorrem em construções predicativas ou atributivas*. Vejamos os exemplos com verbo transitivo, extraídos de Mateus *et al.*:

(16) *O professor reviu os exercícios* → *revistos os exercícios,...*

(17) *A porta [está aberta]_{PRED}* → *[a porta aberta]_{ATRIB} está empenada*

Em nosso trabalho, consideramos apenas os verbos inacusativos, sendo as passivas e copulativas codificadas separadamente. Isso porque acreditamos que estas, embora apresentem propriedades inacusativas, constituem construções específicas e diferenciadas, não se enquadrando no rótulo *tipo de verbo*. Quanto aos verbos de movimento, também foram considerados separadamente, tendo em vista a necessidade, nesse tipo de verbo, da presença do complemento locativo, o que não ocorre com os inacusativos, e, sobretudo, ao fato de seu sujeito possuir um papel temático agentivo, como se pode ver no seguinte exemplo:

(18) *o cara nunca ficô em casa (POS-06)*

O ato de ir a algum lugar implica movimento cinético e intencionalidade, o que o distingue dos processos referidos pelos verbos estritamente inacusativos, tais como: *nascer*, *adoecer*, *envelhecer*, etc. Obviamente, não desconsideramos os empregos do tipo (19), nos quais o argumento externo apresenta um papel temático eminentemente de tema e não de agente, conforme assinalado em (18). Além disso, nesses casos, o locativo é dispensável.

(19) *Aí vêm os problemas. (SP-05)*

Portanto, reservamos o rótulo de *inacusativos* para o subgrupo dos verbos monoargumentais que, conforme vimos, distingue-se dos tradicionais intransitivos por apresentar uma configuração estrutural diferenciada; configuração esta que justifica a presença de sujeito sintático sensivelmente diferente dos sujeitos dos demais verbos intransitivos, já que aqueles apresentam características, sobretudo temáticas, assemelhadas aos objetos e propriedades estruturais específicas, conforme evidenciado nos testes anteriormente apresentados. Além disso, mantivemos sob essa designação verbos de movimento com leitura eminentemente inacusativa (19), cujo único argumento selecionado

apresenta papel de *tema*, e não apresentam complemento locativo. Tendo em vista essas considerações, definimos a tipologia verbal adotada nesta análise da ordem sujeito-verbo, a qual terá sua pertinência rediscutida no desenvolvimento do trabalho.

Se o predicador verbal revelou ser um importante condicionamento gramatical da ordem sujeito-verbo, uma compreensão mais global do fenômeno passa necessariamente pelo processo histórico no qual a ordem dos constituintes na sentença assumiu um valor funcional dentro do sistema gramatical da língua. Esse processo pode ser situado na passagem do latim para as línguas românicas modernas, dentre as quais figura a língua portuguesa.

2.2 A ORDEM DOS CONSTITUINTES NO PORTUGUÊS: BREVE CONTEXTUALIZAÇÃO HISTÓRICA

É sabido que o latim apresentava um caráter bastante flexível quanto à ordenação dos constituintes sentenciais, isto porque as funções sintáticas eram marcadas pelas flexões casuais. Em outras palavras, as relações funcionais estabelecidas entre os constituintes eram expressas morfologicamente. Dessa forma, independentemente de sua colocação na sentença, o sujeito poderia ser reconhecido pela identificação da flexão de nominativo, conforme (20) e (21), a seguir:

(20)

- a. *Puella puerum amat*
- b. *Puerum puella amat*
- c. *Puerum amat puella*

(21)

- a. *Puer puellam amat*
- b. *Puellam puer amat*
- c. *puellam amat puer*

Em (20), o morfema *-a* confere à *puella* o caso nominativo que o identifica como sujeito da sentença; em (21), ao receber o morfema *-am*, recebe caso acusativo, sendo identificado com a função sintática de objeto da sentença.

Conforme Lleal (1990), as flexões latinas de caso sempre caminharam para um contínuo processo de redução. O próprio latim clássico conservava apenas seis, dos oito¹⁶ casos do indoeuropeu. Essa tendência à “desflexionalização” é, de um modo geral, apontada pelos estudos filológicos como uma tendência interna e geral nas línguas românicas e, supostamente, estaria presente desde o latim.

O processo lento que conduziu à perda da flexão casual nas línguas românicas tem causas diversas e, em algum ponto, insondáveis. Entretanto, podemos identificar que o sistema de casos latinos teve, na base de sua mudança, algumas fraquezas de natureza morfológica e fônica. O caso vocativo, por exemplo, em virtude da marcação redundante (flexão e entoação), logo passou a ser expresso pelo nominativo. Acrescenta-se a isso o fato de que existiam coincidências formais entre os casos dativo e ablativo e entre o acusativo e o nominativo (LLEAL, 1990, p. 76-77). A identidade formal de certos casos (o sincretismo dos casos), a expressão de algumas noções por mais de um caso e o uso de casos diferentes para expressar a mesma noção, bem como as mudanças fônicas de perda de quantidade como traço distintivo, apócope do *-m* final do acusativo e a identidade entre *-o* e *-u* finais teriam contribuído para o desaparecimento do sistema flexional. (MATTOS e SILVA, 1993; POGGIO, 2002).

Associado a um sistema morfológico e fonicamente instável, tem-se o forte contato linguístico estabelecido com os povos romanizados. Mesmo na gênese dessas instabilidades, a interferência dos povos aloglotas não pode ser desprezada. Isto porque, se houve redução morfológica e outras determinações internas, por outro lado o contato entre comunidades linguisticamente diferentes esteve sempre presente durante o longo período de expansão do Império Romano.

Independentemente das causas internas ou externas dessa mudança, o fato é que com a perda do caso morfológico, novos procedimentos tornaram-se necessários para a expressão das relações funcionais. Isso explica o desenvolvimento de um sistema de preposições, ao lado de uma tendência de fixação da ordem:

As partículas conectivas ditas “preposições”, e a ordem dos vocábulos na frase tornaram-se o meio de expressão das relações sintáticas, que no tipo linguístico flexional assenta essencialmente em desinências nominais específicas. Assim, as desinências de nominativo, para o nome sujeito, e de acusativo, para o nome objeto direto, que permitiam variações livres e estilísticas de uma frase como *puer vidit lupum* (*lupum puer vidit, lupum*

¹⁶ Acusativo, nominativo, vocativo, dativo, genitivo, locativo, instrumental e ablativo (LLEAL, 1990, p. 76.).

vidit puer, vidit lupum puer, etc), passou-se à ordem românica – *(o) menino viu (o) lobo [...]*.” (CÂMARA JÚNIOR, 1979, p. 23-24)

No português arcaico, essa “remodelação dos padrões sintáticos” já se faz claramente presente. Embora tenha convivido com outros tipos de ordenações, a ordem básica no período arcaico era: sujeito (S) + verbo (V) + complemento (C), (MATTOS e SILVA, 1993).

Com base em Huber (1986) e Pádua (1960), a autora destaca seis possíveis ordenações presentes no referido período e exemplificadas a seguir:

SVC: *o lobo abriu a boca*

SCV: *Quando Eufrosina esto ouvio, prougue-lhe muito*

VSC: *E enton chamou o abade hũũ monge*

VCS: *E cercou a cidade Nabucodonossor*

CVS: *Quando o viu o moço, rogou que veesse.*

CSV: *Todas estas cousas as gentes demandou.(...)*

(MATTOS E SILVA, 1993, p. 124)

Sobre a ocorrência de sujeitos pospostos no português clássico, Paixão (2007, p. 4) afirma que “nos textos portugueses escritos até o século 17 é bastante freqüente a ordem superficial {XVS} (onde X é um complemento do verbo)”. Para fundamentar esta observação, a autora apresenta os seguintes exemplos, extraídos de textos dos séculos XVI e XVII:

(22)

a. *Uma carta de Vossa Reverencia me deu o padre Frei Ene perto de Cascais*

b. *A este livro chamao elles: “Quid pro quo”: quer dizer, huma couza por outra*

c. *Na Corte andou este Rei dous anos*

As seis ordenações apresentadas para o Português arcaico se mantêm no PE moderno, sendo vinculadas, no entanto, a requisitos específicos¹⁷. Salvo a ordem SVC (ou SVO), a qual é considerada neutra em PE, todas as demais encontram-se licenciadas mediante restrições de natureza prosódica, pragmática, sintática ou mesmo semântica (ÂMBAR, 1992, apud FERNANDES, 2007).

¹⁷ Segundo Costa (2000), a ordem SOV seria agramatical no PE.

* *O Paulo a sopa comeu* (COSTA, 2000, p. 94)

Costa (2000) demonstra que as sentenças com sujeito posposto em geral estão vinculadas a uma restrição de natureza pragmática. Para as ordens VOS e VSO, o autor postula as seguintes restrições:

- i) Para a ordem VSO, o sujeito deve receber o “acento pesado¹⁸”, ou seja, recebe o acento prosódico, numa leitura contrastiva.

(23) *Comeu O PAULO maçãs.* (COSTA, 2000, p. 107)

- ii) Na ordem VOS, há duas possibilidades: sobre o sujeito pode recair o “acento setencial *default*(30)¹⁹” ou o “acento pesado (31)”

(24) *Comeu a sopa o Paulo*

(25) *Comeu a sopa O PAULO.* (COSTA, 2000, p. 108)

Assim, é a proeminência prosódica que marca a relação focalização/inversão o contexto licenciador da ordem VS no PE contemporâneo. Essa possibilidade de se lançar mão do mecanismo gramatical da inversão VS não se manteve da mesma forma na outra variedade da língua portuguesa transplantada para o Brasil.

2.2.1 A evolução da ordem VS no Português Brasileiro

A frequência da inversão VS é um dos muitos aspectos gramaticais que revelam o desenvolvimento diferenciado da língua portuguesa no Brasil em relação à variedade europeia. Sobre o contraste entre o português europeu (PE) e o português brasileiro (PB) no tocante à ordem VS, não raro destaca-se um comportamento menos flexível desta variedade (MENUZZI, 2003). Porém, segundo estudo contrastivo realizado por SPANO (2002 *apud* SPANO, 2008) em dados da fala culta do PE e do PB, as duas variedades não se mostraram tão distantes quanto à ocorrência da ordem VS:

¹⁸ O acento forte (*heavy stress*) marca uma leitura contrastiva. A sentença *Comeu a sopa O PAULO* seria uma leitura contrastiva, em resposta a uma pergunta do tipo: *Quem comeu a sopa, o João ou a Joana?*

¹⁹ A opção *default* marca como foco o elemento mais à direita da sentença, posição universalmente proeminente de foco. Assim, em (30), pode ser lido como uma resposta ao contexto *Quem comeu a sopa?*, donde o sujeito *o Paulo* corresponde à informação nova e foco da sentença.

A análise dos dados revela que há mais semelhanças do que diferenças entre o PB e o PE quanto ao comportamento da ordem VS quando se consideram os verbos monoargumentais: a significativa predominância da ordem SV e a restrição de uso da ordem VS, em aproximadamente 30% dos dados, a certos tipos de verbos monoargumentais. (SPANNO, 2008, p. 79)

Vale destacar que o estudo da autora limita-se às construções monoargumentais, as quais parecem, de um modo geral, apresentar menor restrição a VS. Além disso, a análise da autora focaliza a variedade culta do PB, a qual, pelo próprio processo de constituição histórica, mostra-se mais próxima da variedade européia (cf. seção 3.2). Em referência ao estudo diacrônico de Berlinck (1995), Spanno (2008) indica que a diferença entre as duas variedades estaria “na possibilidade de apenas o PE ainda pospor o SN em contextos pluriargumentais na configuração VSX ao contrário do PB”, para o qual esta possibilidade teria se perdido (SPANNO, 2008, p. 87).

Uma investigação na diacronia do PB indica-nos que até meados do século XIX, ainda era possível encontrar no PB uma flexibilidade que o aproximava do PE. Isto passa por alguns processos que desembocam numa ordenação extremamente rígida que caracteriza o PB do século XX.

Os resultados cotejados por Berlinck (1988; 1989) apontam um decréscimo gradual de ordenação VS e crescente enrijecimento de ordem SV no transcurso dos séculos XVIII ao XX (cf. gráfico 1). Vale destacar que os dados correspondentes aos séculos XVIII e XIX são coletados a partir da modalidade escrita (cartas de cunho pessoal), condição já prevista para o estudo em tempo real de longa duração (TARALLO, 1985), mas que pode ter exercido alguma interferência na alta frequência de VS.

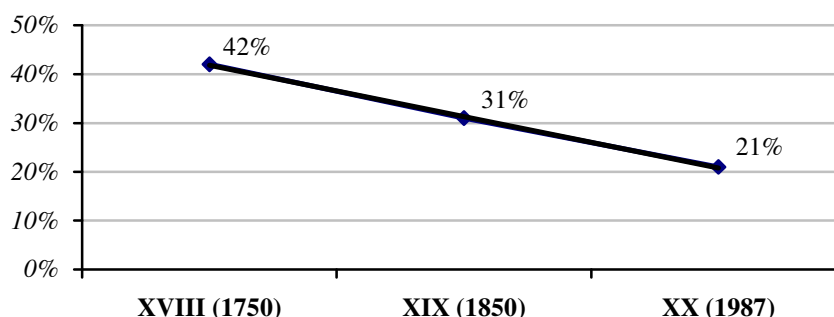


GRÁFICO 1: Ordem VS na diacronia do PB, adaptado de Berlinck (1989, p. 97).

Ao analisar os dados encontrados, a autora elenca os principais condicionamentos para cada corte sincrônico, os quais se mostram levemente diferenciados para cada período, de modo que os fatores condicionantes de VS “ou não são selecionados simultaneamente para os três *corpora*, ou têm seu peso relativo alterado (tornando-se menos ou mais relevantes)” (BERLINCK, 1988, p. 222). Os fatores indicados como relevantes no condicionamento sofrem uma “re-hierarquização”, ou, nas palavras da autora, “o *status* informacional, que atuava num primeiro nível, passa a influenciar a ordem num segundo momento, posterior à ação do verbo predicador” (BERLINCK, 1988, p. 223). Estes resultados levam à conclusão de que o processo de mudança que teria se processado no PB estaria centrado na passagem de uma orientação mais funcional (*status* informacional do SN sujeito) para uma orientação mais formal (a natureza do verbo predicador).

A configuração apresentada no gráfico 1 estaria, segundo Tarallo (1993), ligada ao fato de que no PB houve uma grande mudança no sistema de pronominalização, o que levou a emergência de sujeitos pronominais e, sobretudo, de objetos nulos. No tocante aos sujeitos pronominais, destacam-se as palavras de Kato *et al.* (2006):

Durante todo o século XIX, os sujeitos referenciais de primeira, segunda e terceira pessoas são preferencialmente nulos [...]

[...]

Tal situação, entretanto, muda radicalmente na segunda metade do século XX, quando passam a predominar os sujeitos referenciais (definidos e indeterminados) foneticamente representados” (KATO *et al.*, 2006, p. 416-417).

O incremento do objeto nulo, por sua vez, estaria associado à mudança que desencadeou a perda do clítico de terceira pessoa no PB (KATO *et al.*, 2006; CYRINO, 1993; TARALLO, 1993).

Estes dois fenômenos associados, em última instância, intervieram na restrição de VS, conforme demonstram os resultados do gráfico 2.

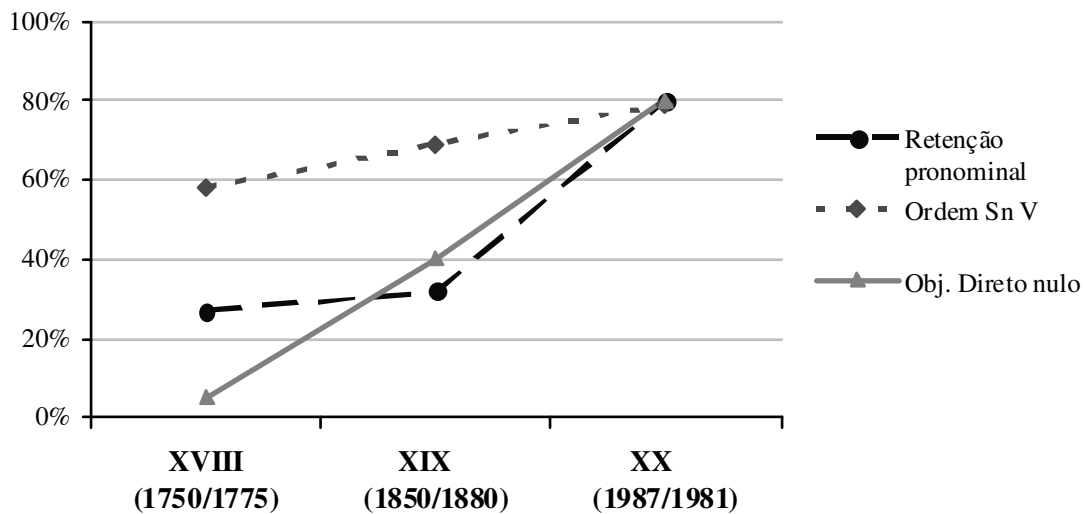


GRÁFICO 2: correlação de frequência de SV, retenção pronominal na posição de sujeito e objeto direto nulo, segundo Tarallo (1993, p. 84) e Berlinck (1989, p. 97). FONTE: Coelho (2006, p. 97)

Os resultados apresentados no gráfico 2 demonstram convergência entre os três fenômenos observados: a ordem rígida SV se instaura ao passo que cresce a tendência à realização do sujeito pronominal e a erosão do objeto.

A correlação dos fenômenos de retenção pronominal e de V SN com verbos transitivos indica que quando o clítico acusativo ainda constituía uma estratégia bastante produtiva na língua (século XVIII), o SN posposto corria um risco menor de ser interpretado como objeto de uma construção transitiva, justamente porque este último era atualizado na forma de clítico

[...]

Também o paralelo aumento da frequência de objeto vazio na passagem de um momento a outro e sua atual predominância sobre as demais estratégias de pronominalização contribuiu para o ‘enrijecimento’ da ordem SN V. Nas estruturas em que a função-objeto está ‘ocupada’ por uma categoria vazia, relevadas as restrições particulares do verbo, é possível que o SN posposto gere uma construção ambígua, sendo percebido como objeto de um verbo com sujeito nulo. (BERLINCK, 1988, p. 242-243).

Sumarizando, com a realização nula do objeto, precedida da perda do clítico acusativo, e a retenção do sujeito pronominal, motivada, sobretudo, pela perda da morfologia flexional, a posição do sujeito anterior ao verbo passou a ser fundamental para sua identificação nesta variedade do idioma. Lembrando que nenhuma mudança acontece no vácuo, mas antes afeta e é afetada pelo sistema linguístico que a circunda (LABOV, 2008 [1972]), podemos concluir que o que se desenha pelos dados e explicações expostas é o encaixamento linguístico do

fenômeno, o qual decorre de um lento processo histórico e encontra-se ensejado num conjunto maior de mudanças que afetaram o português brasileiro de um modo mais geral.

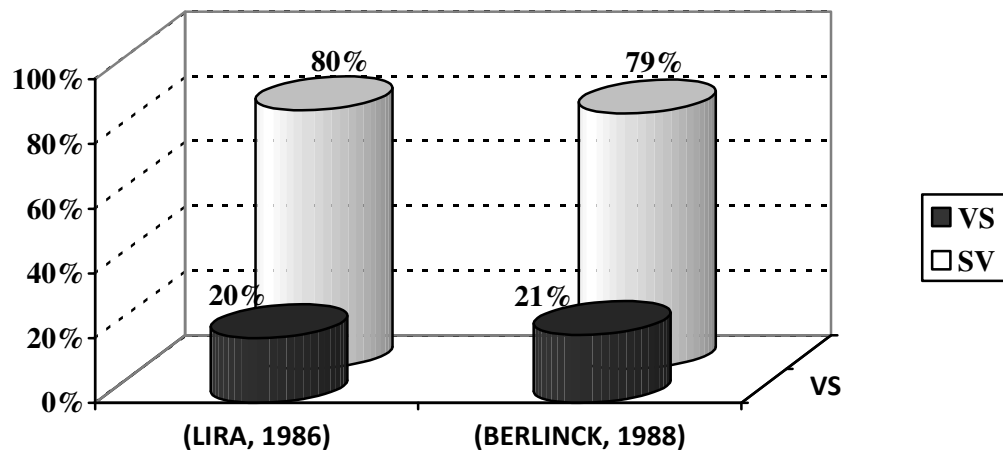
De qualquer forma, todo esse processo caracteriza-se por um crescente enrijecimento da ordem SV no PB tanto nas frases interrogativas quanto nas declarativas.

2.3 A INVERSÃO VS NO PORTUGUÊS BRASILEIRO CONTEMPORÂNEO: CONDICIONAMENTOS ESTRUTURAIIS

Considerando a crescente restrição à inversão VS no PB, torna-se relevante identificar os contextos em que essa possibilidade estrutural ainda se conserva nos padrões de uso linguístico brasileiros. Uma revisão das análises variacionistas que já foram feitas sobre o assunto demonstra que, consoante o tipo de frase que se observa, declarativa ou interrogativa, o fenômeno apresenta uma configuração ligeiramente diferenciada. Desse modo, esta exposição sobre os condicionamentos estruturais da inversão VS no PB está dividida em duas seções, cada uma dedicada a um desses dois tipos de frase.

2.3.1 A inversão VS nas frases declarativas

Os estudos sociolinguísticos sobre o português brasileiro, de modo geral, destacam a baixa ocorrência de sujeitos pospostos nas sentenças declarativas do PB. Vejamos alguns resultados:



GR

ÁFICO 3: A ordem VS segundo alguns estudos sociolinguísticos.

Os dados levantados pelas autoras indicam uma tendência geral do PB à fixação da ordem SV. Os estudos destacados recobrem variedades cultas do PB e observam todos os tipos de verbos em sentenças declarativas. Assim, podemos depreender que a ordem VS nas sentenças declarativas constitui uma variante residual, ou em desuso, no PB contemporâneo, com uma frequência de aplicação relativamente baixa.

Sobre os fatores que ainda estariam condicionando a ordenação VS, os resultados apontados pelos autores são, em certa medida, convergentes e flutuam em torno de aspectos de natureza discursiva e estrutural. Por esse motivo optamos por apresentar as considerações acerca do tema, não por autores, mas sim a partir dos juízos e resultados recorrentemente apontados.

Os estudos sobre o fenômeno nas declarativas apontam a natureza do verbo predicador como principal restrição à VS. Partindo da importância deste fator, Pilati (2006, p. 21) aponta a existência de dois padrões de comportamento para as orações declarativas com ordem VS no PB: (i) inversão menos restrita, para orações com verbos inacusativos (*Morreram as crianças*); (ii) inversão mais restrita se o verbo for inergativo ou transitivo (*Ligou a Maria/ Tomou posse o novo ministro*). Em suma, podemos extrair dois pontos centrais dessa relação: i) a restrição dos verbos transitivos para a ordem VS; ii) o licenciamento desta ordem pelos verbos ditos inacusativos.

O primeiro ponto seria justificado pela “restrição de monoargumentalidade”, proposta por Kato e Tarallo (1988) e retomada por Kato (2000). Com base em dados de estudos variacionistas e gerativistas, Kato (2000, p. 97) chega à seguinte constatação para o PB:

“quanto mais argumentos à direita, menos aceitável a sentença se torna”. Consequentemente, a inversão VS fica quase que restrita aos verbos que não subcategorizam argumentos internos, donde advém o rótulo *restrição de monoargumentalidade*. A autora ainda acrescenta: “A restrição quanto ao número de argumentos não é verdadeira para o Português Europeu (PE)”. Esse comportamento, sintetizado na generalização de que a ordem VS não é possível com mais de um argumento expresso, seria resultado da “perda de duas características do PB: a) o sujeito nulo referencial e b) o clítico acusativo de terceira pessoa” (KATO, 2000, p. 98). Isso explica a baixa ocorrência de sujeitos pospostos com verbos transitivos em estudos como Berlinck (1988), Alberton (2001), Lira (1986), Chaves (1989).

A recorrência de inversões com verbos inacusativos também é atestada em diversos estudos. Mesmo em trabalhos que tratam os verbos monoargumentais indistintamente como intransitivos, são destacáveis as diferenças substanciais entre alguns desses verbos. Berlinck (1988), por exemplo, procede a uma diferenciação entre intransitivos não existenciais e intransitivos existenciais e destaca a proeminência destes como contexto favorecedor de VS. Da mesma forma procede Alberton (2001). A inegável proeminência de VS com inacusativos já foi destacada e oportunamente discutida na seção 2.1.1 deste capítulo.

Além desse condicionamento de natureza estrutural, a presença de um constituinte na posição mais à esquerda da sentença também parece incidir sobre a realização de VS. Esse condicionamento estaria associado ao fato de o português brasileiro apresentar uma tendência ao preenchimento da primeira posição na sentença por um elemento que não seja o verbo, ou seja, por sua tendência a colocar o verbo na segunda posição (SPANNO, 2008). Assim, os resultados de pesquisas empíricas apontam uma tendência à posposição do sujeito quando na presença de um constituinte, sobretudo adverbial, na periferia esquerda de sentença.

São indicados nas análises quantitativas alguns condicionamentos para a ordem VS afetos à natureza do SN, dentre os quais o traço [-animado] do SN sujeito, o *status* informacional *novo* e o traço [-definido] (BERLINCK, 1988; ALBERTON, 2001; SPANO, 2008).

Alguns trabalhos, pautados na relevância dos verbos monoargumentais, têm centrado a análise do fenômeno nesse tipo de verbo, tais como Santos (2008), Coelho (2000) e Carvalho (2006). Estas abordagens, pautadas, em geral, numa perspectiva gerativa, também apresentam resultados convergentes a respeito dos condicionamentos vinculados ao SN.

No tocante aos fatores condicionadores de uma e outra ordem, Berlinck (1989) chega à conclusão de que o rótulo de construção “marcada” imbuído à VS deve ser abandonado, bem como se posiciona contrária à concepção deste fenômeno como um caso de variação

linguística. A autora pauta sua observação no fato de que os contextos apontados para uma e outra ordem são em geral opostos. Assim, enquanto a ordem VS acontece, por exemplo, preferencialmente com sujeitos informacionalmente novos, inanimados e com verbos intransitivos, a ordem SV se processa exatamente de forma contrária, com sujeitos de *status* informacional velho, com o traço semântico [+ animado] e com verbos transitivos. Estas considerações levam a autora a considerar o fenômeno em estudo como um caso de distribuição complementar e não como variação (no sentido laboviano do termo).

Este ponto de vista é reforçado por (SPANNO, 2008, p. 38):

As ordens SV/VS são, de fato, duas estruturas distintas, sem qualquer tipo de equivalência, decorrentes de um processo sintático – referente à estrutura típica de verbos inacusativos/inergativos (v. Perlmutter, 1976; Burzio, 1986) – e/ou discursivo (...)

Podemos concluir da leitura dos estudos sobre o tema, que a ordenação VS nas declarativas do PB atual é orientada, basicamente, porém não exclusivamente, por aspectos de dois níveis:

- a) **Discursivo:** a função pragmático discursiva exercida pelo SN sujeito. Os sujeitos informacionalmente novos, ou focalizados tendem mais à colocação pós-verbal. Associado a este fator, tem-se fatores semânticos como o traço de animacidade e o traço de definitude do SN sujeito.
- b) **Sintático:** A natureza do predicador. De um modo mais irrestrito, a ordem VS é mais natural e recorrente em verbos monoargumentais, especificamente num subgrupo desses verbos, os inacusativos.

Essas generalizações sobre a inversão VS no PB definem o balizamento da análise que se fará nesta pesquisa sobre a ordem sujeito verbo em uma variedade do PB bem determinada em termos sociolinguísticos.

2.3.2 A inversão VS nas frases interrogativas

Estudos linguísticos comparativos tem revelado que as sentenças QU- apresentam um contexto de contraste entre o PB e o PE (KATO; MIOTO, 2005). Dentre essas diferenças, destaca-se o emprego da ordem VS. Segundo Rossi (1993), o padrão de ordenação de sentenças interrogativas no PB difere do PE e de outras línguas românicas por não apresentar a regra obrigatória de inversão VS.

Esta regra, em PE, é obrigatória em interrogativas QU- diretas, introduzidas por um interrogativo simples/leve (sem núcleo nominal preenchido), independente da função sintática desempenhada pelo elemento QU-, como se vê no exemplo abaixo:

- (26) a. *Como reagiu o atleta?*
 b. **Como o atleta reagiu?*

Sentenças do tipo (26b), com o sujeito na posição pré verbal, seriam agramaticais para o PE. Essa obrigatoriedade se processaria nas interrogativas diretas e atingiria mesmo os contextos com verbos transitivos como em: *O que leu a Maria?*. Isso leva Brito *et al.* (2003) a afirmarem que a natureza do mecanismo responsável pela inversão nas sentenças declarativas deve ser diferente das interrogativas.

Segundo Âmbar (1987, *apud* DUARTE, 1992), esta generalização tem como exceções os casos com interrogativo *que* (27), pois este, mesmo nas interrogativas indiretas, ativa a inversão; e os casos em que há o emprego do expletivo *é que* (28), cuja presença anula a obrigatoriedade da inversão (ÂMBAR 1987; ROSSI, 1993; DUARTE, 1992 e BRITO *et al.* 2003).

- (27) *Não sei que comprou a Maria.*
 (28) *O que é que Maria comprou?*

Nas sentenças com QU- complexo/pesado (com núcleo nominal preenchido), a inversão é opcional (KATO; MIOTO, 2005; BRITO *et al.* 2003), como se pode ver nos exemplos abaixo, extraídos de Brito *et al.* (2003, p. 471):

- (29) a. *Que trabalhos a Maria fez?*

b. *Que trabalhos fez a Maria?*

A regra, aplicável, salvo algumas diferenças, também às outras línguas românicas, obedece ao que Rizzi (1996) designou *critério-WH*. Este critério, uma exigência de natureza estrutural, postula, que:

- um operador Q deve estar em uma configuração de especificador-núcleo com um núcleo Q;
- um núcleo Q deve estar em uma configuração de especificador-núcleo com um operador Wh. (KATO e MIOTO, 2005, p. 179)

Em referência a Âmbar (1992), Kato e Miotto explicam que o movimento de V nas interrogativas com Q leve do PE deriva do fato de “uma expressão Q leve apresentar uma categoria vazia que precisa ser regida, exigência que é satisfeita pela subida do verbo flexionado. Uma expressão Q pesada, por sua vez, não apresentando a categoria vazia postulada, dispensa a subida do verbo.” (KATO; MIOTO, 2005, p. 179).

No PB, onde a inversão, como vimos, tem aplicação bastante restrita, a obrigatoriedade da inversão parece não se aplicar. Sentenças como (30), que em PE seriam agramaticais, são plenamente aceitáveis no PB.

(30) *Como (o) Adriano Pinto reagiu?*²⁰

Duarte (1992), a partir de um estudo quantitativo, identifica na diacronia do PB um processo “crescente e constante” de mudança em direção à perda da ordem VS nas sentenças QU-. Os resultados do estudo realizado pela autora com base em textos teatrais dos séculos XVIII, XIX e XX e telenovelas brasileiras de 1989 demonstram que o PB mudou de um padrão com inversão, para um padrão sem inversão, como podemos observar no gráfico 4, extraído de Duarte (1992):

²⁰ Exemplo extraído de Kato e Miotto (2005, p. 173).

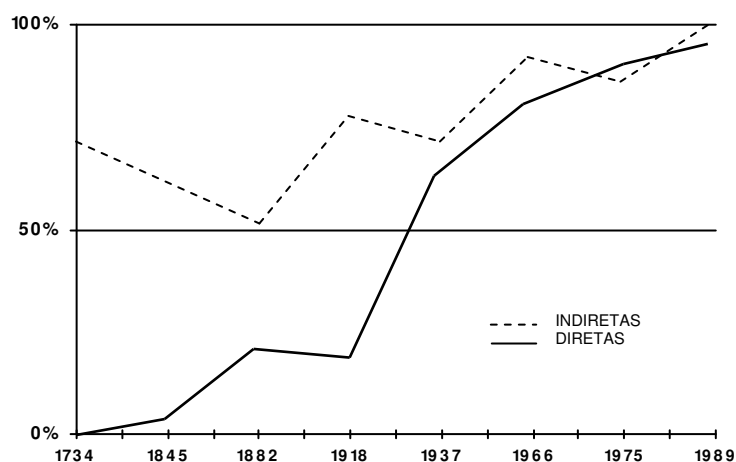


GRÁFICO 4: Ocorrência da ordem SV segundo o tipo de interrogativa *qu-* e o período de tempo. FONTE: Duarte (1992, p. 41).

Destaca-se que, para as interrogativas indiretas, a ordem SV, embora de uso crescente no decorrer do tempo, sempre foi preferida nos *corpora* analisados. Para as interrogativas diretas, entretanto, tem-se uma mudança drástica, passando-se de uma frequência de 100% de VS no século XVIII para um percentual da ordem de quase 100% de SV no final do século XX.

Para a autora, a mudança ocorrida nas interrogativas do PB teria começado paralelamente ao aparecimento do expletivo *é que* em meados do século XIX, passando a constituir o principal condicionamento da ordem SV. Aos poucos, inicialmente circunscrita aos contextos com o expletivo, a ordem SV generaliza-se para as sentenças sem o *é que*, perdendo esse contexto sua exclusividade na “detonação” desta ordem. Diante disso, identificam-se novos condicionamentos para as ordenações. Destacando o comportamento particular das sentenças iniciadas pelos interrogativos *que* e *o que* nos dados de 1995, Duarte identifica o estabelecimento do seguinte paradigma “o interrogativo *o que* detona sempre a ordem SV com o expletivo, enquanto o interrogativo *que* aciona a ordem SV com o expletivo e VS sem ele.” (DUARTE, 1992, p. 44). Nos períodos mais recentes do estudo, nos quais os percentuais de SV já superaram em muito a ordenação VS, destacam-se alguns contextos de resistência da inversão. De um modo geral, as últimas ocorrências de VS encontradas estão associadas a um interrogativo mais adverbial (*onde, como*) e apresentam um sujeito realizado foneticamente por meio de um SN. Além disso, são atestados alguns condicionamentos também atribuídos às ocorrências de VS nas declarativas, tais como verbos inacusativos, verbos copulativos *ser* e *estar*, sujeitos menos animados e sujeitos lexicalmente realizados.

Rossi (1993) demonstra posição contrária à pertinência do condicionamento exercido pelo *é que*, valendo-se do fato de que a interrogativa com o expletivo também aparece no PE, mas não elimina a inversão nos casos de interrogativas sem o expletivo.

Esse tipo de interrogativa aparece nos dados do teatro português europeu na mesma época em que aparece no teatro brasileiro (primeira metade do século XIX). Se o PBM²¹ tivesse passado apenas por essa mudança seria igual ao português europeu. (ROSSI, 1993, p. 330).

Isso leva a autora a propor que as diferenças do PB em relação ao PE seriam fruto de uma mudança no parâmetro do movimento de V processada naquela variedade do português. Segundo a autora, “o passo que concorreu decisivamente para essa mudança foi o enfraquecimento da concordância verbal em PBM”. Isto porque,

o critério Wh proposto por Rizzi (1991) atribui à flexão verbal a propriedade de iniciar a cadeia de licenciamento envolvida no movimento Wh, o que sustenta a hipótese de alguma mudança na flexão verbal do PBM ter contribuído para uma mudança na aplicação do critério Wh” (ROSSI, 1993, p. 331).

Em reexame aos dados apresentados por Duarte (1992), Kato e Duarte (1998) investigam uma possível relação entre a perda do sujeito nulo e a mudança na posição do sujeito nas interrogativas. As autoras, conforme Kato *et al.* (2006), chegam à conclusão de que o processo de mudança em direção à perda do sujeito pronominal teria de alguma forma influenciado a fixação da ordem SV. “À medida que cresce o preenchimento do sujeito pronominal nas declarativas, em consequência da redução de nossos paradigmas flexionais verbais (DUARTE, 1993), cresce também a preferência pela ordem SV nas interrogativas Q” (KATO *et al.* 2006, p. 426).

O sujeito nulo interpretado pelos adultos na posição pós-verbal (31), pode estar sendo interpretado pelas crianças em fase de aquisição como na posição pré-verbal (32).

(31) *Onde se esconderia (ela)?*

(32) *Onde (ela) se esconderia?*

²¹ Português Brasileiro Moderno.

Quanto às construções com SN lexicais, as poucas ocorrências de VS, restringem-se, no século XX, aos contextos com verbos inacusativos, conforme exemplo (33), extraído de Kato *et al.* (2006, p. 426).

(33) *Onde andar*á a *Neiva*?

As interrogativas, portanto, representam um contexto de contraste entre o PE e o PB. Neste, a inversão deixou de ser obrigatória, sobretudo pela inserção do expletivo *é que*. Este fator, no entanto, não é suficiente para explicar a questão nesta variedade do idioma. Outros fatores, como a perda da flexão verbal e a consequente retração do sujeito nulo, teriam favorecido a atual configuração do PB, bastante restritiva quanto à ordem VS.

2.4 A ORDEM VS/SV: CONDICIONAMENTOS SOCIAIS.

A convicção de que fenômenos de natureza sintática não seriam sensíveis aos fatores sociais conduziu alguns autores a descartar este tipo de condicionamento de suas análises (BERLINCK, 1988; SPANO 2008). Para esses autores, o fenômeno da ordem estaria eminentemente determinado por condicionamentos internos, daí a correlação com fatores externos ser, em certa medida, desnecessária:

Na medida em que a pesquisa se concentra no componente interno da língua, não foi necessário observar características sociais variadas na seleção dos informantes. (BERLINCK, 1988, p. 33).

Alguns autores, no entanto, têm mantido fatores de ordem social em suas análises. Na análise de Alberton (2001), a variável sexo é inicialmente indicada como relevante para a análise do fenômeno. Os resultados apresentaram, no entanto, uma diferença sutil e contraditória. Enquanto, nos resultados de frequência, os homens mostraram-se favorecedores de VS, com 47% das realizações (contra 46% das mulheres), os resultados dos pesos relativos apresentaram um resultado contrário: .56 para as mulheres e .45 para os homens. Interpretando o processamento quantitativo dos dados, a autora identifica o fator que teria causado o enviesamento: a animacidade. Numa segunda rodada, com a retirada desse fator, o grupo *sexo* não mais foi indicado como relevante. A autora salienta que o resultado pouco

conclusivo também pode ser atribuído ao baixo número de ocorrências disponíveis no estudo.

Coelho (2000), em análise da alternância SV/VS com verbos monoargumentais observa dois fatores sociais: *faixa etária* (25-39 anos; 40-49 anos; 50-59 anos; e mais de 60 anos) e *escolaridade* (primário; colegial). O controle dessas variáveis teve como objetivo, respectivamente, investigar um possível quadro de mudança em tempo aparente e a influência da escola. Os grupos de fatores não alcançaram significância estatística no processamento quantitativo dos dados pelo VARBRUL.

Carvalho (2006), em análise sobre a ordem em comunidades afro-brasileiras²², observa os fatores sociais *sexo*, *faixa etária*, *escolaridade* e *estada fora da comunidade* no condicionamento do fenômeno. O VARBRUL apontou como relevantes, no conjunto de dados analisados pela autora, os fatores *escolaridade* e *estada fora da comunidade*. A inversão VS mostrou-se favorecida por informantes que tiveram algum contato com o letramento, em oposição a falantes analfabetos, os quais desfavoreceram essa ordem. Segundo a autora, este resultado não era esperado, pois vai de encontro à tendência geral do PB, o emprego de SV. Para nós, que trabalhamos com um *corpus* comum à autora, este tipo de resultado parece contraditório, haja vista o fato de essas comunidades serem marcadas por uma ampla redução de morfologia flexional (SILVA, 2005) e da flexão casual dos pronomes (MENDES, 2009). Assim, acreditamos que esses fatores favoreceriam a baixa ocorrência de VS, já que esse padrão não seria próprio dessas comunidades. A escolarização para esses falantes apareceria, portanto, como via de contato com essas estruturas, incomuns à realidade linguística dessas comunidades.

O fator *estada fora da comunidade*, que deveria reforçar o nosso raciocínio, apresentou um comportamento contraditório, com indivíduos que nunca saíram da comunidade favorecendo a inversão VS. Vale lembrar que a composição da amostra não apresenta uma distribuição homogênea, o que pode ter interferido nos resultados. Em todo caso, somos conduzidos pelo raciocínio da autora de que as ordens SV e VS não apresentam uma avaliação social explícita.

Não podemos negar que, de um modo geral, as análises revelam pouca relevância dos fatores sociais no condicionamento da ordem dos constituintes *verbo* e *sujeito*. No entanto, nos parece muito impositivo afirmar que o fenômeno da ordem encontra-se incólume à

²² A autora procede a uma análise dos verbos inacusativos à luz do modelo gerativista e inclui em sua análise um estudo quantitativo da ordem nesse tipo de contexto. O *corpus* utilizado pela autora inclui duas das comunidades afro-brasileiras aqui analisadas, a saber, Cinzento e Sapé, constituídos pelo Projeto Vertentes.

influência desse tipo de fator. Assim, em nosso estudo, foram analisados os fatores sexo, faixa etária, escolaridade e estada fora da comunidade.

2.5 SUMARIZANDO

Se atualmente a ordem dos constituintes no PB apresenta um rígido padrão SV, haveria mais flexibilidade no passado, com uma probabilidade um pouco maior para a ordem VS. De qualquer forma, o padrão SV atualmente predomina em todas as variedades do português brasileiro, não estando sensível ao filtro social, mas de alguma forma ao estilístico, sobretudo se opormos o registro falado ao registro escrito. No bojo do encaixamento linguístico, destacam-se como catalizadores da rigidez SV mudanças em outros aspectos da sintaxe do PB. Assim, a proeminência de sujeito lexicalmente preenchido, bem como a insurgência de objetos nulos teriam conferido à ordenação sintática do PB o caráter notadamente rígido. As investigações sobre tema são abundantes e ricas. Assim, a simples descrição do fenômeno é tarefa há muito desempenhada no cenário da linguística. Resta-nos observar se, de alguma forma, há diferenças específicas das variedades populares e afro-brasileiras do interior da Bahia e em que medida essas variedades se aproximam ou se diferenciam do PB como um todo, sobretudo como a ordem VS distribui-se pela amostra em tempo aparente, e como (ou se) isso pode de alguma forma estar relacionado ao processo socio-histórico em que se formaram essas comunidades. Empreitados nessa tarefa, acreditamos estar contribuindo, num âmbito mais geral, para a compreensão e caracterização do PB, bem como de sua formação histórica. Esse foi o objetivo perseguido na análise dos dados, a ser apresentada no quarto e último capítulo desta dissertação.

3 O UNIVERSO DE OBSERVAÇÃO

É matéria desta dissertação a descrição e análise linguística do português brasileiro (PB) e de sua formação histórica. Frente à imbricada realidade social que compôs o complexo cenário linguístico de formação do PB, nosso olhar fixa-se em sua variedade popular, especificamente no português popular do interior do estado da Bahia. O interesse particular nesse estado decorre de sua proeminência social, política e cultural na gênese do Brasil, bem como da influência da matriz africana exercida nesse território. Vejam-se, a exemplo disso, as palavras do viajante alemão Robert Ave-Lallemant sobre a Bahia do século XIX, (apud ALBUQUERQUE; FRAGA FILHO, 2006, p. 66):

Quando se desembarca na Bahia, o povo que se movimenta nas ruas corresponde perfeitamente à confusão das casas e vielas. De feito, poucas cidades pode haver tão originalmente povoadas como a Bahia. Se não se soubesse que ela fica no Brasil, poder-se-ia tomá-la sem muita imaginação, por uma capital africana, residência de poderoso príncipe negro, na qual passa inteiramente despercebida uma população de forasteiros brancos puros. Tudo parece negro: negros na praia, negros na cidade, negros na parte baixa, negros nos bairros altos. Tudo que corre, grita, trabalha, tudo que transporta e carrega é negro; até os cavalos dos carros na Bahia são negros.

Acreditamos que em função do papel histórico desempenhado, a Bahia reúne, na atual realidade linguística, vestígios de processos pretéritos cruciais na configuração da realidade linguística brasileira, notadamente plural e polarizada, sobretudo quando fixamo-nos numa análise contrastiva entre comunidades rurais e comunidades rurais etnicamente marcadas, onde se desenvolveram situações particulares de contato linguístico.

Neste capítulo, apresentamos o universo de observação sobre o qual se processa esta análise. Em um primeiro momento, refazemos o encaixamento do nosso recorte no cenário linguístico nacional, apresentando as determinações históricas que teriam imprimido ao português popular o caráter diferenciado das variedades cultas. Para tanto, lançamos mão de algumas formulações teóricas que, a nosso ver, explicam a realidade do PB. No segundo momento, estabelecemos as orientações metodológicas adotadas na composição e tratamento da amostra.

3.1 A FORMAÇÃO DO PORTUGUÊS BRASILEIRO: CONTEXTO SÓCIO-HISTÓRICO.

A realidade linguística do Brasil apresenta uma constituição originalmente difusa e complexa, caracterizada inicialmente pelo multilinguismo generalizado, no qual figuravam, em condições desiguais, as línguas indígenas autóctones, as línguas africanas – trazidas ao Brasil pelo tráfico de escravos – e o português europeu. O cenário inicial sofre várias alterações ao longo do desenvolvimento do território nacional, as quais interferem cabalmente na composição linguística local.

Um breve exame da periodização da história linguística do Brasil revela a interferência de diversos fatores sociais na realidade sociolinguística brasileira. Pautada em três aspectos (demografia, urbanização e escolarização), Lobo (2003) organiza essas influências na configuração linguística do PB em duas grandes fases: a primeira, anterior ao fim do tráfico negreiro (1850), definida como um período de multilinguismo generalizado, marcado pelo povoamento rural, não escolarização e não standardização do português; a segunda, como uma fase de multilinguismo localizado, donde se destacam a acentuação do processo de urbanização, o incremento da escolarização e da normatização da língua. No cenário inicial de multilinguismo generalizado, faz-se referência à coexistência dos idiomas indígenas e europeu, num primeiro momento, e africanos, num segundo momento.

Quando os primeiros colonizadores europeus desembarcaram no Brasil, encontrava-se aqui uma população autóctone heterogênea e linguisticamente diversa. As estimativas contabilizam, para esse período, a existência de aproximadamente 1.273 línguas indígenas (RODRIGUES, 1993). Em verdade, a distribuição geográfica dessas línguas limitou o contato dos colonizadores com a diversidade linguística, uma vez que, na costa oriental, domínio inicial de colonização, predominava o Tupinambá, com algumas variações dialetais (RODRIGUES, 1993). Os índios, falantes dessas variedades, comunicavam-se por meio de uma língua franca, conhecida como *língua geral*, a qual foi inicialmente adotada como veículo de comunicação pelos exploradores e mais tarde viria a servir como o principal meio de comunicação entre mestiços brasileiros, até mais ou menos meados do século XVIII, quando é imposto o uso da língua portuguesa, através da reforma pombalina (MATTOS E SILVA, 2006).

Em síntese, a convivência entre os colonos e a população nativa, graças a uma violenta política de conquista e dominação, encerra, nas palavras de Franchetto e Leite (2006, p. 17), “uma história de danos e perdas”:

O que aconteceu, no século XVII, com a ocupação da Amazônia, foi um processo vertiginoso de extermínio e depopulação (*sic*). As perdas foram, pois bem maiores do que as que figuram nas estatísticas mais conhecidas.

O avanço dos colonos pelo território brasileiro foi acompanhado de um processo de violento domínio e escravização, o que acabou por redesenhar a distribuição demográfica brasileira. Ao passo que a população indígena caía em declínio, introduzia-se no território nacional a presença africana, trazida pelo tráfico negreiro em índices crescentes a partir do século XVI.

Além de uma mão de obra de baixo custo, o tráfico de escravos africanos representava para a sociedade da época um empreendimento altamente lucrativo. O fato importante é que a transplantação de escravos africanos para o Brasil interferiu significativamente na configuração demográfica e linguística local.

É difícil precisar o número exato de escravos africanos trazidos para o Brasil, mas segundo Petter (2006), ao longo dos três séculos da escravidão, teriam sido importados mais de três milhões de africanos. Sobre as regiões de procedência desses escravos, a literatura costuma apontar a existência de diferentes origens, entre elas se destacam as regiões que atualmente correspondem ao Benin, Nigéria, Angola e Moçambique. A respeito da filiação linguística desses escravos, destacam-se dois grandes grupos linguísticos o *kwa*, formada por línguas diversas e diferenciadas, e o *banto*, predominantemente as línguas quimbundo, quicongo e umbundo (LUCCHESI, 2009).

Além de protagonizar a mão de obra da produção agrícola e da extração de minérios, os escravos eram empregados em outras esferas de trabalho da sociedade brasileira, como as atividades domésticas (os escravos domésticos) e alguns dos serviços urbanos (LUCCHESI, 2006). A atuação dos escravos nos diferentes cenários sociais do Brasil ao longo da história, de um modo geral, sempre esteve circunscrita ao domínio do trabalho e da subjugação:

A escravidão foi muito mais do que um sistema econômico. Ela moldou condutas, definiu desigualdades sociais e raciais, forjou sentimentos, valores e etiquetas de mando e obediência. A partir dela instituíram-se os lugares que os indivíduos deveriam ocupar na sociedade, quem mandava e quem devia obedecer. Os cativos representavam o grupo mais oprimido da sociedade, pois eram impossibilitados legalmente de firmar contratos, dispor de suas vidas e possuir bens, testemunhar em processos judiciais contra pessoas livres, escolher trabalho e empregador. (ALBQUERQUE; FRAGA FILHO, 2006, p. 68-69).

Essas determinações impostas pelo sistema escravista à disposição social dos indivíduos repercutem, obviamente, no plano linguístico. Os escravos africanos, desenraizados e dominados, foram forçados a apreender a língua do dominador. Esta, no entanto, era aprendida sob circunstâncias muito adversas (LUCCHESI, 2001; 2003)²³.

A partir da Tabela 2, apresentada por Mussa (1991) e recorrentemente citada pelos linguistas, podemos observar como a constituição demográfica do Brasil colônia alterou-se ao longo dos séculos.

TABELA 2: População do Brasil por etnia do século XVI ao XIX.

Etnias	1538-1600	1601-1700	1701-1800	1801-1850	1851-1890
Africanos	20%	30%	20%	12%	2%
Negros brasileiros	-	20%	21%	19%	13%
Mulatos	-	10%	19%	34%	42%
Branços brasileiros	-	5%	10%	17%	24%
Europeus	30%	25%	22%	14%	17%
Índios integrados	50%	10%	8%	4%	2%

Fonte: Mussa (1991, p. 163).

Uma das constatações imediatas depreendidas dos dados incide sobre o percentual de índios. As frequências apresentadas ilustram o vertiginoso declínio dessa população, a qual correspondia inicialmente à metade da população²⁴ e chega a ínfimos 2% no século XIX. Outro ponto importante a ser observado diz respeito à presença negra (africanos e afro-descendentes). A influência étnica da matriz africana se espalha pelo território nacional de tal forma que no final do século XIX, africanos e afro-descendentes chegam a compor quase 60% da população. Além disso, são eles o combustível básico do desenvolvimento econômico do Brasil colonial, estando presentes desde as *plantations* e casas grandes, durante o ciclo da cana-de-açúcar, até as minas de ouro e, mais tarde, nas fazendas de café.

Sobre o domínio da língua portuguesa nesse período, é mais acentuado durante dois momentos: o ciclo da mineração, século XVIII, no qual se registra uma considerável onda migratória lusófona atraída pela “febre de riqueza”; e a chegada da família real, no século

²³ A interferência do contato na formação do PB será oportunamente discutida.

²⁴ Vale salientar que o percentual diz respeito aos índios “integrados”, ou seja, índios que estavam agregados ao bojo das relações do sistema colonial. Assim, o percentual apresentado não recobre as comunidades indígenas isoladas, ou que se mantiveram à margem da realidade colonial.

XIX, concentrando-se no Rio de Janeiro, como um processo de relusitanização cultural e linguística (LUCCHESI, 2006). Esses eventos, além de aumentar o insumo de falantes da língua alvo, modificaram consideravelmente a configuração linguística e socioeconômica da colônia, de modo a conduzir ao colapso do sistema escravista.

Após a abolição da escravatura, os negros africanos, desqualificados e discriminados, encontraram poucas oportunidades de inserção na sociedade e, sobretudo, no mercado de trabalho. A esse respeito, Oliveira (1988, p. 11) enfatiza: “Tornar-se liberto não era o mesmo que tornar-se livre. Desta distinção encarregava-se a sociedade escravista de modo a perpetuar no ex-escravo as marcas de sua antiga condição servil”.

Assim, os negros libertos protagonizavam um movimento de liberdade aparente, pois são deslocados das senzalas para base da pirâmide social, no que poderíamos designar uma escravidão social, na qual o acesso à cidadania e aos direitos básicos era visivelmente distinto da condição de indivíduo livre. Não obstante, a Abolição da Escravatura, os negros são mantidos à margem do desenvolvimento:

Alguns receberam doações de terra e se dedicaram a uma cultura de subsistência, ficando à margem do processo produtivo. Mais marginalizados foram aqueles que abandonaram as grandes fazendas e se deslocaram ainda mais para o interior, em busca de locais ermos, onde pudessem livremente subsistir; confinando com a situação em que já se encontravam as comunidades quilombolas, de escravos foragidos. Se se dirigiam para as cidades, os ex-escravos eram mantidos nas periferias, em situação de grande miséria e marginalidade. (LUCCHESI, 2009, p. 52)

Esta e outras constatações pautam as conclusões a que chega Mattos e Silva (2006):

A presença maciça dos africanos e afro-descendentes que a demografia histórica demonstra; a atuação constante dos escravos nas grandes frentes de economia da colonização; a mobilidade geográfica, decorrente das vicissitudes da vida econômica de seus senhores e da economia brasileira; os diversificados e múltiplos papéis por eles desempenhados na sociedade rural e urbana; o significado social e linguístico dos espaços ilegítimos da escravidão permitem embasar o meu ponto de vista interpretativo de que é esse segmento numeroso e operante – os africanos e afro-descendentes – o agente principal da difusão do português no território brasileiro, na sua face majoritária, a popular ou vernácula. (MATTOS E SILVA, 2006, p. 106).

Considerações como essas levaram alguns linguistas a atribuir ao contato a responsabilidade pela diferenciação entre o PB e o PE. Mais do que isso, discute-se uma origem bifurcada do PB, haja vista as diferenças substanciais encontradas entre suas variedades culta e popular, conforme discutiremos na próxima seção.

3.2 A REALIDADE LINGÜÍSTICA DO PB COMO RESULTADO DE UM PROCESSO DE POLARIZAÇÃO SOCIOLINGÜÍSTICA

A designação *português brasileiro* sustenta-se no fato de que existem diferenças latentes entre este e a variedade europeia do idioma. As diferenças que separam as duas variedades não estão circunscritas apenas ao léxico, mas atingem também a fonologia, a morfologia e a sintaxe. Conforme destaca Teyssier (1997 [1980]), o português brasileiro figura, ora como inovador, ora como conservador no conjunto das mudanças apresentadas.

Destacadamente, alguns aspectos da morfossintaxe do PB têm chamado a atenção de inúmeros estudos, tais como: a concordância verbal e nominal, a cliticização, a realização do sujeito nulo, a flexão de caso dos pronomes, a ordem dos constituintes.

Mesmo encerrando-nos no PB, identificamos diferenças dialetais que demarcam variações diatópicas e, sobretudo, diastráticas no seu interior. Nesse ponto, recorreremos ao conceito de *norma*, inicialmente esboçado por Coseriu (1979 [1921]) e retomado por Lobo e Lucchesi (1988), através do qual distinguem as normas *cultas* e *populares* brasileiras. Cumpre salientar que o conceito de norma adotado pelos autores delimita-se a partir da identificação de “sistemas de avaliação social da variação linguística, associados a processos independentes de mudança linguística” (LUCCHESI, 2006, p. 90).

Lucchesi (2003, 2004b, 2006) defende, portanto, o argumento de que a realidade linguística brasileira é composta por um todo plural e polarizado. O argumento do autor fundamenta-se na identificação de dois polos com sistemas de avaliação e processos de mudança distintos. Assim, distinguir-se-iam no PB dois subsistemas: a *norma culta*, característica dos segmentos plenamente escolarizados e economicamente favorecidos, e a *norma popular*, concernente à maioria da população, desprovida de recursos financeiros e carente de educação formal.

Em linhas gerais, o que Lucchesi propõe é uma configuração do PB como um diassistema, composto de dois polos, ou subsistemas, que se diferenciam tanto em seus usos, quanto na avaliação subjetiva das variantes e nas tendências de mudança. A identificação desses subsistemas distintos, concebidos como unidades autônomas de análise, parte da observação e análise das evidências empíricas, bem como de eventos e desdobramentos da sócio-história do Brasil, como se lê nas palavras de Lucchesi (2006, p. 87-88):

[...] desde o início da colonização até a Proclamação da República, enquanto uma reduzida elite concentrada nos incipientes centros urbanos guardava uma profunda fidelidade aos modelos de uso da língua proveniente de Portugal, nas imensidões do interior do Brasil grandes contingentes de índios aculturados e negros africanos adquiriam a língua portuguesa em condições as mais precárias; e essa língua segunda defectiva se ia convertendo em modelo para a nativização do português entre os descendentes mestiços e endógamos desses segmentos, desencadeando profundas alterações na gramática da língua portuguesa assim adquirida, socializada e nativizada.

Assim, enquanto as normas cultas, praticadas pelas elites do referido período, espelhavam-se no modelo lusitano, as normas populares foram adquiridas de forma mais irregular, com a interferência do contato com aloglotas e sem uma ação normatizadora.

Por esse motivo, a configuração da sócio-história do PB constitui, para o autor, um diassistema formado por dois polos opostos, do ponto de vista da avaliação e das tendências de mudança. Na fala culta, observa-se um maior emprego da morfologia flexional, como no caso da concordância verbo-nominal, de modo que o afastamento do padrão é avaliado negativamente, podendo chegar a constituir um estereótipo; do outro lado, na norma popular, a não aplicação das regras de concordância constitui o padrão mais comum, não configurando um estereótipo nesse universo.

O processo de urbanização, intensificado a partir de 1940, bem como a democratização do sistema de ensino e o enorme crescimento dos meios de comunicação, especificamente, da mídia televisiva, atuaram como fator de atenuação entre os polos, em função da forma como atingiram os segmentos populares, ampliando o contato dos falantes do polo vernáculo com os modelos cultos do PB. Por outro lado, a norma culta, originalmente decalcada do modelo lusitano, também sofreu a influência das variedades populares, e passou, por volta do século XX, a uma tendência de afastamento da norma padrão. Os dois polos do sistema, teriam, portanto, tendências de mudança distintas: enquanto as normas populares, seja pela escolarização, seja pela influência da mídia, registram uma tendência a adquirir as variantes cultas; as normas cultas, caminham no sentido inverso, exibem uma mudança que conduz ainda mais ao afastamento da norma padrão. Em parte, esse último comportamento nos conduz a uma reflexão social. A educação de nível superior, anteriormente circunscrita aos falantes das elites, atingiu uma considerável expansão, a qual abarcou, independente do nível de qualidade, falantes oriundos das camadas populares, moradores e filhos de falantes de português popular.

Apesar das mudanças mais recentes na configuração social do Brasil, a realidade linguística brasileira ainda pode ser equacionada em termos de uma polarização, conforme

sumarizado na Figura 1, em que as setas (*a* e *b*) representam as tendências opostas de mudança, mas a linha que separa os polos encontra-se descontínua, uma vez que estes apresentam clara inter-influência.

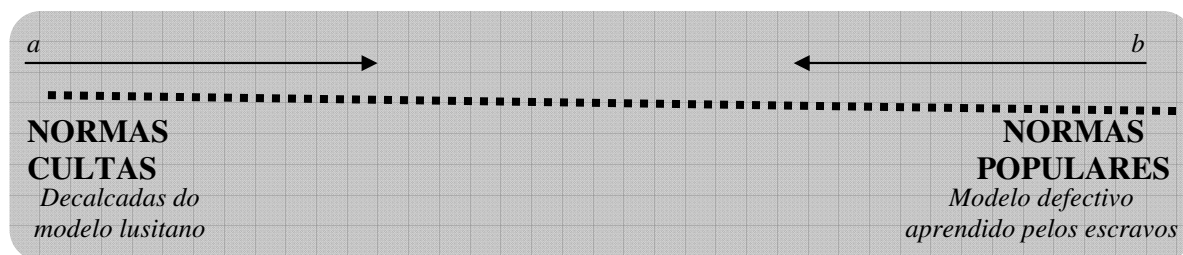


FIGURA 1: A polarização sociolinguística do PB.

Partindo do pressuposto de que as diferenças que marcam as variedades populares do PB seriam fruto do contato linguístico do português com falantes de outras línguas, sobretudo africanas, lançamos mão do conceito de *transmissão linguística irregular* como opção explicativa da realidade linguística nacional. Esta proposta encontra-se inserida numa perspectiva mais ampla de análise e compreensão do processo de contato massivo e abrupto entre línguas e é apresentada por Lucchesi (2000; 2003; 2009).

3.2.1 A transmissão linguística irregular

De um modo geral, a história das línguas humanas não está isenta de contato, seja direto ou indireto, seja intra ou interlinguístico. As línguas pidgins e crioulas, no entanto, são exclusivamente “produzidas” a partir de situações particulares de contato abrupto e massivo. O produto desse tipo de situação, além de original, é marcado por um processo histórico particular. As línguas pidgins e crioulas surgem de uma situação de contato entre falantes de línguas ininteligíveis entre si, obrigados a conviver e interagir por determinado período. O tipo de relação estabelecida, bem como a duração e as condições vão interferir no produto desse tipo de contato linguístico.

De acordo com Lucchesi (2003; 2009), para além da *pidginização* e *crioulização*, as situações de contato entre línguas podem ser definidas de forma mais ampla, tendo em vistas as possíveis mudanças linguísticas oriundas deste tipo de processo histórico. Partindo dessa reflexão, o autor propôs o conceito de *transmissão linguística irregular* – TLI, que recobre as

situações em que uma grande parte da população de adultos de línguas distintas precisa aprender uma segunda língua emergencial em circunstância de sujeição. Esse tipo de situação aconteceu especialmente durante o colonialismo europeu, sobretudo no contexto das *plantations*, nas quais esteve presente um grande contingente de escravos africanos, retirados de suas terras e culturalmente expropriados, sob uma violenta condição de dominação e sujeição.

A designação *transmissão linguística irregular*, que, a nosso ver, não carrega nenhuma conotação pejorativa, distingue-se da transmissão regular pelo seguinte fato: nas situações regulares de transmissão linguística de geração para geração, a língua adquirida pelas crianças corresponde à língua materna dos seus pais. No caso da transmissão linguística irregular, tem-se como modelo de aquisição segunda língua adquirida pelos pais de modo defectivo. Dessa forma, o processo de transmissão linguística irregular pode ser sumarizado em dois planos: a aquisição de uma segunda língua por uma população de falantes adultos sob condições precárias e a nativização desta como língua materna das novas gerações.

Desse modo, nesse processo, estão presentes, de um lado, os dados linguísticos primários (DLP) para a aquisição da língua materna, que são defectivos, no sentido que não contêm certos elementos gramaticais básicos; e, de outro, uma comunidade de falantes adultos que precisa apreender um código emergencial a partir de um modelo pouco acessível da língua do dominador (a *língua alvo*). A fase inicial do processo caracteriza-se pela aquisição precária da língua alvo pelo agrupamento de falantes adultos. O modelo adquirido caracteriza-se pela: i) perda de morfologia flexional e sintaxe aparente²⁵, ii) alteração do valor de parâmetros sintáticos em favor de valores não marcados; iii) e gramaticalização de itens lexicais a fim de substituir gerações. Assim, o resultado do processo é determinado pelo nível de erosão gramatical sofrida pela língua alvo no estágio inicial do contato, associado à reestruturação (ou não) da gramática original. Este modelo defectivo inicialmente gerado determina o grau de alterações que a língua alvo sofrerá na sua transmissão para as novas gerações (Lucchesi, 2003).

Dessa forma, além das situações prototípicas de crioulo e pidgin, a situação de contato pode produzir, também, “processos significativos de variação e mudança na estrutura linguística das variedades de língua alvo que surgiram nesse tipo de situação” (Lucchesi, 2003, p. 276), como aconteceu com o português no Brasil.

²⁵ Marcação explícita, do ponto de vista morfofonológico, de processos gramaticais.

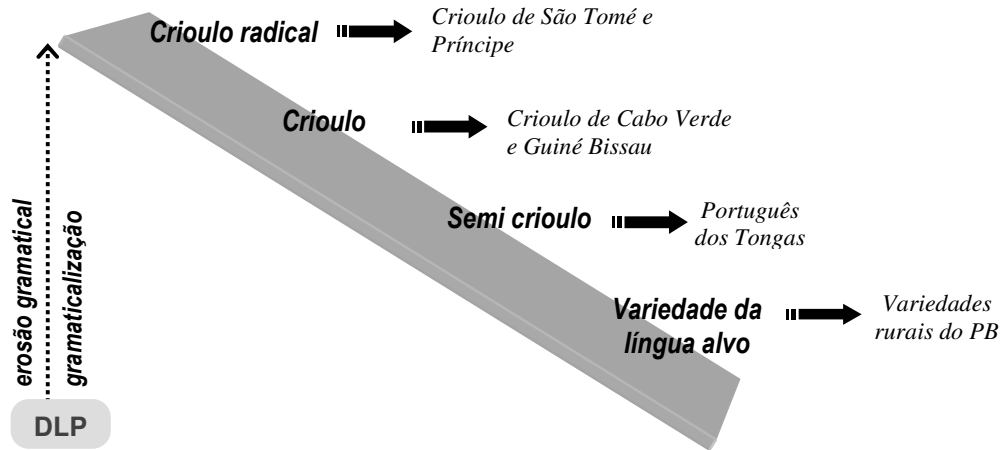


FIGURA 2: Transmissão Linguística Irregular²⁶.

Conforme a Figura 2, nas situações prototípicas de pidginização e crioulição, (situações mais radicais de transmissão linguística irregular) o acesso aos dados linguísticos primários (DLP) é muito baixo, o que conduz à profunda erosão gramatical e, conseqüentemente, à reestruturação da gramática a partir de elementos da língua de substrato. Nos processos de transmissão linguística irregular mais leves, apresentados na base da Figura 2 e exemplificados pelas variedades populares do português do Brasil, a erosão gramatical da língua alvo é menos profunda, deixando algumas marcas, sobretudo na morfologia aparente, mas não chegando a uma reestruturação da nova variedade a partir de elementos do substrato.

Tomando esta proposição como hipótese explicativa para a constituição do PB, buscaremos observar em que medida podem ser encontradas marcas do contato na configuração da ordem dos constituintes V(verbo) e S(sujeito) nas variedades populares do português do interior do estado da Bahia, haja vista que a redução da morfologia flexional constitui um resultado comum das situações de contato, e que essa incide diretamente na fixação do padrão SV. Passemos, pois, à delimitação e apresentação do nosso recorte de observação.

²⁶ A figura foi elaborada com base nas anotações das aulas do professor Dante Lucchesi no curso de *Constituição do português do Brasil*, ministrado na graduação de letras – UFBA.

3.3 O RECORTE DE OBSERVAÇÃO

Os *corpora* analisados nesta pesquisa fazem parte do acervo de fala popular do Projeto Vertentes. Partindo do pressuposto de que o contato com as línguas africanas teria afetado o português brasileiro, o projeto analisa aspectos da morfossintaxe do português popular, buscando identificar a presença de características que o diferenciem do português culto, e que seriam provenientes, sobretudo, do processo de TLI - *transmissão linguística irregular* (cf. Seção anterior).

3.3.1 O Projeto Vertentes

Implementado na Universidade Federal da Bahia a partir de 2002, o Projeto *Vertentes do Português Popular do Estado da Bahia*²⁷, coordenado pelo Professor Doutor Dante Lucchesi, tem se destacado como uma importante fonte para o conhecimento da sócio-história do português brasileiro, mais especificamente do português popular do Brasil.

O Projeto tem formado um expressivo acervo da fala popular do Estado da Bahia. São mais de 150 entrevistas, entre comunidades rurais isoladas, comunidades do recôncavo baiano e comunidades urbanas periféricas. Além disso, os estudos desenvolvidos no âmbito do Projeto em diferentes níveis – iniciação científica, dissertações e teses – têm prestado relevantes contribuições para a investigação e compreensão do português brasileiro (MENDES, 2009; PONTE, 2008; ANTONINO, 2007; SILVA, 2005; SILVA, 2004).

A importância do Projeto se reflete no apoio financeiro e institucional que tem recebido. Instituições como a Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) investiram e investem na instalação e desenvolvimento das atividades do Projeto, bem como na manutenção de uma equipe de bolsistas de iniciação científica e na concessão de bolsas de mestrado, doutorado e de Produtividade em Pesquisa.

Desenvolvendo pesquisas à luz do modelo sociolinguístico e baseado na hipótese da TLI, o Projeto vem fazendo história na tarefa de escrutinar as origens do português brasileiro.

²⁷ Para mais informações consultar www.vertentes.ufba.br.

3.3.2 A constituição do *corpus*

Nessa perspectiva, o Projeto conta com amostras de fala vernácula de duas²⁸ variedades da norma popular brasileira: o *português afro-brasileiro*, representado por comunidades isoladas remanescentes de quilombos, e o *português popular*, representado por municípios do interior do estado da Bahia. Fazem parte do primeiro grupo as comunidades de Helvécia, Cinzento, Sapé e Rio de Contas²⁹. Acredita-se que essas comunidades, por serem constituídas por descendentes diretos de escravos africanos, apresentariam marcas mais proeminentes de processos de mudança derivados da TLI. A segunda variedade é estudada a partir da fala de informantes com pouca ou nenhuma escolaridade dos municípios de Santo Antônio de Jesus e Poções. Nessa amostra, diferenciaram-se falantes da sede e da zona rural dos municípios. Esta estratificação visou observar no português popular do interior do Estado o contínuo de urbanização (BORTONI-RICARDO, 2006).

Os *corpora* são formados por falantes nascidos nas localidades escolhidas e foram estratificados segundo as variáveis *sexo* e *faixa etária*, sendo estabelecidas três faixas: Faixa I (20-40 anos); Faixa II (41-60 anos); Faixa III (mais de 60 anos), de que resultaram seis células, para a amostra de cada localidade. Em cada célula, foram selecionados dois informantes. Seguindo os pressupostos metodológicos da sociolinguística variacionista, compuseram-se as amostras em gradação geracional, de modo a identificar nas três gerações os comportamentos linguísticos correspondentes a três sincronias distintas e, conseqüentemente, a partir do estudo em *tempo aparente* (cf. Seção 1.3 desta dissertação), traçar o processo de mudança que se instaurou nessas variedades linguísticas.

Além das variáveis estratificadas apresentadas acima, são observadas as variáveis *escolaridade* (diferenciando-se falantes analfabetos e semi-analfabetos) e *estada fora da comunidade* (com a distinção entre falantes que nunca viveram fora da comunidade e falantes que estiveram fora por um período superior a seis meses).

Esta dissertação utilizou para sua análise os *corpus* do português *afro-brasileiro* das comunidades Cinzento (12 inquéritos) e Sapé (12 inquéritos) e o *corpus* do português *popular* dos municípios de Santo Antônio de Jesus (24 inquéritos) e Poções (24 inquéritos). Dessa

²⁸ Atualmente o projeto encontra-se numa terceira fase. Após a investigação do português popular do interior, escurtina-se agora o espraioamento dos traços do contato no português popular da capital do estado. Nesta etapa, tem sido constituído um *corpus* com amostras de fala de comunidades periféricas da cidade de Salvador.

²⁹ Das quatro comunidades, serão utilizadas neste trabalho apenas duas: Cinzento e Sapé.

forma, o conjunto amostral deste trabalho perfaz um total de 72 inquéritos, cf. Tabela 3 abaixo:

TABELA 3: *Corpus* delimitado para o estudo.

COMUNIDADE	Nº DE INQUÉRITOS
Cinzeno	12
Sapé	12
Poções	24 (12 sede/ 12 rural)
Santo Antonio	24 (12 sede/ 12 rural)
TOTAL	72 inquéritos

A opção dessa distribuição teve o objetivo de cobrir as variedades do português popular observadas no Projeto Vertentes, buscando identificar possíveis contrastes entre o *português afro-brasileiro* e o *português popular do interior*. Doravante, assume-se essa nomenclatura para a designação dos *corpora* analisados. As comunidades estudadas encontram-se geograficamente dispostas conforme a Figura 3, abaixo:

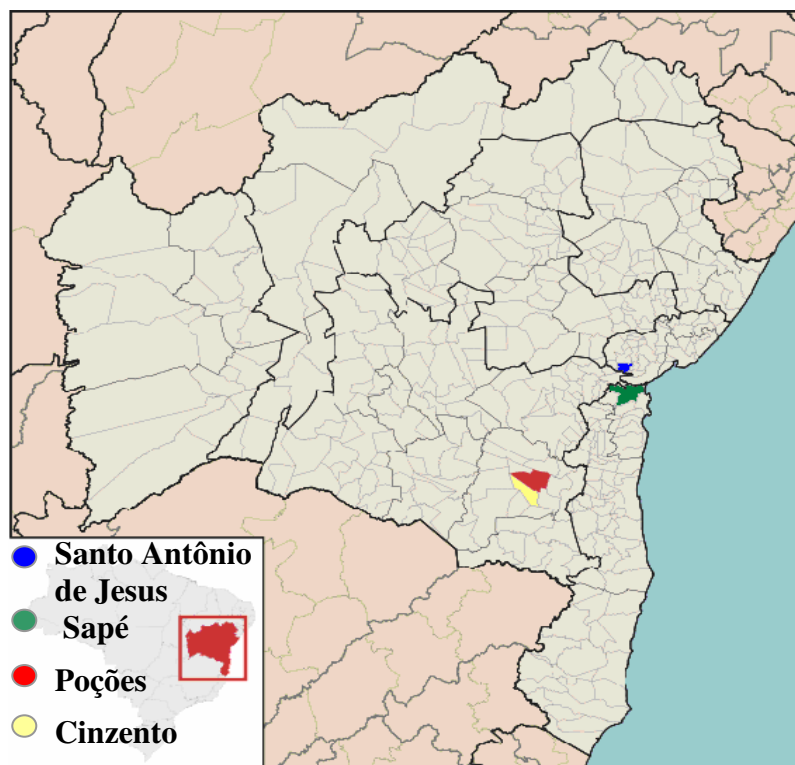


FIGURA 3: Disposição geográfica das comunidades.

3.3.3 O português afro-brasileiro: as comunidades remanescentes de quilombos.

De acordo com o decreto 4887/2003, art. 2º, as comunidades quilombolas são:

grupos étnico-raciais, segundo critérios de auto-atribuição, com trajetória histórica própria, dotados de relações territoriais específicas, com presunção de ancestralidade negra relacionada com a resistência à opressão histórica sofrida.

Tradicionalmente conhecidos como refúgio de escravos fugidos, as comunidades quilombolas tiveram diferentes origens. A maioria das comunidades quilombolas foi formada após a abolição da escravatura. Isto porque, após terem adquirido a liberdade, os ex-escravos ficaram à deriva social, não tendo para onde ir, nem onde trabalhar. Uma das soluções encontradas foi buscar integração nos quilombos estabelecidos por negros fugidos no período colonial. Além disso, migraram para terras inabitadas em áreas isoladas e para fazendas falidas. Algumas comunidades ainda tiveram origem em doações de terras que foram concedidas aos escravos alforriados como pagamento de dívidas de trabalho.

Em geral, essas comunidades apresentam características muito parecidas. A identidade étnico racial³⁰ é um traço característico, a população é sempre formada por indivíduos negros, ocorrendo pouca ou nenhuma miscigenação em virtude da ancestralidade comum, notadamente africana. Além disso, mas do que um grupo de indivíduos ligados por um mesmo processo histórico de formação e que partilham o mesmo território, essas comunidades podem ser definidas como grupos que se fundamentam em torno de uma mesma forma de organização política, cultural e social, além dos aspectos religiosos. Dessa forma, não é apenas a circunscrição territorial e a identidade racial que definem essas comunidades, mas também os aspectos sócio-culturais tangenciais.

A agregação do termo “remanescentes” justifica-se na necessidade de se evidenciar a descontinuidade histórica com o processo de escravidão. Isso fica evidente no fato de os membros dessas comunidades nem sempre terem uma consciência clara de sua ascendência africana, uma vez que esta tem se perdido com a morte dos indivíduos mais velhos da comunidade. Ademais, a adoção oficial do termo **remanescentes de quilombo** surge na

³⁰ A identidade étnica ou etnicidade atualmente é um termo que não concorda necessariamente com as categorias raciais, mas destaca-se como a classificação de grupos ligados por laços culturais em comuns. Conforme Hall (2003, p. 67), “a ‘etnicidade’ gera um discurso em que a diferença se funda sob características culturais e religiosas”.

tentativa de desvencilhar o surgimento exclusivo dessas comunidades como quilombos, ou seja, reduto de escravos fugidos, e garantir às comunidades afro-brasileiras os direitos legais.

Do ponto de vista linguístico, essas comunidades teriam se mantido distantes dos centros difusores da norma culta, tornando-se um ambiente propício para o desenvolvimento de uma variedade diferenciada do português brasileiro. O relativo isolamento dessas comunidades, bem como a forte influência da matriz africana teriam produzido condições favoráveis ao processo de transmissão linguística irregular.

Atualmente, há alguma controvérsia quanto ao número de comunidades no Brasil, mesmo porque nem todas foram oficialmente reconhecidas. Estimativas extra-oficiais apontam números entre 2.000 e 3.000 comunidades. Sabemos, porém, que as comunidades quilombolas encontram-se espalhadas por quase todo o território nacional. Só no estado da Bahia são mais de 300 comunidades, sendo 159 oficialmente reconhecidas pela Fundação Cultural Palmares³¹, órgão do Ministério da Cultura responsável pelo reconhecimento e valorização da cultura afro-brasileira.

A importância da certificação, dentre outros fatores, está na defesa dos direitos dos habitantes dessas comunidades. Às comunidades reconhecidas remanescentes de quilombos é garantida a emissão de título de propriedade de terras (cf. art. 68/ ADCT³²). Além disso, a esses indivíduos socialmente expropriados é garantido o direito de acesso a bens e serviços sociais básicos como infra-estrutura, saúde e educação.

Cabe ressaltar que essas comunidades se encontram cada vez menos isoladas e mais influenciadas pelo entorno urbano, sobretudo após a inclusão do assunto na pauta das políticas públicas, como exemplo do Programa Brasil Quilombola (PBQ) do governo Federal.

Os informantes das comunidades afro-brasileiras ora estudadas apresentam em seu perfil pouca mobilidade e uma rede de relações concentrada na região, com pouca saída do município.

³¹ www.palmares.gov.br

³² Ato das Disposições Constitucionais Transitórias (Constituição Federal 1988)

3.3.3.1 Cinzento

Politicamente ligada ao município de Planalto, no sudoeste baiano, a comunidade afro-brasileira de Cinzento localiza-se há 450 km da cidade da capital do estado. A origem do nome da comunidade não é muito clara. A informante mais antiga do local (107 anos) fornece uma explicação ao documentador (cf. trecho da transcrição CZ-10 abaixo).

DOC: Me conta aí, com'ê... por que que Cinzento chama Cinzento?

INF: Diz o povo que... que, quano eles chegô, achô o Cinzento, mas diz que é poque ele não... os home mais veio diz que sumiu um boi...

DOC: Hum...

INF: ...O boi num 'pareceu mais não... ININT... o Cinzento era mato.

DOC: Hum.

INF: A [cada madêra assim], ó. Aí, o boi sumiu muntos tempo. Quano foi quato... três ano... quano foi com uns quato ano, o boi 'pareceu. Quano 'pareceu, 'pareceu Cinzento, aí eles foi veno o boi Cinzento, boi Cinzento, aí ficô. (CZ-10)

O registro além de revelar a resistência de tradições tipicamente rurais, como as lendas, revela a configuração sócio demográfica do local: “Cinzento era mato”. Este depoimento demonstra o grau de isolamento da comunidade. Segundo Silva (2005), a geografia do local é marcada por um terreno pedregoso e pela escassez de fontes de água.

Neste cenário de difícil acesso, teriam se fixado escravos e crioulos que partilhavam, além dos traços raciais, uma realidade de subjugação e resistência. As marcas do passado quilombola deixaram vestígios evidentes na comunidade. Não raro, é possível encontrar informantes que relatam histórias de pais e avós escravos. Além disso, como bem salienta Silva (2005, p. 189), o fenótipo local é bem marcado, “os cinzentenses são pretos de pele bem escura. Todos muito parecidos devido à endogamia de segundo grau”. O traço de endogamia, segundo o autor, teria sido usado como recurso de auto-proteção da comunidade, a fim de evitar a repartição de terras.

A comunidade recebeu a certificação de remanescente de quilombo em 19/08/2005, conforme informações da Fundação Cultural Palmares.

3.3.3.2 Sapé

Também conhecida como Sapé Grande, a comunidade de Sapé faz parte do município de Valença-BA, no recôncavo baiano.

Segundo Silva (2004), principal responsável pela coleta das entrevistas, a formação histórica da comunidade é marcada pelo processo de escravidão. Grupos de negros, após a abolição da escravatura, teriam se fixado nessa região. Silva (2004, p. 80) nos explica a gênese dessa população:

Os fazendeiros, para quitar suas dívidas com os ex-escravos, “doaram”-lhes pedaços de terra improdutivos, mais altos, sem trato e distantes da água. Esses ex-escravos ali permaneceram e casaram entre si aumentando a população local, e curiosamente, mantendo-se isolados. De acordo com Sr. João Barreto, foram cinco negras de uma família de oito que deram origem à comunidade (...)

O povoado constitui uma região muito carente de infra-estrutura e recursos. Basicamente, as principais atividades são a agricultura e a produção de farinha. Os moradores sobrevivem essencialmente do trabalho nas roças. Somente no dia 05 de maio de 2009, a comunidade recebeu a certificação oficial de reconhecimento como comunidade remanescente de quilombo, conferida pela Fundação Palmares.

3.3.4 O português popular: as comunidades rurais do interior do estado da Bahia.

Historicamente, a ocupação do espaço brasileiro dá-se a partir dos conglomerados rurais. Os engenhos e povoados fixados no período colonial foram as bases de ocupação do território nacional. Isto se reflete no fato de até as primeiras décadas do século XX o Brasil se configurar como um país essencialmente rural (cf. Tabela 4).

Bortoni-Ricardo (2006), em análise aos dados censitários identifica que a mudança na configuração demográfica brasileira se processa gradualmente, acentuando-se sensivelmente a

partir da década 60, graças às mudanças engendradas no seio da sociedade brasileira no período, que compeliram a forte tendência de migração urbana³³.

TABELA 4: Distribuição da população brasileira – dados dos censos brasileiros.

Período	POPULAÇÃO	URBANA
1890	14.333.915	6,8%
1920	30.635.605	10,7 %
1940	41.236.315	31,29%
1950	51.944.397	36,16 %
1980	120 milhões	67,60 %

Fonte: IBGE

A urbanização do país acontece em paralelo ao desenvolvimento industrial e sociocultural da sociedade brasileira, assumindo um ritmo acelerado, de forma que, atualmente podemos pensar não somente numa migração dos indivíduos do campo para a cidade, mas na migração da própria cidade para o campo. Em outras palavras, falamos da urbanização extensiva, que, segundo Sathler (2005, p. 81), baseado na formulação de Montemór (1994), “refere-se ao avanço do tecido urbano que extrapola os limites das cidades com a geração de novas centralidades urbanas, sendo um processo econômico-espacial”. Dessa forma, a delimitação de territórios rurais e urbanos brasileiros configura-se atualmente uma tarefa de grande dificuldade, mesmo porque há pouca concordância teórica na aceção desses termos.

Atualmente, a delimitação entre urbano e rural não é tópico pacífico na geografia. Oficialmente, a diferenciação é estabelecida no decreto Lei 311 de 1938, o qual define como urbanas as cidades (ou seja, as sedes dos municípios) e as vilas (sedes dos distritos), sendo o resto do território considerado rural. Ainda segundo o decreto, cumpre às prefeituras a responsabilidade de definir legalmente as áreas incluídas no perímetro urbano. De acordo com a classificação oficial, as sedes dos municípios de Santo Antonio e Poções seriam tão urbanas quanto a capital do estado, Salvador.

Essa definição, adotada pelo IBGE, não alcança concordância no âmbito dos estudos sócio-espaciais, sendo, ao contrário, fortemente rejeitada pela sua motivação essencialmente política e anacronismo.

³³ Tais como a industrialização

No âmbito teórico, de acordo com Sather (2005), destacam-se duas correntes: uma orientada por uma visão dicotômica que opõe campo e cidade; outra fundamentada numa noção de *continuum*, que propõe uma aproximação entre o espaço rural e a realidade social e econômica urbana.

A primeira corrente não dá conta da complexa relação estabelecida nos espaços sócio-geográficos da sociedade moderna, sobretudo tendo em vista o fenômeno da globalização e o extremado desenvolvimento e alcance dos meios de comunicação. Isto nos leva a crer que esta não é uma posição adequada à realidade brasileira.

Na esteira da idéia de *continuum*, podemos identificar uma posição dual que opõe dois polos diferenciados por traços específicos marcantes. Dentre os traços essencialmente rurais, destaca-se a proeminência da atividade agrícola e o modo de vida essencialmente campestre. Essa abordagem, predominante nas primeiras décadas do século XX, vai dando lugar a uma visão mais gradualista e interpenetrada do processo, sobretudo com o incremento da mecanização da agricultura.

Conforme salienta Sather (2005, p. 82):

O crescimento das cidades, a industrialização da agricultura e o transbordamento do urbano nas áreas rurais, verificados em vastas regiões do mundo, no decorrer do século XX, sugerem que a transição entre os espaços rural e urbano deve ser entendida de acordo com a formulação teórica do espaço *continuum*. Nessa perspectiva, a polarização antagônica é substituída por um gradiente de variações espaciais.

Embora se apresente mais coerente e reflexivo que o modelo dicotômico, a proposta de *continuum* não encerra de todo a carência metodológica de classificação, mesmo porque a própria definição do binômio rural/urbano é conceitualmente incompatível com a noção de *continuum*. No entanto, acreditamos que é esta a noção que melhor define a realidade sócio-demográfica brasileira e, num âmbito mais estendido, o próprio panorama de distribuição das variedades linguísticas do Brasil, tendo em vista a proposta de Bortoni-Ricardo para o *continuum* de urbanização (cf. Figura 4).

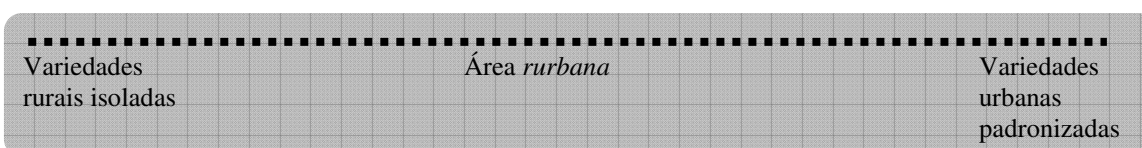


FIGURA 4: Distribuição das variedades linguísticas do PB, conforme o *continuum* de urbanização.
FONTE: Bortoni-Ricardo (2006, p. 52).

De acordo com a autora, no polo extremo esquerdo do *continuum*, estão as variedades rurais do português, usadas pelas comunidades isoladas (no nosso caso, as comunidades afro-brasileiras supracitadas); no polo oposto, temos as variedades urbanas diretamente influenciadas pelo padrão culto da língua (os falantes da capital Salvador, por exemplo); no centro do *continuum*, temos as comunidades *rurbanas* que seriam formadas:

[...] pelos migrantes de origem rural que preservam muito de seus antecedentes culturais, principalmente no seu repertório lingüístico, e as comunidades interioranas residentes em distritos ou núcleos semi-rurais, que estão submetidos à influência urbana, seja pela mídia, seja pela absorção de tecnologia agropecuária (BORTONI-RICARDO, 2006, p. 52).

Sob esse rótulo estariam as sedes e as zonas rurais dos municípios de Poções e Santo Antônio de Jesus e, numa adaptação da Figura 4, teríamos assim dispostas as nossas comunidades:

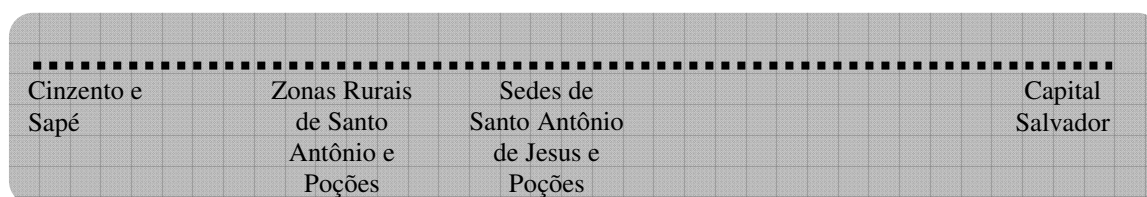


FIGURA 5: Disposição das comunidades rurais e afro-brasileiras no *continuum* de urbanização.

No polo extremo esquerdo do *continuum* de urbanização, temos as comunidades representativas do *português afro-brasileiro*; no extremo direito, a capital do estado, representativa, nesse plano, das variedades *cultas* do idioma; ao centro, as variedades do *português popular do interior*, as sedes e zonas rurais de Santo Antônio de Jesus e Poções.

3.3.4.1 Santo Antônio de Jesus

O município de Santo Antonio de Jesus é um dos mais importantes do recôncavo baiano, destacando-se como importante polo comercial e de serviços de sua microrregião. Situada à aproximadamente 200 Km da capital do estado, a cidade conta com uma população

de 77.368 habitantes³⁴. Segundo dados do IBGE, o município apresentou grau de urbanização³⁵ de 85,62% em 2000. Vale destacar, que tem sido cada vez mais frequente o deslocamento dos indivíduos da zona rural para a sede do município, motivados sobretudo pela necessidade de serviços básicos como saúde, educação e também pela busca por emprego.

Além do comércio, ponto alto da região, destacam-se como atividades econômicas do município a produção de gêneros agrícolas (produção de amendoim e limão e laranja), a pecuária (criação de bovinos e muares) e a produção de fogos de artifício.

Inicialmente, o território que deu origem ao município de Santo Antônio de Jesus pode ter sido habitado por índios cariris, tupinambás e aimorés (SANTOS, 1999). Os primeiros assentamentos populacionais aconteceram por volta do século XVIII. A princípio, o território vinculava-se ao município de Nazaré, alcançando independência política e institucional em 1880.

O desenvolvimento urbano da cidade se deu com a insurgência de atividades comerciais, desencadeado, sobretudo, pelo incremento das vias de transporte ferroviário e rodoviário.

Ao desenvolvimento alcançado pelo município estão atrelados problemas de crescimento populacional desordenado, tais como violência urbana e carência de infraestrutura. O município registrou, em 2000, densidade demográfica de 306,99 h/Km², taxa muito alta se comparada à taxa do estado no mesmo período (23,16 h/Km²).

3.3.4.2 Poções

O topônimo Poções, segundo a tradição oral, tem origem na grande quantidade de poços na região mais baixa do município.

A história do povoamento do município é marcada pelo processo de colonização, e data da segunda metade do século XVIII. Além da presença européia, há registros da presença

³⁴ As entrevistas do município foram coletadas em 2002, motivo pelo qual são utilizados aqui os dados do censo demográfico de 2000. Em 2007, a população do município apresentou um crescimento de 9%, segundo contagem do IBGE.

³⁵ Indica a porção da população residente nas áreas urbanas do município, no período considerado.

indígena e africana (SILVA, 2005). Inicialmente, Poções era um distrito vinculado a Vitória da Conquista. Sua independência política só é alcançada em 1883.

Situado na microrregião de Vitória da Conquista, o município dista 444 km de Salvador e conta com uma população de 44.213 habitantes, sendo a maioria concentrados na zona urbana (cf. IBGE, grau de urbanização de 71,93% em 2000). O município apresenta um desenvolvimento bem inferior ao município de Santo Antonio de Jesus, com densidade demográfica de 45,86 h/km².

O município de Poções tem como principais atividades econômicas a agricultura (produção de café, feijão, tomate, mandioca, banana, milho, aipim e mamona) e a pecuária, (sobretudo o rebanho). Ademais, destaca-se no setor de bens minerais pela produção de ferro e amianto. Do ponto de vista cultural, o município caracteriza-se pela manutenção de algumas tradições folclóricas, como o Reisado e a tradicional festa do Divino Espírito Santo, frequentemente referida pelos informantes durante as entrevistas.

Ao contrário de Santo Antônio de Jesus, Poções ainda guarda muitos traços da tradição rural. De acordo com a prefeitura do município³⁶, mesmo com o crescimento da cidade, pouco se modificou a paisagem local, a cidade ainda conserva um clima de calma, com pouca circulação de pessoas, e um forte apego às tradições religiosas.

Na zona rural do município, destaca-se o distrito de Morrinhos, região eminentemente agrária, mas de grande importância para a região, pois além dos insumos agrícolas que produz, concentra a barragem que abastece de água a cidade e regiões circunvizinhas.

3.4 A COLETA E A PREPARAÇÃO DOS DADOS

Toda pessoa que comece a estudar a língua em seu contexto social imediatamente se depara com o clássico problema metodológico: os meios empregados para coletar os dados interferem nos dados a serem coletados.

Labov (2008 [1972], p. 63)

³⁶ <http://www.poco.es.ba.gov.br/historia.htm>

As entrevistas coletadas nas comunidades apresentadas são do tipo diálogos espontâneos entre informante e documentador, com duração de aproximadamente uma hora, e visam captar o registro menos monitorado do informante.

Tendo em vista a necessidade de se captar o registro vernáculo do falante face ao problema metodológico que consiste a gravação das entrevistas, optou-se por algumas estratégias já consagradas no âmbito dos estudos linguísticos.

A fim de mitigar os efeitos do *paradoxo do observador*, as entrevistas foram gravadas conforme as seguintes orientações:

1º) O adentramento à comunidade de fala deu-se por intermédio de um membro da comunidade. Dessa forma, acreditou-se estabelecer uma relação mais próxima e confiável com o informante e assim dirimir o constrangimento e a formalidade durante a entrevista.

2º) As entrevistas foram realizadas de forma distensa e versaram sobre tópicos que motivassem a emoção do falante, tais como a exposição do informante a algum tipo de situação de risco, como morte e acidente, os problemas de saúde e outros enfrentados na vida pessoal, as condições de vida e moradia, as lembranças da infância e as festas locais.

Como se trata de um tipo de entrevista sem roteiro predefinido, contamos ainda com a sagacidade e percepção dos documentadores na identificação e abordagem de temas que demonstraram interesse do informante ao longo do diálogo.

As entrevistas de Cinzento e Sapé foram coletadas em 2002, em mídia analógica (fita magnética). As entrevistas de Poções e Santo Antônio de Jesus foram gravadas em 2003, em mídia digital (gravadores de *minidisc*).

Todas as entrevistas gravadas foram digitalizadas e submetidas ao tratamento acústico, a fim de mitigar os ruídos e melhorar a qualidade do som. Vale salientar, que esse processo inicia-se durante a gravação dos inquéritos, evitando, por exemplo, o manuseio do microfone de forma excessiva e dando preferência à gravação em ambientes fechados e silenciosos. O processo de edição, levado a cabo através da aplicação de *software* específico, removeu, quando possível, a interferência de ruídos tais como vento, animais, dentre outros, que poderiam de alguma forma, interferir na clareza do áudio e dificultar a fidedignidade da transcrição.

Após esta etapa, as entrevistas foram transcritas, seguindo uma chave de critérios previamente estabelecida pelo coordenador do projeto (Professor Doutor Dante Lucchesi). Dentre os critérios definidos, destaca-se o registro de variantes morfológicas e fonológicas típicas das variedades populares do português brasileiro. Quanto à pontuação, além da adoção de critério eminentemente sintático, são marcadas hesitações e intervenções pragmáticas típicas da linguagem falada (cf. Chave de Transcrição do Projeto Vertentes, disponível na Internet, no *site*: www.vertentes.ufba.br/chave.htm).

3.5 A VARIÁVEL DEPENDENTE

Definiu-se como *variável dependente* desta análise a ordem do sujeito em relação ao verbo nas sentenças declarativas e interrogativas, considerando como variantes as colocações pré ou pós-verbal do sujeito, exemplificadas a seguir:

- (34) Ordem direta (sujeito-verbo – SV): ... *a mãe de Ana Maria morreu* (POR-12)
 (35) Ordem inversa (verbo-sujeito – VS): há dois anos atrás, **morreu** *meu sogro*.
 (POS- 06)

A variável binária foi coletada nos 72 inquéritos selecionados para o estudo. Tendo em vista a abundância de ocorrências válidas para o estudo do fenômeno, optamos por estipular uma cota de 50 ocorrências por inquérito.

Em face da importância relativa que possui o contexto interrogativo no condicionamento da ordem do sujeito, foram recolhidas, para além das 50 ocorrências da cota fixada, todas as sentenças interrogativas –QU (*Wh-questions*) de cada entrevista, conforme exemplos abaixo:

- (36) *Como vai sai(r) uma mocinha de doze anos, catorze ano, pedindo carona na estrada?* (POR-07)
 (37) Tu mora onde? (POS-01)

O levantamento exaustivo das interrogativas está associado também à baixa frequência desse tipo de frase no gênero de texto que constitui a amostra de fala analisada: a entrevista

monitorada por um pesquisador, em que o entrevistado (i.é., o informante, de cuja fala são eliciadas as ocorrências) faz poucas perguntas, já que o seu papel nesse gênero de texto é, ao contrário, o de responder às perguntas.

3.5.1 Dados descartados

Durante o processo de levantamento da variável dependente fez-se necessário estabelecer alguns critérios para a delimitação do universo de observação, alguns previamente estabelecidos, outros identificados no desenrolar do trabalho. Desse modo, não foram consideradas as seguintes ocorrências:

a) Construções com verbos existenciais, já que nesses casos a ordem VS é categórica.

(38) *Hoje existe a raposa ainda também.* (SAR- 07)

b) construções declarativas clivadas ou pseudo clivadas com o expletivo *é que*, tendo em vista o fato de essas construções representarem uma estratégia específica de focalização.

(39) *A Florêncça é que num me abusa muit.o* (SAR-01)

c) sentenças interrompidas ou incompletas.

(40) *por todo lado 'cê dava... 'cê encontrava serviço* (POR-09)

Neste caso, foi considerada apenas a retomada ... *'cê encontrava serviço.*

d) Construções identificacionais (ou equativas).

Construções copulativas com dois SNs introduzidos pelo artigo definido e frases copulativas em que um dos constituintes é um pronome pessoal:

(41) *O difícil é a ladêra* (SAR 02)

(42) *meu primêro filho é ela* (SAR 06)

Embora sejam apresentados alguns testes para a identificação do sujeito dessas construções (MATEUS, 2003, p. 544-546), pela dificuldade desta identificação e pela particularidade do contexto, optamos por desconsiderar estas ocorrências que poderiam enviesar os dados da análise.

3.6 A CHAVE DE CODIFICAÇÃO: ALGUMAS CONSIDERAÇÕES

Devidamente definido o universo de observação e delimitado o fenômeno, partiu-se para a análise dos dados. As ocorrências levantadas foram codificadas com base em uma chave de codificação previamente definida. A partir do estudo da bibliografia e de reflexões pessoais, foram definidos grupos de fatores que estariam condicionando a realização do fenômeno. Definiu-se assim uma chave de codificação com os principais fatores elencados como possíveis condicionadores do fenômeno em estudo. Inicialmente foram definidos 18 grupos de fatores, os quais foram submetidos ao teste de relevância pelo VARBRUL.

Os fatores definidos contemplam fatores de diferentes níveis de funcionamento da linguagem além de fatores extralinguísticos, listados a seguir.

A) VARIÁVEIS LINGUÍSTICAS

1. Constituição do SN sujeito
2. Presença de constituinte à esquerda
3. Tipo de verbo
4. Tipo de frase
5. Estatuto da oração (matriz ou encaixada)
6. Flexão verbal (finita ou infinita)
7. Concordância verbal
8. Presença ou ausência do objeto
9. Animacidade do SN sujeito
10. Definitude do SN sujeito
11. *Status* informacional do SN sujeito
12. Estrutura de focalização com *só* ou *mesmo*

B) VARIÁVEIS EXTRA-LINGUÍSTICAS

13. Faixa etária
14. Sexo
15. Estada fora da comunidade
16. Nível de escolaridade
17. Comunidade
18. Localidade

Como é comum em análises sociolinguísticas de orientação quantitativa, foram realizadas algumas rodadas de dados até se chegar a um conjunto de fatores estatisticamente significativos para o condicionamento do fenômeno. Os resultados, bem como os fatores, seguem descritos e interpretados no capítulo 4.

4 ANÁLISE DOS DADOS

A análise empreendida pelo pacote estatístico computacional se processa em etapas (cf. seção 1.3.2). A chegada até os resultados finais percorre algumas etapas e é permeada de algumas decisões. O caminho até a obtenção dos fatores estatisticamente relevantes a seguir apresentados e interpretados compreendeu uma série de rodadas e ajustes, uma vez que, alguns dados apresentaram comportamentos categóricos ou sobreposições nas primeiras rodadas.

Em nossa rodada inicial, com os 72 inquiridos das comunidades rurais e afro-brasileiras, foram computadas 3.602 ocorrências. As ocorrências foram rodadas separadamente, conforme os dois grupos de *corpora*: português popular (Santo Antônio de Jesus e Poções) e português afro-brasileiro (Cinzento e Sapé)³⁷. Cumpre salientar que os resultados ora apresentados foram observados em função da variante ordem VS.

No tocante aos resultados gerais, a disposição dessa ordem nos dois *corpora* confirmou a hipótese inicial de que essas comunidades apresentariam um percentual extremamente reduzido dessa variante. Os resultados gerais encontrados para a variável dependente encontram-se no Gráfico 5:

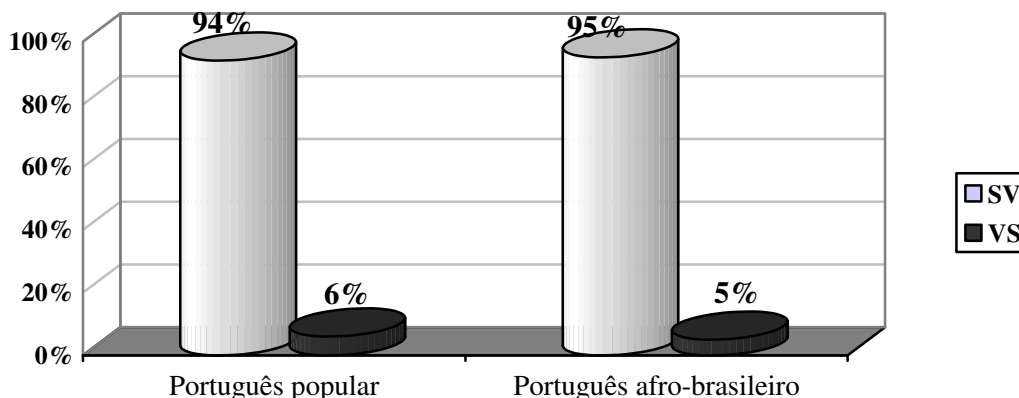


GRÁFICO 5: Resultado geral da variável dependente no português popular e no português afro-brasileiro.

³⁷ Vide delimitação e designação apresentada na seção 3.2.2.

De acordo com os resultados, não foram encontradas diferenças significativas entre os dois *corpora*. No que tange as normas populares, o português popular do interior do estado da Bahia parece ter sido afetado de forma geral pela restrição à inversão VS.

Comparando-se os resultados encontrados nesse estudo com os percentuais encontrados em outros estudos variacionistas do fenômeno, percebemos que as comunidades analisadas mostram-se correspondentes à tendência geral de do português brasileiro, preferência da ordem SV e restrição a VS (cf. Gráfico 6). Para fins de comparação, apresentamos neste gráfico apenas os resultados encontrados para as sentenças declarativas, por esse motivo, os percentuais para o português popular e afro-brasileiro são diferentes dos apresentados no Gráfico 5. Os percentuais de ordem VS, no entanto, apontam uma sensível variação entre as variedades cultas (20% e 21%) e as variedades populares (5% e 4%).

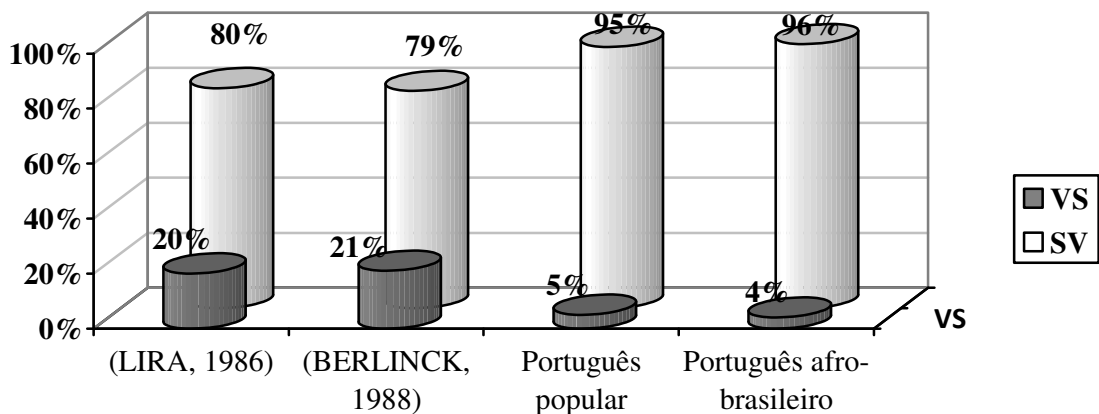


GRÁFICO 6: A ordem VS nas declarativas em variedades cultas e populares do PB.

Assim, embora os resultados apontem tendências equivalentes, as ocorrências de VS demonstram uma diferença percentual muito acentuada, o que confirma a nossa expectativa de que as variedades populares e afro-brasileiras apresentariam uma rigidez mais acurada do que as variedades cultas.

4.1 OS CONDICIONAMENTOS ESTRUTURAIS

Nesta análise, foram definidas 12 variáveis estruturais (cf. 3.5) . Essas variáveis dizem respeito, em geral, a fatores discursivos e sintáticos afetos ao sujeito, ao verbo e à oração. Os resultados encontrados serão apresentados conforme o nível em que se situam os condicionamentos mensurados. Assim, analisaremos inicialmente as variáveis ligadas ao sujeito (*constituição do SN sujeito, definitude do SN sujeito, traço de animacidade, status informacional do SN sujeito*), depois as variáveis ligadas ao verbo (*tipo de verbo, presença ou ausência do objeto, flexão finita ou infinita e concordância verbal*) e por último as variáveis relacionadas à oração (*tipo de frase, estatuto da oração, estrutura de focalização com só ou mesmo, presença de constituinte à esquerda*).

4.1.1 Condicionamentos relacionados ao sujeito

Foram definidos quatro fatores afetos ao sujeito: *constituição do SN sujeito, definitude do SN sujeito, traço de animacidade, status informacional do SN sujeito*. Desses fatores, apenas a variável *status informacional* não foi selecionada no *corpus* das comunidades rurais – Santo Antônio de Jesus e Poções.

4.1.1.1 Tipo de sujeito

Diferenciamos, com esta variável, sujeitos com as seguintes constituições: *Sujeitos pronominais (pronomes pessoais), SN com núcleo nominal preenchido, sujeitos com núcleo nominal vazio, SN modificado por quantificador, sujeitos constituídos por pronomes indefinidos, sujeitos oracionais e sujeito modificado por oração relativa*. Os valores desta variável seguem exemplificados.

a) Sujeito pronominal (pronome pessoal):

(43) *Promete eles assim, né?* (SAR-05)

(44) *nós vinha e desteiava Poções* (POR-01)

b) SN com núcleo nominal preenchido:

(45) *aí passô **um caminhão** assim* (POR-04)

(46) *quando os **homi** passasse a sê sabido* (POR-09)

c) SN com núcleo vazio:

(47) *Tinha sete, morreu **uma**, né, com onze ano...* (POR-04)

(48) ***Aquela** 'tá dirigino.* (SAR 01)

d) SN com quantificador todo(s):

(49) *Saía daí, vinha **todo mundo**, encontrava nessa igreja aí.* (POR-01)

(50) ***todo mundo** tinha presépi.* (POR-12)

e) Tudo como núcleo do SN:

(51) *hoje, totalmente, é **tudo** diferente.* (POR-03)

(52) ***Tudo** era na mesa* (PO-05)

f) Pronome indefinido

(53) *num passô **ninguém*** (POR-07)

(54) ***ninguém** vai passa* (POR-07)

g) Sujeito oracional

(55) *É difícil **ele ir mais a gente**.* (SP-01)

(56) ***quem qué dêxa uns dois dia ô três lá... e quem não qué tira hoje*** (SP-12)

h) SN com relativa

(57) *Mas a pessoa **que leva** é uma pessoa tão bondoso!* (SAR-10)

(58) *e vem um material **que compra pa a terra*** (SAR-05)

Os pronomes pessoais têm se mostrado o contexto mais refratário para à inversão VS (COELHO, 2000; BERLINCK 1989; CARVALHO, 2006). Esse resultado pode estar relacionado ao fato de os pronomes pessoais estarem, por seu caráter anafórico, vinculados a

uma menção anterior, carregando, portanto uma referência [+ definida], o que constitui fator desfavorecedor para a colocação pós-verbal (conforme veremos a seguir).

Quanto aos contextos favorecedores, destacam-se os pronomes indefinidos e os sujeitos quantificados. O condicionamento exercido pelos pronomes indefinidos, segundo Coelho (2000), decorre do traço de indefinidade que apresentam, o que favoreceria a inversão VS. Os sujeitos quantificados, conforme resultados encontrados por Carvalho (2006) também favorecem amplamente a posposição do sujeito.

Os resultados encontrados nos dois *corpora* relativos a essa variável são apresentados nas Tabelas 5 e 6 abaixo:

TABELA 5: A ordem VS no português afro-brasileiro, segundo a variável *constituição do SN sujeito*.

	<i>Oco./ total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Oracional	2/6	33%	0,98
SN com núcleo nominal preenchido	53/318	17%	0,80
<i>Tudo</i> como núcleo do SN	1/23	4%	0,57
Modificado por quantificador	1/2	50%	0,55
Sujeito pronominal	4/829	0%	0,38
SN com núcleo vazio	1/37	3%	0,24
TOTAL	63/1226	5%	

Nível de significância 0,000

De acordo com os resultados encontrados para o português afro-brasileiro, o contexto que mais favorece a ordenação VS foi o sujeito oracional, com peso relativo (P.R.) 0,98. Esse resultado pode estar relacionado com o fato de que esse tipo de sujeito em geral é uma estrutura pesada, o que, segundo Coelho (2000), favorece a posposição do sujeito.

A presença do quantificador mostrou-se, como esperado, um fator favorecedor da colocação pré-verbal, sendo este condicionamento mais evidente quando este é o núcleo do SN sujeito (P.R. 0,57).

Como contextos refratários à inversão VS, o programa apontou os fatores *SN com núcleo vazio* (P.R. 0,24) e *sujeito pronominal* (P.R. 0,38).

Os resultados encontrados para o português popular apresentaram algumas diferenças, conforme podemos observar na Tabela 6.

TABELA 6: A ordem VS no português popular, segundo a variável *constituição do SN sujeito*.

	<i>Oco./ total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Modificado por quantificador	11/24	46%	0,93
<i>Tudo</i> como núcleo	7/30	23%	0,88
SN com núcleo nominal preenchido	95/638	15%	0,87
SN com núcleo vazio	10/41	24%	0,87
Pronome indefinido	2/28	7%	0,86
Modificado por relativa	1/7	14%	0,62
Sujeito pronominal	10/1626	1%	0,29
TOTAL	136/2394	6%	

Nível de significância 0,018

Os resultados da mesma variável no português popular mostraram-se mais condizentes com as hipóteses propostas. O quantificador constituiu um fator extremamente importante no favorecimento da ordem VS, com peso relativo 0,93, como modificador e 0,88, como núcleo do SN.

O SN com núcleo vazio, ao contrário do resultado encontrado no português afro-brasileiro, foi apontado como um fator relevante para a posposição do sujeito, com peso relativo 0,87.

Da comparação dos dois resultados, podemos compreender que há fatores que atuam regularmente como condicionamentos de VS, como acontece com os pronomes, que, em todas as comunidades aqui estudadas, mostrou-se refratário à inversão.

As poucas ocorrências de inversão com pronome pessoal são casos com os verbos *discendi* (59), (60), (61); estruturas com interpretação de lista (62) ou são casos de focalização com o *só* (63).

(59) *Chega ela de gaiata.* (SAR-01)

(60) *um dia, **chegô ela**: "Vem cá Edézio"...* (SAR-01)

(61) *Promete eles assim, né?* (SAR-05)

(62) *Trabalha ela e ele, todos dois trabalha.* (SAR-04)

(63) *É, só é ruim ele no inverno.* (SAR-04)

Os exemplos, de um modo geral, constituem construções em que o sujeito encontra-se focalizado, podendo ser ainda marcado pelo focalizador *só*.

Alguns fatores, no entanto, mostraram comportamentos diferentes e até contrários em cada um dos *corpora*, conforme sintetizamos no Quadro 4, a seguir:

	<i>Fatores favorecedores</i>	<i>Fatores desfavorecedores</i>
Português afro-brasileiro	<ul style="list-style-type: none"> • Sujeito oracional³⁸ • SN com núcleo nominal preenchido • Tudo como núcleo do SN • SN modificado por quantificador 	<ul style="list-style-type: none"> • Sujeito pronominal • SN com núcleo vazio
Português popular	<ul style="list-style-type: none"> • SN modificado por quantificador • SN com núcleo vazio • Tudo como núcleo do SN • SN com núcleo nominal preenchido • Pronome indefinido • Sn modificado por relativa³⁹ 	<ul style="list-style-type: none"> • Sujeito pronominal

QUADRO 4: A relação dos fatores condicionadores com a ordem VS no português afro-brasileiro e no português popular.

A variável, portanto não apresentou uma distribuição muito clara. Destacando-se apenas a restrição do pronome pessoal à colocação pós-verbal. Essa restrição, sobretudo no português afro-brasileiro, pode estar associada à falta de flexão de caso nos pronomes pessoais, com um emprego muito amplo das formas do chamado caso reto como complementos verbais⁴⁰ (e.g., *Maria viu ele*; *Maria encontrou nós*). Nesse contexto, a inversão VS com pronomes pessoais tende a gerar frases ambíguas.

Ao contrário do português europeu, no qual ainda se conservam as formas acusativas dos pronomes, no português popular, o uso indistinto do pronome reto como nominativo e acusativo deixa a posição como único elemento identificador da função sintática. Uma frase do tipo *viu ele*, em geral, carrega a compreensão de que um sujeito oculto viu um objeto *ele*. Dificilmente os falantes do português popular e do PB, de modo mais genérico, irão interpretar o pronome *ele*, numa posição pós-verbal, como sujeito da sentença.

³⁸ A ocorrência desse tipo de fator nas comunidades rurais Santo Antônio de Jesus e Poções, apresentou apenas um dado: *Não, comigo aconteceu de eu sai assim*. (SAS-01).

³⁹ Este fator, nas comunidades afro-brasileiras Cinzento e Sapé, apresentou apenas uma ocorrência na ordem VS.

⁴⁰ Ver Mendes (2009).

4.1.1.2 Animacidade do SN sujeito

O traço [-animado] do sujeito tem sido apontado como importante condicionador da inversão VS por alguns trabalhos (ALBERTON 2001, BERLINCK 1989, COELHO 2000, dentre outros). Sustentando a hipótese perseguida e comprovada pelas autoras, definimos as seguintes variáveis:

a) [+ animado]: seres vivos com mobilidade voluntária.

(64) *Minha mãe assistia novela* (SAR 04)

(65) *Promete eles assim, né?* (SAR 05)

b) [- animado]: seres não vivos ou desprovidos de mobilidade voluntária.

(66) *que o pé de laranja, ele só vai esmoreceno* (SAR 04)

(67) *Aí vai tudo ficano mais caro* (SAR 05)

Os resultados encontrados nos dois *corpora* (cf. Gráfico 7) confirmam a hipótese consagrada em outros trabalhos.

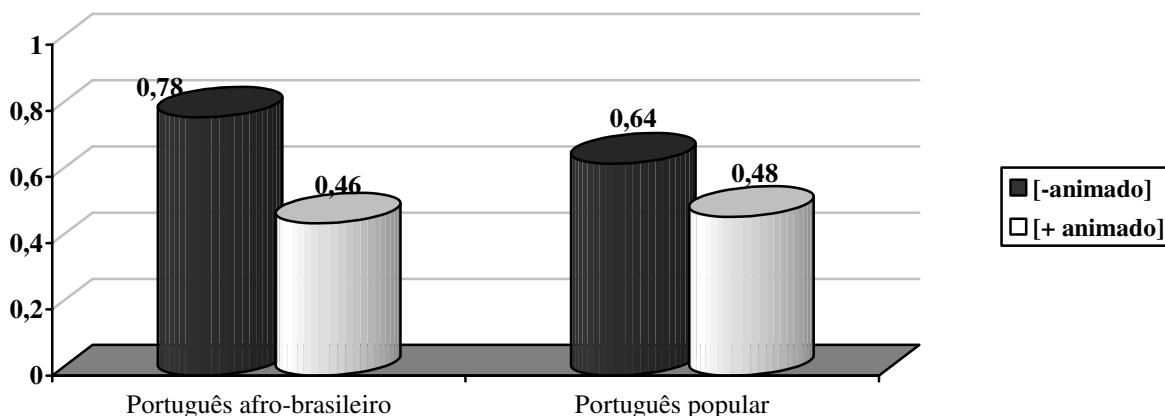


GRÁFICO 7: A ordem VS no português afro-brasileiro e no português popular, segundo a variável *animacidade do SN sujeito*.

Assim, o comportamento da variável vai ao encontro da hipótese inicial nas duas variedades do português do interior da Bahia. Com P.R. 0,78 no português afro-brasileiro e 0,64 para o português popular, o traço [-animado] do SN sujeito apresentou-se como fator favorecedor da ordem inversa VS. Vale destacar que o contraste relativo à animacidade ficou mais evidente nas comunidades afro-brasileiras.

O comportamento aqui encontrado pode ser explicado pelo fato de que os sujeitos com o traço [+animado] são, via de regra, agentes da ação verbal, uma vez que a ação implica movimento voluntário. Essa agentividade seria um aspecto semântico próprio da posição de sujeito, o que enrijeceria a ordem SV. Ao contrário, os sujeitos não agentivos, em geral, de verbos inacusativos, tendem a apresentar o traço [-animado], com o papel temático de experienciadores ou mesmo pacientes da ação verbal, o que os identifica aos complementos verbais, favorecendo a inversão VS.

Esta relação entre o traço de animacidade e natureza do verbo predicador foi apontada por Berlinck (1989, p. 233), segundo a qual “a animacidade do SN é um dos traços especificados pela grade temática do verbo. Portanto, sua relevância não é própria, mas derivada dessa relação. (BERLINCK, 1989, p. 233)

A formulação desse fator, portanto encontra-se ancorada em outras instâncias, o que pode nos indicar que se trata de um epifenômeno. Em outras palavras, não seria exatamente o traço semântico [-animado] que condicionaria a ordem VS, mas sim o fato de esse traço ser característico de sujeitos normalmente subcategorizados por determinado tipo de verbos, estes, sim, favorecedores da ordem VS (como veremos nos resultados da variável tipo de verbo).

4.1.1.3 Traço de definitude do SN sujeito

Para esta variável, foram definidos três valores: **definido**, quando o referente é identificável no discurso e representa uma entidade específica; **indefinido**, o referente é um grupo ou uma entidade específica, não referido anteriormente, ou não identificável pelo discurso do falante; **genérico**, quando o referente é uma entidade inespecífica, de forma que o referente do sujeito pode ser identificado como qualquer indivíduo de um grupo ou espécie.

a) Definido:

(68) *cresceu muito Poções* (POS-05)

(69) *meu pai tem um problema de amnésia* (POS-06)

b) Indefinido:

(70) *ôtos já casaro* (POR-04)

(71) *juntô um bocado de home e cercô* (POS-06)

c) Genérico

(72) *Ficava home e mulher* (POS-07)

(73) *você podia 'manhecê o dia na rua aí com o bolso cheio de dinheiro* (POS-06)

Com esta variável, buscamos investigar a pertinência do traço de definitude no condicionamento da ordem. Como hipótese de trabalho, esperamos que sujeitos *indefinidos* favoreçam a ordem VS.

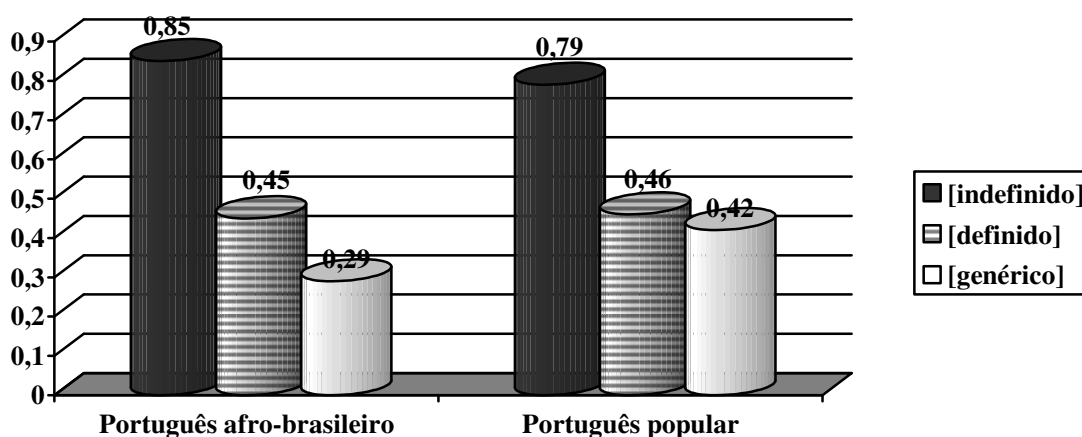


GRÁFICO 8: A ordem VS no português popular e no português afro-brasileiro, segundo a variável *definitude do SN sujeito*.

Os resultados encontrados confirmaram a hipótese inicial. Os sujeitos *indefinidos* favorecem a variante VS, com pesos relativos 0,85 e 0,79 para o português afro-brasileiro e popular respectivamente. Os sujeitos com referência *definida* foram apontados pelo VARBRUL como desfavorecedores da ordem VS nas duas variedades do PB aqui analisadas, com pesos relativos 0,45 e 0,46. O contexto mais refratário à inversão foram os SN *genéricos*, como pesos 0,29 e 0,42.

A relação entre a referência indefinida e a inversão VS pode ser atribuída a uma possível focalização do sujeito nesses casos, por se tratar de uma primeira menção. No desenvolvimento do discurso, os elementos com o *status* informacional *novo* tendem a ser realçados por mecanismos linguísticos, como a clivagem. Dessa forma, a inversão VS pode ser interpretada como um mecanismo de realce. Essa relação entre o estatuto informacional do sujeito e a ordem dos constituintes na oração será mais bem explicitada na próxima variável.

4.1.1.4 Status informacional do SN sujeito

A formulação deste fator objetivou explicitar a relação entre o *status* informacional do SN sujeito e a posição deste SN na sentença. Esta variável diz respeito, portanto, à estrutura informacional do discurso. Para o estudo, foram definidos dois valores *dado*, quando o sujeito é uma informação referida anteriormente, ou seja, é uma informação textualmente pressuposta, e *novo*, para os casos em que o sujeito é uma informação ainda não mencionada, inserida pela primeira vez no discurso.

Os valores definidos para esta variável encontram-se exemplificados abaixo:

a) Dado: Faz parte do conhecimento partilhado por ouvinte e falante e já foi citado no discurso anteriormente.

(74) *um dia, minha professora num pôde ir, aí te... teve uma **ôta professora** que foi ensiná, chamava até dona Tonha, né? Aí ela me deu umas reuguada...* (POR-02)

(75) *A **mãe** todo dia dá uma piada a você, você entra com sua esposa no lugá, chega ela de gaiata...* (SAR 01)

b) Novo: Não faz parte do conhecimento partilhado por ouvinte e falante e ainda não foi citado no discurso.

(76) *Não, eu... poque às vez eu fico aí em casa. Aí, o povo sai por' aqui tomano licô e tudo, aí eu fico em casa.* (SAR 11)

(77) *com' é que vai sai(r) **uma mocinha de doze ano**, treze ano, catoze ano, pedino carona na estrada, né?" Passa um caminhonêro aí, que [ninguém] sabe nem d' onde que vem* (POR-07)

Espera-se com esta variável que os sujeitos informacionalmente novos favoreçam a ordem VS, tendo em vista a relação que esta posição tem com função semântico-discursiva de foco.

A variável foi selecionada como estatisticamente relevante apenas no português afro-brasileiro, sendo descartada no português popular, no entanto seus resultados também serão apresentados tendo em vista a clareza dos resultados de frequência absoluta.

TABELA 7: A ordem VS no português afro-brasileiro, segundo a variável *status informacional do SN sujeito*.

	<i>Oco./ total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Dado	20/1098	2%	0,45
Novo	43/128	34%	0,86
TOTAL	63/1226	5%	

Nível de significância 0,000

Embora não tenha sido selecionada no português popular, os resultados de frequência apontam um comportamento semelhante, conforme Tabela 8.

TABELA 8: A ordem VS no português popular, segundo a variável *status informacional do SN sujeito*.

	<i>Oco./ total</i>	<i>Freq.</i>
Dado	75/2085	4%
Novo	61/310	20%
TOTAL	136/2395	6%

Os resultados de frequência demonstram que os SNs com *status novo* apresentam um percentual 14% superior à frequência geral de VS no *corpus*, o que indica que esse fator favorece a variante. O *status dado*, por sua vez, está abaixo da frequência geral, com 4% de ocorrência, isto indica que o fator desfavorece a ordem VS.

Conforme os resultados, SNs com estatuto informacional *novo* tendem a favorecer a posposição do sujeito. Este resultado nos remete a uma análise retomada da noção de foco discursivo e sua relação com a posição na sentença.

De acordo com Lambrecht (1994), as noções foco e tópico consistem em noções sintáticas com função pragmática correlata às noções cognitivas *pressuposição/asserção* e às noções semânticas *informação velha/informação nova*. Ainda segundo o autor, o tópico de uma sentença é o constituinte da proposição sobre o qual se fala, ou seja, sobre o que é a sentença. Essa explicação é definida pelo autor como *aboutness*, ou seja, o caráter do tópico de “ser sobre”. O foco, por sua vez é o elemento acrescentado a uma pressuposição, sendo, portanto parte da asserção. A função de foco pode ser desempenhada por qualquer sintagma ou até mesmo por uma oração inteira. O foco diferencia-se do tópico tanto em sua função pragmática quanto em sua disposição formal. Ao contrário do tópico, o elemento focalizado está relacionado ao predicado, estabelecendo uma ligação entre este e o indivíduo, o que caracteriza seu papel na predicação. Dessa forma, podemos dizer que o foco é a função

pragmática relacionada ao estatuto informacional *novo* dos constituintes de uma sentença, mantendo, dentro da estrutura da informação, íntima relação com os demais elementos.

Ainda segundo o autor, a posição canônica do foco é à direita da oração, na última posição. Isto porque nessa posição o esforço entoacional é maior, ou seja, possui proeminência prosódica. A ocorrência do foco nesta posição é construção mais natural, sendo, portanto, a opção não marcada.

Portanto, conforme sintetizam Mateus et al. (2003), a função do foco está correlata à introdução de um elemento novo no discurso, dando-lhe ênfase.

Estas constatações justificam os resultados encontrados aqui. Quando correspondentes a uma informação *nova*, os sujeitos tendem a ocupar a posição mais à direita, posição sobre a qual, em língua portuguesa, recai o acento proeminente de foco, identificável com a ênfase conferida ao elemento novo do ponto de vista pragmático- discursivo.

4.1.2 Condicionamentos relacionados ao verbo

Do conjunto inicial de quatro variáveis relacionadas ao verbo (*Tipo de verbo, presença ou ausência do objeto, flexão finita ou infinita e concordância verbal*), apenas uma foi apontada como relevante no condicionamento da ordem VS: *Tipo de verbo*.

4.1.2.1 Tipo de verbo

A natureza do verbo tem sido apontada como principal condicionamento para a inversão VS. Diferentes estudos, sob diferentes orientações teóricas, têm destacado a pertinência deste fator, estruturado nesta análise com base nos seguintes valores:

a) Transitivo:

(78) *Eh, a crasse báxa num **tem** moral!* (SAS-02)

(79) ***Tem** celulá o pessoal.* (SAR-04)

b) Intransitivo:

(80) *Eu **costurava** pra mim* (CZ-05)

(81) *Convesa os dois pra vê se chega em algum lugá* (SAS-05)

c) Inacusativo:

(82) Agora mesmo, há dois anos atrás, *morreu meu sogro*. (POS-06)

(83) *A mãe de Ana Maria morreu* (POR-12)

d) Ligação

(84) *é raso o rio...* (SP-03)

(85) *Jequié é quente!* (CZ-10)

e) Movimento/ Locativo

(86) *Ele foi pá São Paulo* (CZ-06)

(87) *e bem ali a diante vinha um rapaz de lá pra cá...* (SP-08)

No âmbito dos estudos gerativos, a natureza do predicador é apontada como principal motivação para a inversão VS. Partindo da distinção entre verbos monoargumentais intransitivos e inacusativos, destacam-se as suas diferentes configurações estruturais como condicionantes de comportamentos diferentes no que tange a inversão (cf. seção 2.1.2).

Os verbos transitivos, por sua vez, devido à restrição de monoargumentalidade (KATO; TARALLO, 1988; KATO, 2000), constituiriam o contexto mais refratário à inversão.

Os resultados foram semelhantes para todas as comunidades, como se pode ver nas tabelas abaixo:

TABELA 9: Ordem VS no português afro-brasileiro, segundo a variável *tipo de verbo*.

	<i>Oco./total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Inacusativo	23/98	23%	0,89
Movimento/ Locativo	12/212	6%	0,72
Intransitivo	0/75	0%	---
Ligação	19/177	11%	0,58
Transitivo	8/658	1%	0,33
TOTAL	62/1220	5%	

Nível de significância 0,000

TABELA 10: Ordem VS no português popular, segundo a variável *tipo de verbo*.

	<i>Oco./total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Inacusativo	62/172	36%	0,94
Movimento/ Locativo	25/378	7%	0,84
Intransitivo	10/178	6%	0,72
Ligação	18/292	6%	0,64
Transitivo	6/1271	0%	0,24
TOTAL	121/2291	5%	

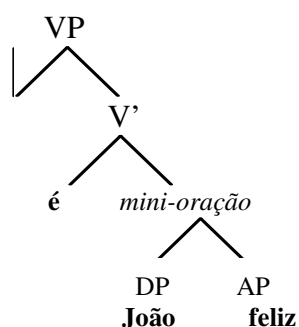
Nível de significância 0,018

De modo geral, a ordem VS mostrou-se desfavorecida pelos verbos transitivos. No português afro-brasileiro apresentaram peso relativo 0,33, e no português popular 0,24. O verbo intransitivo, entretanto, apresentou comportamento contraditório, sendo favorecedor da ordem VS no português popular (P.R. 0,72) e categoricamente refratário à inversão (0% de 75 ocorrências) no português afro-brasileiro.

Como contexto favorecedor da ordem VS, como já esperado, os verbos inacusativos foram os principais condicionadores. Com este tipo de verbo, os resultados encontrados foram 0,89 (afro) e 0,94 (popular).

Os verbos de ligação também apresentaram resultados acima de 0,5, o que estatisticamente indica favorecimento à inversão VS. Os pesos 0,58 (afro) e 0,64 (popular) podem estar associados ao fato de esses verbos, de acordo com a teoria gerativa, serem construções nas quais se manifesta o fenômeno da inacusatividade. Os verbos copulativos selecionariam apenas um argumento interno, uma mini-oração, constituída por um predicador (o predicativo do sujeito) que seleciona semanticamente o sujeito sintático da sentença, conforme estrutura representada a seguir.

(88) João é feliz.



Os verbos de movimento e locativo também favoreceram à ordem VS, com pesos 0,73 (afro) e 0,84 (popular). Assim como os verbos de ligação, os verbos de movimento e locativos, segundo a teoria gramatical, também apresentam a mesma estrutura dos verbos inacusativos, e, portanto, apresentam a mesma justificativa.

4.1.3 Condicionamentos no nível da oração

Os condicionamentos ligados à oração distribuem-se em dois tipos: o primeiro relaciona-se à natureza da estrutura frasal em que o sujeito se encontra, o segundo diz respeito aos constituintes que figuram na oração. No primeiro caso, temos as variáveis *tipo de oração* e *tipo de frase*. No segundo caso, temos as variáveis *presença de constituinte à esquerda* e *estrutura de focalização com só ou mesmo*. Dessas variáveis, apenas duas foram selecionadas:

- (i) *Tipo de frase*
- (ii) *Presença de constituinte à esquerda*

4.1.3.1 Tipo de frase

Para esta variável, distinguiram-se inicialmente quatro fatores:

a) afirmativa

(89) *O da Ilha eu dava dinheiro* (SAR 01)

b) negativa

(90) *you não pode fazer muito filho mode disso* (SAR 01)

c) interrogativa direta

(91) *quando eu morrer, eu vô levá o quê?* (SAR 01)

d) Interrogativa indireta

(92) *aí pergunta à pessoa qual é o nome de Deus* (CZ-01)

De acordo com Duarte (1992), em referência ao estudo de Âmbar (1987), a inversão VS é obrigatória em interrogativas diretas QU- no português europeu. Assim como no PE, no espanhol figura uma regra sintática em que a inversão é obrigatória nos contextos interrogativos. No português brasileiro, a inversão apresenta um caráter mais restrito, o qual, segundo Duarte (1992)⁴¹, teria passado por um processo de mudança de um padrão com ordem VS para as interrogativas diretas no século XVIII⁴², para uma ordenação SV no século XX.

Dessa forma, acreditamos que as interrogativas favorecem a ordem VS, enquanto será desfavorecida nas sentenças declarativas (negativas e afirmativas).

TABELA 11: A ordem VS no português afro-brasileiro, segundo a variável *tipo de frase*.

	<i>Oco./ total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Interrogativa indireta	7/11	64%	0,99
Interrogativa direta	7/34	21%	0,91
Afirmativa	48/1034	5%	0,52
Negativa	1/147	1%	0,18
TOTAL	63/1226	5%	

Nível de significância 0,000

TABELA 12: A ordem VS no português popular, segundo a variável *tipo de frase*.

	<i>Oco./ total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Interrogativa direta	10/81	12%	0,89
Interrogativa indireta	4/20	20%	0,85
Afirmativa	116/2026	6%	0,49
Negativa	6/268	2%	0,36
TOTAL	136/2395	6%	

Nível de significância 0,018

A hipótese inicial foi confirmada pelos resultados. A variável parece apresentar um comportamento contrastivo, opondo-se as sentenças declarativas, desfavorecedoras da ordem

⁴¹ O estudo da autora baseia-se na análise empírica de peças teatrais escritas nos séculos XVIII, XIX e XX.

⁴² Vale destacar que os dados referentes ao século XVIII dizem respeito à peça *O Judeu*, escrita em 1734 por Antônio José, brasileiro de nascimento, mas residente em Portugal.

VS às sentenças interrogativas, favorecedoras de VS. Os dados das Tabelas 11 e 12 confirmam esse comportamento.

4.1.3.2 Presença de constituinte à esquerda

Selecionada apenas no *corpus* do português popular, esta variável baseia-se na tendência de o PB preencher a primeira posição do sintagma (cf. DUARTE, 1993). Levantamos a hipótese de que o preenchimento da primeira posição com um sintagma adverbial favoreceria a colocação posposta do sujeito. Foram definidos os seguintes valores para esta variável:

a) sem constituinte

(93) *Eu já 'tava sonhando com a menina* (SAR 01)

b) tópico

(94) *A pêra, ela é sempre mais coisa* (SAR 02)

(95) *o pé de laranja, ele só vai esmoreceno* (SAR-04)

c) Adjunto adverbial

(96) *Perto da cidade eu tomo* (SAR 01)

(97) *Cada dia passa um diferente* (SAR 04)

d) Marcador conversacional

(98) *Agora, o difícil é a ladêra*

Os resultados encontrados confirmaram a hipótese predefinida.

TABELA 13: A ordem VS no português popular, segundo a variável *presença de constituinte à esquerda*.

	<i>Oco./ total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Adjunto adverbial	22/215	10%	0,71
Marcador conversacional	9/134	7%	0,62
Sem constituinte	105/2004	5%	0,47
Tópico	0/42	0%	---
TOTAL	136/2395	6%	

Nível de significância 0,018

A presença do constituinte favoreceu a ordem VS com P.R. 0,71. Podemos concluir que a presença de constituinte incide na posição do sujeito e do verbo na sentença. Conforme discutido na seção 2.3.1, o preenchimento da primeira posição por um elemento diferente do verbo influencia na colocação desse na segunda posição, ficando o sujeito, portanto, na terceira posição. O resultado vai ao encontro das pesquisas discutidas em 2.3.1 (cf., por exemplo, SPANO, 2008). A função de *tópico* desfavoreceu categoricamente a posposição do sujeito. Vale destacar que o número de ocorrências foi muito baixo (42 ocorrências).

4.2 OS CONDICIONAMENTOS EXTRALINGUISTICOS

Conforme Lucchesi (2009), em análise a estudos dedicados ao português popular, para algumas variáveis, destacadamente as variáveis não marcadas por um julgamento social explícito, o efeito das variáveis sociais não tem se mostrado muito nítido. É o caso, por exemplo, de variáveis como a *realização do sujeito pronominal, construções relativas e processos de topicalização e clivagem*, e parece ser o caso da ordem dos constituintes Verbo e Sujeito.

Os dados aqui analisados parecem confirmar a hipótese de Lucchesi. Dentre os fatores sociais controlados pela análise, apenas a variável escolaridade foi selecionada nas comunidades rurais. As demais variáveis foram descartadas na seleção dos fatores condicionantes e não apresentaram resultados de frequência significativos.

4.2.1 A variável escolaridade

A relevância dessa variável para o estudo da variação e da mudança já foi oportunamente discutida na seção 1.3.1. Em síntese, a proeminência concedida ao fator está na medida direta de sua influência no contato e aquisição das variantes prestigiadas, ao passo que atua como fator inibidor das variantes estigmatizadas. Sendo assim, espera-se que seu efeito seja mais expressivo em variáveis socialmente sensíveis, cuja realização esteja ao alcance do nível de consciência do falante.

Em nosso estudo, levando em consideração a realidade social-cultural dos falantes estudados (em geral informantes com baixo acesso à escolarização), definimos como valores dessa variável as variantes:

- a) analfabetos: indivíduos sem contato com o letramento;
- b) semi-analfabetos: indivíduos com pouco contato com o letramento.

Em geral, os informantes classificados como semi-analfabetos, apesar de terem frequentado a escola, sabem apenas assinar o próprio nome e não apresentam muita fluência em leitura.

Acreditamos que as diferenças estabelecidas nesse *corpus* no tocante à variável escolaridade são muito sutis para que uma diferença sensível seja encontrada. Isso porque a ausência de contato com o letramento e o contato incipiente podem constituir níveis pouco distintos de escolarização. Não obstante, nessas localidades a qualidade de ensino é bastante precária, sobretudo nas comunidades afro-brasileiras, seja pela carência de qualificação e de políticas educacionais específicas, seja pela ausência de infra-estrutura.

Ademais, é necessário destacar que o fenômeno em estudo nesta dissertação parece apresentar pouca ou nenhuma carga de valoração social, sendo, portanto, menos sensível ao processo de escolarização. Apesar disso, levantamos a hipótese de que a variante VS está sendo adquirida por via indireta, através da aquisição de outras regras, como a flexão de caso dos pronomes pessoais, que viriam a favorecer a inversão VS. Só assim podemos pensar o efeito da escolarização sobre a ordem dos constituintes na oração.

TABELA 14: A ordem VS no português popular, segundo a variável *escolaridade*.

	<i>Oco./ total</i>	<i>Freq.</i>	<i>Peso relativo</i>
Semi-analfabeto	90/1436	6%	0,56
Analfabeto	46/959	5%	0,42
TOTAL	136/2395	6%	

Nível de significância 0,018

Embora os resultados indiquem o contato com letramento como favorecedor de VS, P.R. 0,56, é preciso destacar que os pesos relativos e, principalmente, os resultados de frequência encontram-se muito próximos, o que inibe um diagnóstico mais conclusivo.

Nas comunidades afro-brasileiras a variável não foi selecionada, e apresentou resultados percentuais idênticos para os dois fatores (5%).

Os resultados de modo geral demonstram certa coerência e conformidade com a visão que assumimos aqui sobre a escolarização. O fato de apresentar relevância apenas nas comunidades de fala dos municípios de Santo Antônio de Jesus e Poções decorre do fato de nesses centros existir um padrão de escolarização com mais consistência do que o (não) encontrado nas comunidades afro-brasileiras, em função do isolamento e do baixo desenvolvimento social destas.

4.2.2 As demais variáveis sociais

Embora não tenham sido selecionadas, apresentamos os resultados das demais variáveis sociais, haja vista a importância que apresentam para o estudo sociolinguístico.

4.2.2.1 A variável faixa etária

O estabelecimento dessa variável orientou-se pela hipótese de que a variante VS seria inovadora para a comunidade, sendo introduzida pela escolarização e pelo contato dialetal. Como os mais jovens, nessas comunidades, apresentam maior escolarização, esperávamos que a variável fosse favorecida por essa faixa. Como vemos nos resultados do Gráfico 9.

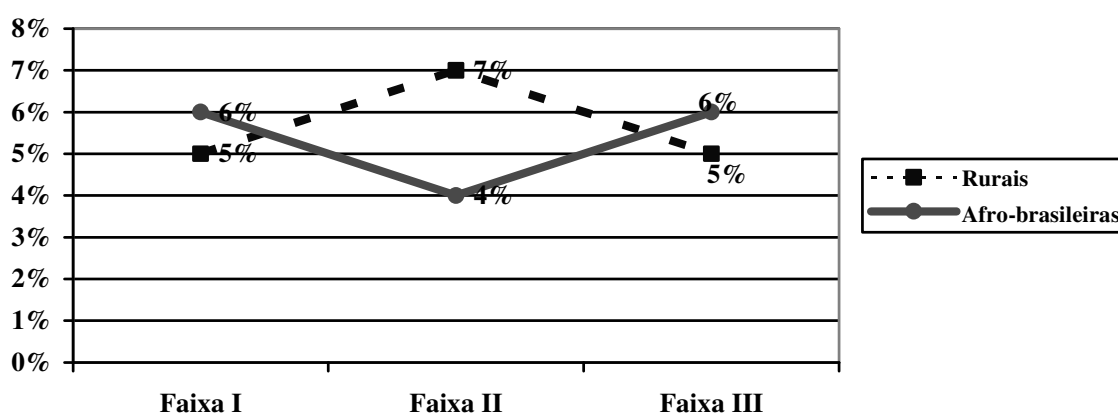


GRÁFICO 9: A ordem VS no português popular e no português afro-brasileiro, segundo a variável *faixa etária*.

Os resultados indicam uma situação de variação estável, com orientações opostas para os dois tipos de comunidades. Os resultados observados, no entanto, não nos permitem grandes explicações, uma vez que são considerações feitas a partir de frequências absolutas.

4.2.2.2 A variável *estada fora da comunidade*

Com esta variável, buscamos compreender em que medida o contato com outros dialetos interferiria na realização da ordem. Levantamos a hipótese de que a ordem VS, incomum à norma dessas comunidades estaria sendo influenciada pelo contato dos falantes com outras normas.

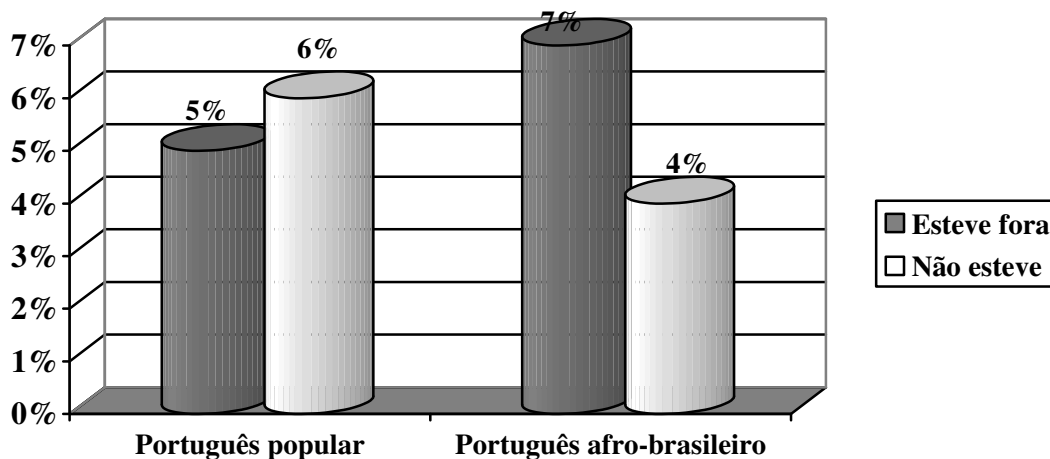


GRÁFICO 10: A ordem VS no português popular e no português afro-brasileiro, segundo a variável *estada fora da comunidade*.

Os resultados da variável no português popular apresentaram uma diferenciação insignificante (1%). No português afro-brasileiro, por sua vez, a diferença entre os percentuais indicou a favorecimento da variante *estada fora da comunidade* no condicionamento de VS, com frequência 7%.

4.2.2.3 A variável sexo.

Esta variável mostrou-se irrelevante para o condicionamento do fenômeno. Nas comunidades afro-brasileiras, apresentou percentuais idênticos para homem e mulher (5%), nas comunidades rurais, a diferença foi de 1%, com os homens com 6% e as mulheres com 5%. Assim, podemos concluir que esta variável não exerce influência sobre o fenômeno.

CONSIDERAÇÕES FINAIS

Desde a sociolinguística, constitui lugar comum, a assunção da heterogeneidade linguística. Como entidade a serviço de uma comunidade complexa, como é toda organização social humana, a língua tem em sua heterogeneidade ordenada o fundamento de sua funcionalidade. Esse objeto se entrecruza com o processo que a envolve, de modo a constituir, ao mesmo tempo, produto e processo cultural. Em outras palavras, na mesma medida em que a língua é resultado do processo histórico circundante, é também um elemento que interfere nessa realidade. Assim sendo, não seria diferente com a língua portuguesa, que, percorrendo oceanos, “namorou no chão e na poeira do Brasil” e constituiu o complexo português brasileiro.

A realidade sociolinguística brasileira, conforme brevemente discutido no capítulo 3, constitui uma clara polarização, com normas cultas e populares, caracterizadas e orientadas por tendências diferentes. Essa realidade é materializada em alguns fenômenos linguísticos como a morfologia flexional, amplamente reduzida no português popular; redução esta estigmatizada pelas camadas mais elevadas e escolarizadas. Orientado por essa constatação, esta pesquisa investigou como esse quadro geral do PB se atualiza na configuração da ordem dos constituintes *Verbo* e *Sujeito* na sentença, em duas de suas variedades populares: o português afro-brasileiro, das comunidades rurais isoladas oriundas de antigos quilombos, e o português popular de dois municípios do interior do Estado da Bahia.

A análise do fenômeno ancorou-se no aparato teórico metodológico da sociolinguística variacionista, conforme apresentado no capítulo 1. O referencial teórico adotado mostrou-se adequado ao tratamento do fenômeno sintático, constituindo importante ferramenta para a compreensão do funcionamento da língua. O suporte tomado, no entanto, não desprezou as considerações da teoria gerativa para importantes aspectos do fenômeno. Pelo contrário, estas foram de extrema importância para a compreensão de alguns condicionamentos, tais como a interferência do tipo de verbo na seleção da ordem VS.

Os fatores condicionadores, em geral, foram os mesmos em todas as comunidades. No entanto registrou-se uma diferença no tocante a prioridade dos fatores. No Quadro 5, apresentamos os condicionamentos por ordem de seleção, ou seja, pela importância atribuída pelo VARBRUL como fator condicionante.

Português afro-brasileiro	Português popular
1. <i>status</i> informacional do SN sujeito	1. Tipo de verbo
2. Animacidade do SN Sujeito	2. Constituição do SN sujeito
3. Tipo de frase	3. Traço de definitude
4. Tipo de verbo	4. Tipo de frase
5. Traço de definitude	5. Animacidade do SN Sujeito
6. Constituição do SN sujeito	6. Presença de constituinte à esquerda
	7. Escolaridade

QUADRO 5: Fatores condicionantes da ordem VS no português afro-brasileiro e no português popular.

De modo geral, as duas variedades do PB se aproximam quanto aos condicionamentos. Vale destacar, no entanto, que, enquanto no português afro-brasileiro o condicionamento mais proeminente é discursivo – *status informacional do SN sujeito* –, no português popular é a natureza do predicador que em primeira estância orienta a ordenação dos constituintes *Sujeito* e *Verbo*. Portanto, pode-se dizer que o processo variável no português popular afigura-se como um fenômeno eminentemente sintático, ao passo que está relacionado ao discurso, no português afro-brasileiro.

Sobre as comunidades e a proposta de *continuum*, a restrição à ordem VS parece ter afetado o português brasileiro de modo geral, haja vista a tendência à fixação observada pelos demais estudos. No entanto, a rigidez encontrada no português popular, de modo geral, é, notadamente, mais forte do que nas variedades cultas (5% de ordem VS no português popular, contra 20% das variedades cultas). O que nos leva a considerar a influência do contato entre línguas nessas variedades, o qual, se não influenciou diretamente a ordenação dos constituintes, atuou em fatores tangenciais dessas variedades.

Nesse sentido, a restrição da ordenação do SV, atuante no PB de modo geral, e associada a outros fenômenos como a perda do sujeito nulo e a realização nula do objeto, conforme encaixamento proposto por Tarallo (1993)⁴³, tem no português popular um fator catalisador: a perda da flexão dos pronomes. Conforme resultados apontados por Mendes (2009), no português popular do interior do estado da Bahia, o uso dos pronomes flexionados está limitado à primeira e segunda pessoa do singular e, mesmo nesse contexto, as formas flexionadas convivem com formas não flexionadas. Isto explica, sobretudo, a forte restrição de posposição com pronomes pessoais nessas variedades.

⁴³ Ver seção 2.2.1.

Sentenças como *Encontrou ele o João* são, no português brasileiro, construções ambíguas, uma vez que o pronome assume a mesma forma tanto para a função de sujeito quanto para a função de objeto. Diferentemente da variedade lusitana na qual a flexão do pronome dirime qualquer confusão ao realizar-se na função de objeto, como em *Encontrou-o o João*. Essa ambiguidade se estende, no português popular, a quase todos os pronomes pessoais (e.g.: *Encontrou nós o João* ao invés de *Encontrou-nos o João*). Isso limita tremendamente qualquer inversão com sujeitos pronominais nas variedades do PB.

Quanto ao encaixamento social do fenômeno, em geral fatores extralinguísticos não foram apontados como relevantes, o que nos leva a refletir sobre a afirmação de Labov de que alguns fenômenos linguísticos se processam abaixo do nível de consciência do falante. Somos, portanto, levados a concordar com LUCCHESI (2009), quando afirma:

Mas se o cenário da mudança aquisicional é nítido nos processos de variação decorrentes da erosão gramatical provocada pelo contato entre línguas, não se encontrou nas comunidades de fala estudadas uma distribuição clara que indicasse um processo de mudança em curso na análise sociolinguística para as quais o efeito do contato não é nítido. (LUCCHESI, 2009, p. 540).

O autor destaca, a partir da análise de alguns estudos empreendidos no âmbito do Projeto Vertentes, que alguns fenômenos, tais como, a realização do sujeito pronominal, os processos de topicalização e clivagem, as construções relativas e as estruturas de negação não apresentam marcas de explícitas de julgamento social, o que justificaria os resultados irrelevantes ou contraditórios apontados pelas pesquisas (ver LUCCHESI, 2009, p. 540).

Os resultados da variável tipo de frase, por fim, indicaram uma necessidade de uma análise mais aprofundada sobre o tema, que busque investigar dentro deste grupo quais os fatores estariam atuando na preferência pela ordem VS. Numa análise exclusiva das interrogativas, que possa contemplar um número maior de ocorrências, fatores como *tipo do pronome interrogativo*, *a presença do expletivo é que*, bem como outros a serem levantados, podem ser mais bem observados.

Esta constatação alimenta a motivação para a possível continuidade deste trabalho, que embora termine nas considerações finais, certamente não se esgota aqui.

REFERÊNCIAS

ALBUQUERQUE, Wlamyra R. de; FRAGA FILHO, Walter. **Uma história do negro no Brasil**. Salvador: Centro de estudos Afro Orientais; Brasília: Fundação Cultural Palmares, 2006.

AVE-LALLEMANT, Robert. **Viagens pelas províncias da Bahia, Pernambuco, Alagoas e Sergipe (1859)**. Belo Horizonte: Itatiaia; São Paulo: Editora da Universidade de São Paulo, 1980.

ALMEIDA, Manuel; CAMPOY, Juan Manuel Hernández. Metodología para el estudio de la variación: planificación inicial. In: _____. **Metodología de la investigación sociolingüística**. Málaga: Editorial Comares, 2005. p. 37-81.

ALBERTON, Cristiane. **O português falado no Rio Grande do Sul: a ordem verbo sujeito**. Passo Fundo: UPF, 2001.

ÂMBAR, Manuela. Gouvernement et inversion dans les interrogatives QU- en portugais. 16 ed. [s.l.]: **Recherches Linguistiques**, 1987. p. 5-47.

ÂMBAR, Manuela. **Para uma sintaxe da inversão sujeito verbo em português**. Lisboa: Colibri, 1992.

ANTONINO, Vivian. **A concordância nominal em predicativos do sujeito e em estruturas passivas no português popular do interior do estado da Bahia**. 2007. Dissertação (Mestrado) - Universidade Federal da Bahia, Salvador, 2007.

BANDEIRA, M. As funções sintáticas de nós e a gente no português afro-brasileiro. In: CONGRESSO INTERNACIONAL DA ABRALIN, 6., 2009. João Pessoa. **Anais...** João Pessoa: Idéia, 2009. v. 1, p. 2701-2709.

BARBOSA, J. S. **Gramática filosófica da língua portuguesa**. 5. ed. Lisboa: Tipografia da Academia Real das Ciências, 1871.

BECHARA, Evanildo. **Moderna Gramática Portuguesa**. 37. ed. Rio de Janeiro: Lucerna, 2001.

BELLETTI, Adriana. The case of unaccusatives. **Linguistic Inquiry**, [S.l.], v. 19, n 1, p. 1-34, inverno 1988.

BERLINCK, Rosane de Andrade. **A ordem VS no português do Brasil: sincronia e diacronia**. 1988. 264 f. Dissertação (Mestrado) Universidade Estadual de Campinas, Campinas, 1988.

BERLINCK, Rosane de Andrade. A Construção VS no português do Brasil: uma visão diacrônica do fenômeno da Ordem. In: TARALLO, Fernando (Org.). **Fotografias sociolingüísticas**. Campinas: Pontes, 1989. p. 95-112.

BERLINCK, Rosane de Andrade. **La position du sujet en portugais**: etude diachronique des variétés brésilienne et européenne. 1995. Thèse (Doctorat) Katholieke Universiteit Leuven, 1995.

BERLINK, Rosane Andrade et al. Sintaxe. In: MUSSALIM, Fernanda; BENTES, Ana Cristina (Org.). **Introdução à lingüística**: domínios e fronteiras. 2. ed. São Paulo: Cortez, 2001. p. 207-244.

BORTONI-RICARDO, Stella Maris. **Educação em língua materna**: a sociolingüística na sala de aula. São Paulo: Parábola, 2006.

BRITO, Ana Maria; DUARTE, Inês; MATOS, Gabriele. Estrutura da frase simples e tipos de frase. In: MATEUS, M. H. M. et al. **Gramática da língua portuguesa**. 5. ed. Lisboa: Caminho editorial, 2003. p. 433-506.

BURZIO, Luigi. **Italian Syntax**: a government-binding Approach. Dordrecht: Reisel Publishing Company, 1986.

BRIGHT, William. As dimensões da sociolingüística. In: FONSECA, Maria Stela V.; NEVES, Moema F. (Org.). **Sociolingüística**. Tradução: Livraria Eldorado Tijuca Ltda. Rio de Janeiro: Eldorado, (1974 [1966])

CÂMARA JÚNIOR, Joaquim Mattoso. **História e estrutura da língua portuguesa**. 2. ed. Rio de Janeiro: Padrão, 1979.

CARVALHO, Guilhermina Maria Bastos Mendes de. **A inacusatividade na fala de comunidades afro-brasileiras isoladas**. 2006. 167 f. Dissertação (Mestrado) - Universidade Federal da Bahia, 2006.

CEDERGREN, H; SANKOFF, David. Variable rules: performance of a statistical reflection of competence. **Language**, n. 50, p. 33-55, 1974.

CHAVES, Arlete Saddi. A ordem VS no português da fronteira. In: TARALLO, Fernando (Org.). **Fotografias sociolingüísticas**. Campinas: Pontes, 1989. p. 65-79.

COELHO, Izete. **A ordem V-DP em construções mono argumentais**: uma restrição sintático semântica. 2000. 244 f. Tese (Doutorado) - Universidade Federal de Santa Catarina, Florianópolis, SC, 2000.

COELHO, F. A. **Noções elementares de gramática portuguesa**. Porto: Lemos & Cia. Editores, 1891.

COSERIU, Eugênio. **Sincronia, diacronia e história**: o problema da mudança linguística. Tradução: Carlos Alberto da Fonseca e Mário Ferreira. Rio de Janeiro: Presença, 1979 [1921].

COSTA, João. Word Order and Discourse-Configurationality in European Portuguese. In: COSTA, João (Ed.). **Portuguese Syntax**: new comparative studies. New York: Oxford University Press, 2000. p. 94-115.

CUNHA, Celso; CINTRA, Lindley. **Nova Gramática do português contemporâneo**. 2. ed. Rio de Janeiro: Nova fronteira, 1985.

CYRINO, Sonia M. L. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. ROBERTS, Ian; KATO, Mary A. (Org.). **Português Brasileiro: uma viagem diacrônica**. Campinas: Unicamp, 1993. p. 163-184.

DUARTE, Inês. Relações gramaticais, esquemas relacionais e ordem de palavras. In: MATEUS, M. H. M. et al. **Gramática da língua portuguesa**. 5. ed. Lisboa: Caminho editorial, 2003. p. 275-321.

DUARTE, Maria E. L. A perda da ordem V(verbo) S(sujeito) em interrogativas QU- no português do Brasil. [S.l.]: **D.E.L.T.A**, 1992. v. 8, n. especial, p. 37-52.

DUARTE, Maria E. L. O sujeito expletivo e as construções existenciais. In: RONCARATI, C. ABRAÇADO, J. (Org.). **Português brasileiro: contato linguístico, heterogeneidade e história**. Rio de Janeiro: Viveiros de castro; 7 Letras, 2003. p. 123-131.

DUARTE, Maria E. L.. A estrutura da oração. In: VIEIRA, Silvia Rodrigues; BRANDÃO, Silvia Figueiredo (Org.). **Morfossintaxe e ensino do português: reflexões e propostas**. Rio de Janeiro: Faculdade de Letras da UFRJ, 2004. p. 27-39.

ELISEU, A. M. G. S. **Verbos ergativos do português: descrição e análise**. 1984. Trabalho de síntese (Aptidão Pedagógica e Capacidade Científica) – Universidade de Lisboa, Lisboa, 1984.

FERNÁNDEZ, Francisco Moreno. **Principios de sociolingüística y sociología del lenguaje**. Barcelona: Editorial Ariel, 1998.

FERNANDES, Flaviane Romani. **Ordem, focalização e preenchimento em português: sintaxe e prosódia**. Campinas. 2007. 415f. Tese (Doutorado) - Universidade Estadual de Campinas, 2007.

FRANCHETTO, Bruna; LEITE, Yonne. 500 anos de línguas indígenas no Brasil. In: CARDOSO, Suzana; MOTA, Jacyra; MATTOS E SILVA, Rosa Virgínia (Org.). **Quinhentos Anos de História Lingüística do Brasil**. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. p. 15-61.

FUNDAÇÃO CULTURAL PALMARES. Disponível em: <<http://www.palmares.gov.br>>. Acesso em: 25 de maio de 2009.

FIGUEIREDO SILVA, Maria Cristina. **A posição do sujeito em português brasileiro: Frases finitas e infinitivas**. Campinas: UNICAMP, 1996.

GUY, Gregory R.; ZILLES, Ana. **Sociolingüística quantitativa: instrumental de análise**. Parábola. São Paulo: Parábola, 2007.

HAEGEMAN, L. **Introduction to Government and Binding Theory**. Oxford: Blackwell, 1996.

HUBER, J. **Gramática do português antigo**. Lisboa: Gulbenkian, 1986.

HALL, Stuart. **Da diáspora e mediações culturais**. Tradução: Adelaine La Guardia Resende et al. Belo Horizonte: Editora UFMG, 2003.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Rio de Janeiro, 2006.
Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 08 maio 2009.

KATO, Mary A.; TARALLO, Fernando. Restrictive VS syntax in Brazilian Portuguese: its correlation with invisible clitics and visible subjects. In: GEORGETOWN ROUNDTABLE IN LANGUAGES AND LINGUISTICS, 1988, [S.l.]. **Anais...**[S.l.: s.n], 1988.

KATO, Mary A. A restrição de monoargumentalidade da ordem VS no português do Brasil. **Fórum Lingüístico**, Fpolis, n. 2, p. 97-127, out-dez. 2000.

KATO, Mary A.; MIOTO, Carlos. As interrogativas Q do português europeu e do português brasileiro atuais. **Revista ABRALIN**, Belo Horizonte, v. 4, n. 1 e 2, p. 171-196, dez. 2005.

KATO, Mary A. et al. O português brasileiro no fim do século XIX e na virada do milênio. In: CARDOSO, Suzana; MOTA, Jacyra; MATOS e SILVA, Rosa Virgínia (orgs.). **Quinhentos Anos de História Lingüística do Brasil**. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. p. 413-438.

KATO, Mary A.; DUARTE, M. Eugênia Lamoglia. **The loss of VS order in wh-questions in Brazilian portuguese and the null subject parameter**. Paper presented at the 1996 NWAWE, Las Vegas, 1998.

LABOV, William. **Padrões Sociolingüísticos**. Tradução: Marcos Bagno, Maria Marta Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola, 2008[1972].

LABOV, William. Contraction, deletion and inherent variability of the English copula. **Language**, n. 45, p. 715-62, 1969.

LABOV, William. Where does the linguistic variable stop? A response to Beatriz Lavandera. **Working Paper in sociolinguistics**, Texas, n. 44, p. 5-23, 1978.

LABOV, William. **Principles of linguistic change: v. 2 Social factors**. Massachusetts/Oxford: Blackwell, 2001.

LABOV, William. The gender paradox. In: **Principles of linguistic change. v. 2**. Massachusetts/Oxford: Blackwell, 2001, p. 260-293.

LAMBRECHT, Knud. **Information Structure and Sentence Form: Topic, Focus, and the mental representations of discourse referents**. Cambridge: Cambridge University Press, 1994.

LAVANDERA, Beatriz. Where does the sociolinguistic variable stop? **Working Paper in sociolinguistics**, Texas, n. 40, p. 5-24, 1977.

LIRA, Solange. Subject posposition in Portuguese. **D.E.L.T.A.** v. 2, n. 1, 1986.

LLEAL, Coloma. **La formación de las lenguas romances peninsulares**. Barcelona: Barca Nova, 1990.

LOBO, Tânia. A questão da periodização da história lingüística do Brasil. In: CASTRO, IVO; DUARTE, Inês (Org.). **Razões e emoção**: miscelânea de estudos em homenagem a Maria Helena Mira Mateus. Lisboa: Imprensa Nacional; Casa da Moeda, v. 1, 2003. p. 395-409.

LOBO, Tânia; LUCCHESI, Dante. Gramática e ideologia. **Sitientibus**, Feira de Santana, v. 8, p. 73-81, 1988.

LUCCHESI, Dante. As duas grandes vertentes da história sociolingüística do Brasil. São Paulo: **D.E.L.T.A.**, 2001. p. 97- 130.

LUCCHESI, Dante. O conceito de transmissão lingüística irregular e o processo de formação do português do Brasil. In: RONCARATI, Cláudia; ABRAÇADO, Jussara (orgs.). **Português Brasileiro**: contato lingüístico, heterogeneidade e história. Rio de Janeiro: 7 Letras. 2003. p. 272-283.

LUCCHESI, Dante. **Sistema, Mudança e Linguagem**: um percurso na história da lingüística moderna. São Paulo: Parábola, 2004a.

LUCCHESI, Dante. Norma lingüística e realidade social. In: BAGNO, Marcos (Org.). **Lingüística da norma**. São Paulo: Loyola, 2004b.

LUCCHESI, Dante. Parâmetros sociolingüísticos do português brasileiro. **Revista da ABRALIN**, Belo Horizonte, v. 5, n. 1 e 2, p. 83-112, 2006.

LUCCHESI, Dante. Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão de conjunto. In: RONCARATI, Cláudia; ABRAÇADO, Jussara. (Org.). **Português Brasileiro II**: contato lingüístico, heterogeneidade e história. Niterói: EdUFF, p. 366-390, 2008a.

LUCCHESI, Dante. História do contato entre línguas no Brasil. In.: BAXTER, Alan; LUCCHESI, Dante; RIBEIRO, Ilza (Org.). **O português afro-brasileiro**. Salvador: EDUFBA, 2009. p. 41-73.

MATEUS, Maria Helena Mira et al. **Gramática da Língua Portuguesa**. 5. ed. Lisboa: Caminho, 2003.

MATTOS E SILVA, Rosa Virgínia. **O português aracaico**: morfologia e sintaxe. São Paulo: Contexto, 1993.

MATTOS E SILVA, Rosa Virgínia. **Ensaio para uma socio-história do português brasileiro**. São Paulo: Parábola, 2006.

MENDES, Elisângela dos Passos. **A flexão de caso dos pronomes pessoais no português popular do interior do estado da Bahia**. 2009, 130 f. Dissertação (Mestrado) - Universidade Federal da Bahia, Salvador, 2009.

MENUZZI, Sérgio. A ordem verbo-sujeito no português do Brasil: algumas abordagens e questões em aberto. **Novas Perspectivas sobre a Ordem VS em Pernambuco**, Rio de Janeiro, Revista da ANPOLL, v. 16, p. 349-384, jan 2003. Disponível em: <http://www.geocities.com/smenuzzi/download/ordem_vs_pb_abralin_2003.pdf.htm> Acesso em 03 set. 2007.

MIOTO, Carlos; SILVA, Maria Cristina Figueiredo; LOPES, Ruth Elizabeth Vasconcelos. **Novo Manual de sintaxe**. Florianópolis: Insular, 2004.

MONTEIRO, José Lemos. **Para compreender Labov**. Petrópolis: Vozes, 2002.

MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). **Introdução à sociolinguística: o tratamento da variação**. 2. ed. São Paulo: Contexto, 2004.

MONTE-MÓR, Roberto Luís de Melo. Urbanização extensiva e lógicas de povoamento: um olhar ambiental. In: **Território, Globalização e Fragmentação**. São Paulo: HUCITEC-ANPUR, 1994.

MUSSA, Alberto. **O papel das línguas africanas na história do português do Brasil**. 1991. Dissertação (Mestrado) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1991.

NARO, Anthony Julius. Modelos quantitativos e tratamentos estatísticos. In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). **Introdução à sociolinguística: o tratamento da variação**. [S.I.:s.n], 2004. p. 15-25.

NARO, Anthony Julius. O dinamismo das línguas. In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). **Introdução à sociolinguística: o tratamento da variação**. [S.I.:s.n], 2004. p. 43-50.

NEVES, Maria Helena de Moura. A questão da ordem na gramática tradicional. In: CASTILHO, Ataliba de (Org.). **Gramática do português falado**. Vol 1. São Paulo: FAPESB, 1990. p. 187-310.

OLIVEIRA, Maria Inês Côrtez. **O liberto: seu mundo e os outros**. Salvador, 1790-1890. São Paulo: Corrupio, 1988.

OMENA, Nelise Pires de. A referência variável da primeira pessoa do discurso no plural. In: NARO, Anthony et al. **Relatório Final de Pesquisa: Projeto Subsídios do Projeto Censo à Educação**. Rio de Janeiro, UFRJ, 1986. v. 2. p. 286-319.

PAIXÃO, M. C. A. **A sintaxe da ordem no Português, 1400-1900**. Projeto. 2007. Disponível em: <http://www.ime.usp.br/~tycho/participants/psousa/projeto/projeto_ordem.pdf>. Acesso em: 12 set. 2009.

PÁDUA, M. P. de. **A ordem das palavras no português arcaico: frases de verbo transitivo**. Coimbra: Universidade de Coimbra, 1960.

PAIVA, Maria Conceição A. de; DUARTE, Maria Eugênia Lamoglia. Quarenta anos depois: a herança de um programa na sociolinguística brasileira. In: WEINREICH, Uriel; LABOV, William; HERZOG, Marvin I. **Fundamentos empíricos para uma teoria da mudança linguística**. [S.l.: s.n.], 2006. p. 131-151.

PERLMUTTER, D. Impersonal passive and the unaccusative hypothesis. In: JAEGER, J. et al. (Ed.). **Proceedings from the IV Annual Meeting of the BLS**. Berkley: Califórnia, 1978.

PETTER, Margarida Maria Taddoni. Línguas Africanas no Brasil. In: CARDOSO, Suzana; MOTA, Jacyra; MATOS e SILVA, Rosa Virgínia (Org.). **Quinhentos Anos de História Lingüística do Brasil**. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. p. 17-142.

PILATI, Eloísa Nascimento Silva. **Aspectos sintáticos e semânticos das orações com ordem Verbo-Sujeito no português do Brasil**. 2006. Tese (Doutorado) - Universidade de Brasília, Brasília, 2006.

PILATI, Eloísa Nascimento Silva. Considerações sobre sujeitos pós-verbais em português do Brasil, francês e italiano.[S.l.]: ANPOLL, 2004.

PINTO, Ivone Isidoro; FIORETI, Maria Thereza Gomes. **Tutorial para o pacote VARBRUL**. [S.l.: s.n.], 1992.

PINTZUK, Suzan. **Programas Varbrul**. Tradução: Ivone Isidoro Pinto, revisão de Maria Thereza Gomes Fioreti e Maria Marta Pereira Scherre (coord.). Rio de Janeiro: UFRJ, Faculdade de Letras, 1988.

POGGIO, Rosauta Maria Galvão Fagundes. **Processo de gramaticalização de preposições do Latim ao Português: uma abordagem funcionalista**. Salvador: EDUFBA, 2002.

PONTE, Vanessa. **A indeterminação do sujeito no português rural do interior do estado da Bahia**. 2008. 125f. Dissertação (Mestrado) - Universidade Federal da Bahia, Salvador, 2008.

PROGRAMA Brasil Quilombola. Disponível em : <<http://bvms.saude.gov.br/bvs/publicacoes/brasilquilombola2004.pdf.htm>>. Acesso em: 10 jun. 2009.

PROJETO Vertentes. Disponível em: <<http://www.vertentes.ufba.br/>>. Acesso em: 10 abr. 2009.

RIZZI, L. Residual verb second and the Wh-criterion. In BELETI, A; RIZZI, L. (Ed.) **Parameters and functional heads: essays in comparative syntax**. New York; Oxford: Oxford University Press , 1996. p. 64-90.

ROCHA LIMA, Carlos Henrique da. **Gramática normativa da língua portuguesa**. 38 ed. Rio de janeiro: José Olympio, 1972.

RODRIGUES, Aryon D. **Línguas indígenas: 500 anos de descobertas e perdas**. São Paulo D.E.L.T.A. 9 (1), 1993. p. 83-103.

ROMAINE, Suzane. **Language in society: an introduction to socio-linguistics**. Oxford: Oxford University Press, 1994.

ROSSI, Maria Aparecida Garcia Lopes. Estudo diacrônico sobre as interrogativas do português do Brasil. In: ROBERTS, Ian e KATO, Mary A. (Org.). **Português Brasileiro: uma viagem diacrônica**. Campinas: Unicamp, 1993, p. 307-342.

SANTOS, Miguel Cerqueira dos. **A dinâmica urbana de Santo Antônio de Jesus – Bahia**. 1999. 86f. Dissertação (Mestrado) - Universidade Federal da Bahia, Salvador. 1999.

SANTOS, Danielle Resende. **A ordem VS/SV com verbos inacusativos um estudo diacrônico**. 2008. 110f. Dissertação (Mestrado) - Universidade Federal do Rio de Janeiro, 2008.

SATHLER, Douglas. O rural e o urbano no Brasil. **Caderno de Geografia**, Belo Horizonte, v. 15, n. 25, p. 77-92, 2005.

SAUSSURE, F. **Curso de Lingüística Geral**. Tradução: Antônio Chelini, José Paulo Paes e Izidoro Blikstein. 28. ed. São Paulo: Cultrix, 2006 [1916].

SCHERRE, Marta; NARO, Anthony Julius. Análise quantitativa e tópicos de interpretação do Varbrul. In: MOLLICA, Maria Cecília; BRAGA, Maria Luísa. **Introdução à Sociolingüística: O tratamento da variação**. São Paulo: Contexto, 2003, p. 43-50.

SCHERRE, Maria Marta Pereira; NARO, Antony Julius. Análise quantitativa e tópicos de interpretação do VARBRUL. In: MOLLICA, Maria Cecília; BRAGA, Maria Luíza (Org.). **Introdução à sociolingüística: o tratamento da variação**. 2004, p. 147-177.

SCHERRE, Maria Marta Pereira. **Introdução ao pacote VARBRUL para microcomputadores**. Brasília: UnB. 1993.

SILVA, Jorge Augusto Alves. **A concordância verbal de terceira pessoa do plural no português popular do Brasil: um panorama sociolingüístico de três comunidades do interior do estado da Bahia**. 2005. 324 f. Tese (Doutorado) - Universidade Federal da Bahia, Salvador, 2005.

SILVA, Maria Cristina Vieira de Figueiredo. **O objeto direto anafórico no dialeto rural afro-brasileiro**. 2004, 149 f. Dissertação (Mestrado) - Universidade Federal da Bahia, Salvador, 2004.

SPANO, Maria. **A ordem Verbo sujeito no português brasileiro e europeu: um estudo sincrônico da escrita padrão**. 2008. 190 f. Tese (Doutorado) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

TARALLO, Fernando. **A pesquisa sociolingüística**. São Paulo: Ática, 1985.

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. In: ROBERTS, Ian e KATO, Mary A. (Org.). **Português Brasileiro: uma viagem diacrônica**. Campinas: Unicamp, 1993, p. 69-105.

TEYSSIER, Paul. **História da língua portuguesa**. Tradução de Celso Cunha. São Paulo: Martins Fontes, 1997 [1980].

VASCONCELOS, A. R. G. de. **Gramática portuguesa**. Paris, Lisboa: Guilhard, Aillaud, 1898).

VOTRE, Sebastião José. Relevância da variável escolaridade In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). **Introdução à Sociolinguística: o tratamento da variação**. 3. ed. São Paulo: Contexto, 2003, p. 51-58.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. **Fundamentos empíricos para uma teoria da mudança linguística**. Tradução: Marcos Bagno. São Paulo: Parábola, 2006 [1968].