



**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](mailto:pgletba@ufba.br)



**LUANDA ALMEIDA FIGUEIREDO DE OLIVEIRA**

**O USO VARIÁVEL DO ARTIGO DEFINIDO NA COMUNIDADE RURAL AFRO-  
BRASILEIRA DE HELVÉCIA-BA.**

Salvador  
2011

**LUANDA ALMEIDA FIGUEIREDO DE OLIVEIRA**

**O USO VARIÁVEL DO ARTIGO DEFINIDO NA COMUNIDADE RURAL AFRO-  
BRASILEIRA DE HELVÉCIA-BA.**

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  
2011

Sistema de Bibliotecas da UFBA

Oliveira, Luanda Almeida Figueiredo de.

O uso variável do artigo definido na comunidade rural afro-brasileira de Helvécia-BA / Luanda Almeida Figueiredo de Oliveira. - 2011.

127 f. : il.

Orientador: Prof. Dr. Dante Lucchesi.

Dissertação (mestrado) - Universidade Federal da Bahia, Instituto de Letras, Salvador, 2011.

1. Língua portuguesa - Artigo. 2. Sociolinguística. 3. Língua portuguesa - Português falado - Helvécia (BA). I. Lucchesi, Dante. II. Universidade Federal da Bahia. Instituto de Letras.  
III. Título.

**Luanda Almeida Figueiredo de Oliveira**

**O USO VARIÁVEL DO ARTIGO DEFINIDO NA COMUNIDADE RURAL AFRO-  
BRASILEIRA DE HELVÉCIA-BA.**

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:**

**BANCA EXAMINADORA:**

---

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

---

**Prof. Dinah Callou – Universidade Federal do Rio de Janeiro (UFRJ)**

---

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

*Ao meu marido Flávio Simões, companheiro  
de todas as horas que sempre incentivou  
minhas escolhas.*

## **Agradecimentos**

A Jesus Cristo, Senhor da minha vida, por tê-lo conhecido um dia e, desde então, não ter sido a mesma.

Aos meus pais, Wellington Figueiredo e Maria da Conceição Almeida, por todo apoio e carinho. Destacar a importância na busca do conhecimento foi essencial em minha vida.

Ao meu tio Agnaldo Almeida pelo exemplo de perseverança nos estudos.

Ao meu amado esposo Flávio que não poupou esforços em ajudar. Obrigada por sua compreensão nos momentos em que eu estive ausente. Amo muito você!

Ao meu orientador, Dante Lucchesi, pela paciência diante das minhas limitações e por incentivar meu caminho na Sociolinguística. Sua dedicação aos estudos linguísticos e seu conhecimento sobre o português do Brasil são inspiradores.

À minha avó Neca pelas palavras de sabedoria.

Às minhas irmãs, Quênia e Lise, pelos momentos de descontração e pelos conselhos nos momentos difíceis.

Às amigas Lanuza Lima, Darliane Rocha e Maria Cecília Guimarães pelas palavras doces. Os encontros, especialmente com Lanuza, foram imprescindíveis.

À Elisângela Mendes, Vívian Antonino e Telma Assis, por saberem ouvir e acompanhar de perto minhas dúvidas e aflições. A disposição de Vívian em ajudar foi um alento.

Aos irmãos em Cristo, Rosana Assis, Lourdes Albergaria e Edvaldo Santos, pelas orações.

À Ana Paula Carneiro companheira sempre presente nas aulas do Mestrado.

Aos primos, Rívia e Rodrigo, por lembrar que minha “vida acadêmica” não corria risco de ir à UTI.

Às amigas de pesquisa do Projeto Vertentes, Renata, Camila, Shirley, Manuele e Heide pelo carinho e estímulo.

Às crianças, Gabriel, João Pedro, Luciano Júnior, Lucas, Matheus e Marcelo por suavizar meus dias com afagos. Em especial, agradeço à Clara, minha pequena orientadora e amiga, por cobrar agilidade e eficiência na entrega dos capítulos.

À professora Ilza Ribeiro por ceder seu tempo, ajudando significativamente na elaboração de algumas variáveis.

Aos professores da pós-graduação em Letras, que me acompanham desde a época da graduação, pelas valiosas sugestões.

Aos funcionários da secretaria da pós-graduação, em especial ao Sr. Wilson, pelo apoio e “avisos” durante dois anos.

Ao CNPq por custear os meus estudos e por possibilitar a realização desta dissertação.

Por fim, agradeço a todas as pessoas que contribuíram para a realização deste trabalho.

*Quase me apetece dizer que não há uma língua portuguesa;  
há línguas em português.*

José Saramago (no depoimento apresentado no documentário  
Língua: vidas em português, dirigido por Victor Lopes).



## RESUMO

A ausência do artigo definido em sintagmas nominais (SNs) com referência definida é uma característica proeminente de línguas crioulas. Trabalhos como os realizados por Baxter e Lopes (2004, 2009) e Ribeiro (2010) já identificaram diferenças no uso do artigo definido na fala de indivíduos pertencentes a uma comunidade do português rural afro-brasileiro. Neste trabalho, com base no aporte teórico-metodológico da Sociolinguística Variacionista, investigam-se os fatores linguísticos, discursivos-pragmático e sociais que presidem a escolha do falante em empregar ou não o artigo definido. A comunidade analisada é o município de Helvécia, localizada no extremo sul do estado da Bahia, onde foram observadas amostras de falas de 12 pessoas, distribuídas pelos dois sexos, em quatro faixas etárias: faixa 1 (25 a 35 anos), faixa 2 (45 a 55 anos), faixa 3 (65 a 75 anos) e faixa 4 (mais de 80 anos), considerando-se dois níveis de escolaridade: analfabetos e semi-analfabetos. Apenas as funções sintáticas de *sujeito* e *objeto direto* foram analisadas. Os dados apontam para o perfil aquisicional do fenômeno em direção à variedade padrão e revela que os mais velhos conservam vestígios que testemunham os processos de reestruturação morfossintática promovidos pelo contato entre línguas no passado da comunidade. A variável social *estada fora da comunidade* e a variável *escolaridade* também foram selecionadas como significativas pelo pacote de programas estatísticos VARBRUL. Quanto aos fatores de natureza sintática, semântica e discursiva foram apontados como fortes favorecedores do emprego do artigo definido as variáveis *referencialidade*, *função sintática*, *número do SN*, *determinante na menção anterior* e *contabilidade do núcleo do SN*.

Palavras chave: Artigo definido. Sociolinguística. Contato entre línguas. Português popular brasileiro.

## ABSTRACT

The absence of the definite article in nominal phrases (NPs) with definite references is an outstanding feature of Creole languages. Similar works to those made by Baxter and Lopes (2004, 2009) and Ribeiro (2010) have already identified some differences in the use of definite articles on the speech of individuals from an afroBrazilian Portuguese speaking rural community. This work investigates the linguistic, discursive, pragmatic and social factors which govern the choice of speaker on the moment of the decision of either using a definite article or not, based on the theoretical and methodological framework of Sociolinguistic Variations. The chosen community for the analysis is Helvécia Town which is located in the south of Bahia state. The work observed speech samples of 12 people, distributed by both sexes within four age groups: age group 1 (25 to 35 years old), age group 2 (45 to 55 years old), age group 3 (65 to 75 years old) and age group 4 (over 80 years old), considering two schooling levels: illiterate and semi-literate. Only the syntactic functions of *subject* and *direct object* were analyzed. The data indicate the phenomenon aquisicional profile towards the standard variety and show that the oldest people preserve traces which witness the morphosyntactic restructuring processes promoted by contact among languages spoken in the past of the community. The social variable *being out of the community* and the variable *schooling* were also chosen as significant by the VARBRUL statistic software package. Regarding the factors of syntactic, semantic and speech nature, the variables, *referentiality*, *syntactic function*, *NP's number*, *determinant of the previous mention* and *NP's core accounting* were identified as strong promoters of the use of the definite article

Key words: Definite Article. Sociolinguistics. Contact among languages. Popular Brazilian Portuguese.

## LISTA DE QUADROS

QUADRO 1: Gradação etária segundo a hipótese clássica.....	30
QUADRO 2: Relações entre tempo aparente x tempo real .....	31
QUADRO 3: Sistema referencial do crioulo de Cabo Verde, segundo Baptista (2007).....	87

## LISTA DE TABELAS

TABELA 1: Perfil social dos inquiridos selecionados para a realização da pesquisa – <i>corpus</i> do português afro-brasileiro.....	38
TABELA 2: População do Brasil por etnia do século XVI ao XIX.....	55
TABELA 3: Semelhanças e diferenças no uso do artigo definido diante de nomes próprios e de possessivos – comparação diacrônica e sincrônica – depreendidos de Costa (1998) .....	79
TABELA 4: A presença do artigo definido em SNs, segundo a variável <i>nível de referencialidade</i> .....	99
TABELA 5: A presença do artigo definido em SNs, segundo a variável <i>função sintática</i> .....	100
TABELA 6: A presença do artigo definido em SNs, segundo a variável <i>número do SN</i> .....	102
TABELA 7: : A presença do artigo definido em SNs, segundo a variável <i>determinante na menção anterior</i> .....	104
TABELA 8: A presença do artigo definido em SNs, segundo a variável <i>contabilidade do núcleo do SN</i> .....	106
TABELA 9: A presença do artigo definido em SNs, segundo a variável <i>faixa etária</i> .....	107
TABELA 10: A presença do artigo definido em SNs, segundo a variável <i>estada fora da comunidade</i> .....	109
TABELA 11: A presença do artigo definido em SNs, segundo a variável <i>escolaridade</i> .....	110
TABELA 12: A variação do artigo definido em SNs, segundo o cruzamento das variáveis <i>referencialidade x função sintática</i> .....	111
TABELA 13: A presença do artigo definido em SNs, segundo a variável <i>gênero do núcleo do SN</i> .....	112
TABELA 14: A presença do artigo definido em SNs, segundo a variável <i>animacidade do núcleo do SN</i> .....	113
TABELA 15: A presença do artigo definido em SNs, segundo a variável <i>modificadores pós-nucleares</i> .....	115
TABELA 16: A presença do artigo definido em SNs, segundo a variável <i>animacidade do núcleo do SN</i> .....	116
TABELA 17: A presença do artigo definido em SNs, segundo a variável <i>estatuto informacional</i> .....	117

## LISTA DE FIGURAS

FIGURA 1: Rota dos escravos do século XVI ao século XIX.....	52
FIGURA 2: Periodização da história linguística do Brasil, segundo dados depreendidos de Lobo (2003) e Lucchesi (2008) .....	54
FIGURA 3: <i>Língua, norma e fala</i> , segundo a teorização de Coseriu (1979 [1921]) .....	60
FIGURA 4: Distribuição das variedades linguísticas do PB, conforme o <i>continuum</i> de urbanização .....	63
FIGURA 5: <i>Continuum</i> de definitude entre o demonstrativo e os artigos definidos e indefinidos .....	69

## LISTA DE GRÁFICOS

GRÁFICO 1: Frequência do artigo definido diante de possessivo, adaptado de Callou e Silva (1996, p. 3). .....	76
GRÁFICO 2: Frequência do artigo definido diante de nome próprio, adaptado de Callou e Silva (1996, p. 3). .....	77
GRÁFICO 3: Frequência do artigo definido segundo a variável faixa etária, segundo Baxter e Lopes (2004, 2009). .....	89
GRÁFICO 4: A presença do artigo definido em SNs na comunidade de Helvécia-BA. ....	96
GRÁFICO 5: Presença/ausência do artigo definido em SNs com relação a <i>faixa etária</i> . ....	107
GRÁFICO 6: Cruzamento das variáveis <i>referencialidade x função sintática</i> . ....	111

## SUMÁRIO

<b>INTRODUÇÃO</b> .....	<b>17</b>
<b>1 PRESSUPOSTOS TEÓRICO-METODOLÓGICOS</b> .....	<b>20</b>
1.1 FUNDAMENTAÇÃO TEÓRICA .....	20
<b>1.1.2 Variação e mudança linguística</b> .....	<b>23</b>
<b>1.1.3 Mudança em tempo aparente e em tempo real</b> .....	<b>28</b>
1.2 A TRANSMISSÃO LINGUÍSTICA IRREGULAR.....	32
1.3 METODOLOGIA.....	36
<b>1.3.1 Seleção e caracterização do <i>corpus</i></b> .....	<b>36</b>
<b>1.3.2 Procedimentos da pesquisa</b> .....	<b>39</b>
<b>1.3.3 Tratamento estatístico dos dados</b> .....	<b>40</b>
1.4 A VARIÁVEL DEPENDENTE.....	42
<b>1.4.1 Casos desconsiderados na pesquisa</b> .....	<b>43</b>
<b>1.4.2 Variáveis linguísticas e extra-linguísticas analisadas</b> .....	<b>46</b>
<b>2 HISTÓRIA SOCIOLINGUÍSTICA DO BRASIL</b> .....	<b>48</b>
2.1 AS ORIGENS DO PORTUGUÊS BRASILEIRO.....	48
<b>2.1.1 Reflexos da polarização linguística: convivência de normas x ensino</b> .....	<b>59</b>
2.2 A COMUNIDADE DO PORTUGUÊS AFRO-BRASILEIRO ESTUDADA .....	64
<b>2.2.1 O português afro-brasileiro do interior do estado da Bahia</b> .....	<b>64</b>
2.2.1.1 <i>Helvécia</i> .....	66
<b>3 O ARTIGO DEFINIDO EM PORTUGUÊS</b> .....	<b>68</b>
3.1 FORMAÇÃO E DEFINIÇÃO TRADICIONAL DO ARTIGO DEFINIDO. ....	68
3.2 AS FUNÇÕES DO ARTIGO DEFINIDO .....	71
3.3 PROPRIEDADES SEMÂNTICAS E SINTÁTICAS DO ARTIGO DEFINIDO ..	73
<b>3.3.1 Sobre a ausência de artigo definido</b> .....	<b>74</b>
3.4 A VARIAÇÃO NO USO DO ARTIGO DEFINIDO EM PORTUGUÊS.....	76
3.5 O ARTIGO DEFINIDO NAS LÍNGUAS CRIOLAS DE CABO VERDE E SÃO TOMÉ E PRÍNCIPE .....	81
3.6 O USO DO ARTIGO DEFINIDO NO PORTUGUÊS AFRO-BRASILEIRO .....	88
<b>4 ANÁLISE DOS DADOS</b> .....	<b>94</b>
4.1 VARIÁVEL DEPENDENTE.....	94
4.2 VARIÁVEIS LINGUÍSTICAS SIGNIFICANTES .....	96

4.2.1 Referencialidade .....	96
4.2.2 Função sintática do SN.....	100
4.2.3 Número do SN.....	101
4.2.4 Determinante na menção anterior.....	102
4.2.5 Contabilidade do núcleo do SN.....	104
4.3 VARIÁVEIS EXTRALINGUÍSTICAS SIGNIFICANTES.....	106
4.3.1 Faixa etária.....	106
4.3.2 Estada fora da comunidade.....	108
4.3.3 Escolaridade.....	109
4.4 CRUZAMENTO SIGNIFICATIVO.....	110
4.4.1 Referencialidade x função sintática.....	110
4.5 OUTRAS VARIÁVEIS ANALISADAS.....	112
4.5.1 Gênero do núcleo do SN.....	112
4.5.2 Tipo de núcleo do SN.....	113
4.5.3 Presença de modificadores pós-nucleares.....	114
4.5.4 Animacidade do núcleo do SN.....	115
4.5.5 Estatuto informacional.....	116
CONSIDERAÇÕES FINAIS.....	118
REFERÊNCIAS.....	121



## INTRODUÇÃO

O português falado no Brasil apresenta uma gama de transformações decorrentes de um passado marcado pelo encontro de povos de línguas e culturas distintas. A herança do período colonial fez do Brasil um país pluriétnico, sendo os africanos e afro-brasileiros os principais agentes do *português geral brasileiro, antecedente histórico do chamado português popular brasileiro*. (SILVA, 2001).

A análise do sistema de determinantes na comunidade de Helvécia, localizada no extremo sul do estado da Bahia, originada de um remanescente quilombola, tem assinalado um intenso processo de reformulação em direção ao padrão culto da língua. Os resultados obtidos por Baxter e Lopes (2004, 2009) e, mais recentemente, por Ribeiro (2010) afirmam a existência de características similares em línguas crioulas, fato que incitou maior profundidade sobre o assunto.

Esta dissertação, assim, volta-se para o estudo do artigo definido em SNs somente nas funções de *sujeito* e de *objeto direto*. O estabelecimento em priorizar apenas essas duas funções sintáticas decorreu da observação de que a presença da preposição motivaria o emprego do artigo definido. Além disso, os resultados encontrados por Baxter e Lopes (2004, 2009) não revelaram resultados que contrastassem as duas posições. A outra diferença concentrou-se nos aspectos semânticos do substantivo. Em relação à variável *nível de referencialidade*, elegemos cinco condições de referência: *definido pragmaticamente, definido anaforicamente, genérico, existencial e entidade única*. As variáveis *tipo de núcleo do SN, presença de modificadores pós-nucleares e estatuto informacional* também são investigadas pela primeira vez.

A inclusão das variáveis sociolinguísticas *sexo, estada fora da comunidade e nível de escolaridade* foi de extrema importância para determinar a configuração linguística da comunidade de Helvécia. Observar a variável *faixa etária* e reafirmar o perfil aquisicional da comunidade em direção à variante padrão é um dos objetivos desta dissertação.

Diante do exposto, traça-se como objetivo principal desta dissertação descrever o uso variável do artigo definido em sintagmas nominais (SNs) na comunidade de Helvécia, ao qual se juntam outros mais específicos, quais sejam: a) estabelecer a influência de fatores linguísticos e extralinguísticos na expressão da variável e b) confrontar os resultados com estudos sobre os crioulos de base lexical portuguesa (LUCCHESI, 1993; BAPTISTA, 2007).

Com vistas a se conseguirem os objetivos traçados, o presente trabalho parte das seguintes hipóteses:

- i) a variação de uso do artigo definido em SNs na comunidade de Helvécia está relacionada à situação histórica de contato linguístico;
- ii) o emprego do artigo definido tem maior produtividade na fala dos mais jovens, demonstrando o perfil aquisicional do fenômeno;
- iii) a escolha do falante por uma das formas alternantes seria motivada por fatores linguísticos, discursivos/pragmáticos e sociais.

Seguindo o modelo teórico da Sociolinguística Variacionista, este estudo será feito com base em amostras vernáculas de 12 indivíduos nascidos em Helvécia, ambos os sexos (feminino e masculino), seguindo dois níveis de escolaridade (analfabeto e semi-analfabeto), informantes que se ausentaram da comunidade ou que nunca saíram e com quatro faixas etárias: faixa I (25 a 35 anos), faixa II (45 a 55 anos), faixa III (65 a 75 anos) e IV acima de 80 anos.

No capítulo 1 é apresentada a base teórica sobre a qual se apóia este trabalho. São discutidos os pressupostos básicos da Sociolinguística Variacionista (WEINREICH; HERZOG; LABOV, 2006 [1968]; LABOV, 2008 [1972]), bem como a hipótese da *transmissão linguística irregular* que direciona as explicações das variações observadas no português brasileiro. É também no primeiro capítulo que são apresentados os procedimentos metodológicos empregados: a variável dependente e as variáveis independentes; os dados descartados da pesquisa e, por fim, o modelo estatístico utilizado para o processamento dos dados.

No capítulo 2, intitulado *História Sociolinguística do Brasil*, são apresentados alguns eventos históricos que explicam a afirmação de que a realidade linguística brasileira é “heterogênea, plural e polarizada” (SILVA, 2006; LUCCHESI, 2006). Há uma breve consideração sobre a relação entre ensino x convivência de normas no Brasil.

O capítulo 3, trás uma breve abordagem da tradição gramatical sobre o artigo definido, sua formação e suas funções. Explicitam-se também estudos sociolinguísticos relacionados ao assunto e que ajudaram na elaboração das variáveis independentes, bem como são descritos e ponderados alguns estudos realizados em línguas crioulas.

No quarto capítulo, apresentamos a análise qualitativa e quantitativa dos dados, seguindo a ordem de relevância abalizada pelo pacote de programas quantitativo VARBRUL.

Por fim são apontadas as considerações finais, sintetizando os principais resultados.

## INTRODUÇÃO

O português falado no Brasil apresenta uma gama de transformações decorrentes de um passado marcado pelo encontro de povos de línguas e culturas distintas. A herança do período colonial fez do Brasil um país pluriétnico, sendo os africanos e afro-brasileiros os principais agentes da formação e difusão do *português geral brasileiro, antecedente histórico do chamado português popular brasileiro*. (MATTOS E SILVA, 2001). Com base nisso, o Projeto *Vertentes do Português Popular do Estado da Bahia* (doravante, Projeto *Vertentes*), coordenado pelo Prof. Dante Lucchesi, no Departamento de Letras Vernáculas da Universidade Federal da Bahia, tem desenvolvido uma ampla pesquisa de campo junto a variedades populares do português brasileiro para identificar processos de mudança que o contato entre línguas teria desencadeado nessas variedades linguísticas em sua formação histórica.

A primeira etapa dessa pesquisa está concentrada nas comunidades rurais afro-brasileiras isoladas, muitas delas oriundas de antigos quilombos, com base no pressuposto de que, se o contato entre línguas afetou a formação histórica de variedades do português brasileiro, os reflexos mais notáveis desse processo apareceriam exatamente na fala dessas comunidades por suas particularidades históricas, étnicas e sócio-econômicas. Já veio a luz um conjunto de resultados expressivos que demonstram que muitas das características da fala dessas comunidades têm a sua origem em processos de mudança desencadeados pelo contato do português com línguas africanas no passado (LUCCHESI; BAXTER; RIBEIRO, 2009). Entretanto, o quadro sociolinguístico observado atualmente nessas comunidades aponta para o desaparecimento dessas características, com processos de mudança em curso em direção ao padrão urbano culto ou semi-culto do português brasileiro (LUCCHESI, 2009).

Integrada nesse programa de pesquisa, esta dissertação focaliza a variação no uso do artigo definido em Sintagmas Nominais não preposicionados, tendo como universo de observação a comunidade afro-brasileira de Helvécia, situada no extremo sul do Estado da Bahia. Essa comunidade é aquela que, no universo de quatro comunidades afro-brasileiras de diferentes regiões do estado da Bahia estudadas pelo Projeto *Vertentes*, parece ter passado pelo processo mais profundo de mudanças em função do contato linguístico (LUCCHESI, 2009). Com base nos pressupostos teóricos e metodológicos da Sociolinguística Variacionista, foi feita uma análise da variação no

uso do artigo definido, tendo como base empírica 12 entrevistas informais, realizadas em janeiro de 1994, com membros da comunidade de Helvécia, distribuídos equitativamente pelos dois sexos e por quatro faixas etárias. A análise buscou identificar os condicionamentos estruturais do fenômeno variável em foco, bem como os seus condicionamentos sociais, visando a estabelecer um diagnóstico nos termos da dicotomia laboviana *variação estável* ou *mudança em progresso*. Foi explorado também um aspecto já ressaltado em estudos anteriores sobre essa variável (BAXTER; LOPES, 2004, 2009; RIBEIRO, 2010) a similaridade entre estruturas encontradas na gramática dessa comunidade de fala e estruturas próprias de línguas crioulas de base lexical portuguesa da África. Nessa linha de investigação, os resultados dessa análise foram cotejados com estudos sobre o sistema de artigos de crioulos de base lexical portuguesa da África (LUCCHESI, 1993; BAPTISTA, 2007).

Dessa forma, o objetivo principal desta dissertação é descrever o uso variável do artigo definido em sintagmas nominais (SNs) na comunidade de Helvécia, buscando identificar os reflexos de processos pretéritos de mudanças induzidas pelo contato entre línguas. Para alcançar esse objetivo, o presente trabalho parte das seguintes hipóteses:

- i) a variação de uso do artigo definido em SNs na comunidade de Helvécia está relacionada à situação histórica de contato linguístico, o que teria levado a uma erosão do emprego dessa partícula gramatical no passado;
- ii) estaria em curso na comunidade um processo de mudança que se caracteriza pelo incremento do uso do artigo definido em função de influência externas;
- iii) no bojo desse processo variável, a escolha do falante por uma das formas alternantes não seria aleatória, mas motivada por fatores linguísticos, discursivos/pragmáticos e sociais.

Assim concebida, esta dissertação apresenta a seguinte estrutura. No capítulo 1 é apresentada a base teórica sobre a qual se sustenta este trabalho. São discutidos os pressupostos básicos da sociolinguística variacionista (WEINREICH; HERZOG; LABOV, 2006 [1968]; LABOV, 2008 [1972]), bem como a hipótese da *transmissão linguística irregular* que direciona as explicações das mudanças induzidas pelo contato entre línguas no português brasileiro. É também no primeiro capítulo que são

apresentados os procedimentos metodológicos empregados: a variável dependente e as variáveis independentes; os dados descartados da pesquisa e, por fim, o modelo estatístico utilizado para o processamento dos dados.

No capítulo 2, intitulado *História sociolinguística do Brasil*, são apresentados alguns eventos históricos que explicam a afirmação de que a realidade linguística brasileira é “heterogênea, plural e polarizada” (LUCCHESI, 2001, 2002, 2006). Há uma breve consideração sobre a relação entre ensino x convivência de normas no Brasil.

O capítulo 3 trás uma breve abordagem da tradição gramatical sobre o artigo definido, sua formação e suas funções. Explicitam-se também estudos sociolinguísticos relacionados ao assunto e que ajudaram na elaboração das variáveis explanatórias, bem como são descritos e ponderados estudos realizados em línguas crioulas.

No quarto capítulo, apresentamos a análise qualitativa e quantitativa dos dados, seguindo a ordem de relevância abalizada pelo pacote de programas quantitativo VARBRUL.

Apontam-se, por fim, as considerações finais, sintetizando os principais resultados.

## 1 PRESSUPOSTOS TEÓRICO-METODOLÓGICOS

### 1.1 FUNDAMENTAÇÃO TEÓRICA

*Se uma língua tem de ser estruturada, a fim de funcionar eficientemente, como ela funciona enquanto a estrutura muda? Weinreich, Labov e Herzog (2006 [1968])*

Durante anos, as correntes linguísticas formais impulsionaram as pesquisas sobre a língua distanciando-a de qualquer implicação social. Para a maioria dos estudiosos que seguiam o arcabouço teórico do modelo estruturalista, era natural deixar a margem os fatores sociais e enfatizar os fatores internos no estudo da língua. Somente durante a década de sessenta, com Willian Labov e seus estudos sobre mudança em progresso na Ilha de Martha's Vineyard (1963) e sobre a estratificação social do inglês da cidade de Nova Iorque (1966), sob a orientação de Uriel Weinreich, teremos a descrição e a análise dos fenômenos linguísticos dentro da perspectiva da Sociolinguística (CHAMBERS, 1995; TARALLO, 2005). Em 1969, é publicado o *Empirical Foundations for a Theory of Language Change*, de autoria de Weinreich, Labov e Herzog, onde são apresentadas as bases para uma nova teoria da mudança linguística,<sup>1</sup> dando respostas às questões olvidadas em detrimento dos limites impostos da concepção de língua difundida pela corrente que deu origem ao *Cours de Linguistique Générale*.

Os autores que abriram caminho para o avanço da Sociolinguística ou Teoria da Variação posicionaram-se contra a identificação entre homogeneidade e sistematicidade, propondo a identificação entre heterogeneidade e sistematicidade. Assim, por trás do aparente “caos” linguístico há uma sistematização, delimitada por meio da estrutura linguística e social, elucidando o fato de as línguas se modificarem sem afetar a eficácia da interação comunicativa. Com a aceitação do “axioma da heterogeneidade ordenada”, dá-se início uma nova forma de estudar a língua, atrelada a

---

<sup>1</sup> Segundo Mattos e Silva (2008, p. 95), o texto fundador do que veio a ser chamado de “teoria da variação e mudança”.

um instrumental metodológico que confere uma rigorosa técnica na apreciação de fenômenos variáveis.

Alguns pesquisadores como Gauchat, André Martinet e Antoine Meillet, anos antes, já apresentavam questionamentos ao rígido suporte estruturalista e à concepção de mudança restrita ao interior do sistema linguístico. Meillet (1921) advertia que, ao separar a variação linguística das categorias externas, Saussure ostentava uma abstração que privava os estudos linguísticos de uma realidade incontestável: o contexto social interfere na variação e na mudança de uma língua. Segundo o autor, *“por ser a língua um fato social, resulta que a lingüística é uma ciência social, e o único elemento variável ao qual se pode recorrer para dar conta da variação lingüística é a mudança social”*. Já não fazia sentido para o linguista francês a dicotomia saussuriana entre sincronia e diacronia, muito menos a asseveração de que *“a lingüística tem por único e verdadeiro objeto a língua considerada em si mesma e por si mesma”* (Curso de Lingüística Geral, 2006). Segundo Calvet (2002, p. 15-17), quase sempre Meillet foi apresentado como discípulo de Saussure, entretanto, *“tudo opõe os dois homens tão logo os situamos no terreno da lingüística geral”* e foi assim *“durante meio século, esses dois discursos se desenvolvem de modo paralelo, sem nunca se encontrar”*.

Apesar de todos os progressos alcançados pela linguística formal, de orientação saussuriana, a verdade é que as abordagens estruturalistas já não davam conta de responder as indagações sobre a mudança linguística e, além disso, o método empregado ignorava o tratamento da variação. Apenas com os avanços da Sociolinguística, teremos o rompimento das dicotomias estanques cunhadas pelas teorias formalistas da Escola de Genebra. A língua, antes concebida como homogênea, autônoma e estática, passa a ser analisada como heterogênea, social e variável, sendo central para a Sociolinguística a preocupação com a contraparte social da língua, já que *“numa língua que serve a uma comunidade complexa (i.e. real), a ausência de heterogeneidade estruturada seria disfuncional”* (WEINREICH, LABOV e HERZOG (2006 [1968]) ou nas palavras de Tarallo (1995): *“[...] é na heterogeneidade refletida através do desempenho que se deve buscar estrutura, sistema e funcionamento”*.

O foco dessa nova abordagem da língua direciona-se para o valor social da linguagem dentro da comunidade de fala, priorizando, em sua investigação, a relação entre aspectos linguísticos e sociais (MOLICA; BRAGA, 2004). Ao trabalhar com o conceito de *comunidade de fala*, a Sociolinguística busca estabelecer quais as características que um grupo de falantes compartilha e, desta forma, relacionar quais os



fatores que agem ou não na variação e na mudança dentro dessa comunidade de fala. Diz Labov (1972) acerca da comunidade de fala:

[...] não é definida por nenhum acordo/contrato no uso de elementos de língua, mas pela participação em um jogo de normas compartilhadas; tais normas podem ser observadas em tipos claros de comportamentos avaliativos e pela uniformidade de modelos abstratos dos padrões da variação que são invariáveis em relação aos níveis particulares de uso.

Assim sendo, o objetivo de uma análise fundamentada na Sociolinguística é examinar os fenômenos variáveis encaixados na estrutura social e na estrutura linguística, buscando observar uma mudança em curso ou uma variação estável do fenômeno focalizado. Nesse modelo, a variação é um fenômeno previsível e sistematizável, o que implica estabelecer os limites do fenômeno variável em foco, bem como os fatores que o regulam, formalizados, respectivamente, nos termos de uma *variável dependente* e de *variáveis independentes* ou *explanatórias*.

É oportuno salientar que sempre existiu, em diferentes enfoques no estudo da língua, a constatação de que a variabilidade é característica geral de todas as línguas humanas. Assim, o que muda no panorama da década de 1960 é o entendimento de que essa variabilidade está estruturada no próprio sistema linguístico. Há duas ou mais formas alternantes para expressar um mesmo significado nos diversos níveis da língua: lexical (ex: *tangerina* ~ *mexirica*), morfossintático (ex: concordância nominal – *as meninas bonitas* ~ *as meninaØ bonitaØ*); fonético-fonológico (ex: vocalismo átono – *melhor* – *m[e]lhor* ~ *m[ɛ]lhor* ~ *m[i]lhor*); discursivo-pragmático (forma do tratamento do interlocutor – *tu* ~ *você* ~ *o senhor/a senhora* ~ *omissão de qualquer pronome*)<sup>2</sup>.

Todas as línguas, dessa maneira, apresentam formas diferentes para expressar um mesmo conteúdo informacional, denominadas *variantes*. As variantes são, assim, escolhas possíveis, dentro de um mesmo contexto linguístico, configurando um fenômeno variável ou uma *variável linguística*. Essa variável linguística pode ser descrita e analisada, pois apresenta uma sistematicidade determinada por fatores estruturais e sociais denominados *variáveis independentes* – são esses grupos de fatores que regulam os fenômenos variáveis e atuam no sentido de motivar ou restringir a frequência de uso de cada variante que forma a variável linguística. A variação e a

<sup>2</sup> Exemplos de fenômenos variáveis no português do Brasil.

mudança são, então, contextualizadas por parâmetros diversos, de natureza linguística e extralinguística que atuam concomitantemente.

Constituem os parâmetros de natureza linguística as **variáveis internas** (fatores fonético-fonológicos; morfossintáticos, léxico-semânticos e discursivos) e os parâmetros extralinguísticos as **variáveis externas**, que abarcam os fatores estritamente sociais e inerentes ao falante como idade, sexo, etnia, localização geográfica, exposição à mídia, e diversas outras variáveis que possibilitam investigar que fato externo exerce influência nas escolhas linguísticas.

Desse modo, o avanço da Teoria da Variação assenta na análise da língua como um sistema heterogêneo, em que se entrecruzam fatores intralinguísticos e extralinguísticos, produzindo padrões sociolinguísticos coletivos (SILVA, 2002). Sob esse ponto de vista, a variação é atributo fundamental da linguagem humana, sendo as estruturas variantes categorias que apresentam padrões de regularidade. Portanto, a Sociolinguística se ocupa da descrição das línguas em sua diversidade funcional e social, revelando como a heterogeneidade linguística se apresenta de forma organizada e regulada por um conjunto de regras funcionais. Com isso, esse modelo teórico contribui consideravelmente para a compreensão da realidade da língua e da própria sociedade.

### **1.1.2 Variação e mudança linguística**

A relação entre tempo, língua e mudança passa por uma nova reflexão com o advento da Sociolinguística. Dentro dessa perspectiva, para avaliar a mudança linguística, é necessário estar consciente de que se trata de um processo complexo, que abarca tanto fatores linguísticos quanto fatores sociais, geográficos, cognitivos e pragmáticos. De acordo com Labov (2008 [1972]), a mudança linguística não pode ser elucidada exclusivamente por argumentos internos, absolutamente, as relações sociolinguísticas precisam ser adotadas como agentes condicionantes. A afirmação de que a variação faz parte do sistema e o reconhecimento do estudo da mudança em progresso são aspectos fundamentais para a Sociolinguística.

Segundo Labov (2008 [1972], p. 192), os principais problemas da mudança poderiam ser resumidos em cinco perguntas:

1. Existe uma direção geral para a evolução lingüística?
2. Quais são os condicionamentos universais da mudança lingüística?
3. Quais as causas do surgimento contínuo de novas mudanças lingüísticas?
4. Por meio de que mecanismo as mudanças ocorrem?
5. Existe uma função adaptativa na evolução lingüística?

Em seguida, Labov elenca três problemas que servem para responder as indagações acima: o problema da transição, o problema do encaixamento e o problema da avaliação.

O **problema da transição** consiste em “*encontrar o caminho pelo qual um estágio de uma mudança lingüística evoluiu a partir de um estágio anterior*”. Busca, desta forma, esclarecer como ocorre uma mudança lingüística, tentando encontrar o estágio intervisor que determina “*a trilha pela qual a estrutura A evoluiu para a estrutura B*” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 122). A Sociolingüística postula que a variação e a mudança se revelam em um *continuum* ininterrupto, contestando a concepção saussuriana de estado de língua – uma sucessão de sistemas homogêneos com fases intervenientes de instabilidade e mudança. O problema da transição associa-se à definição de como se produz a variabilidade de uma língua em determinada comunidade.

O **problema do encaixamento** incide em “*encontrar a matriz contínua de comportamento social e lingüístico em que a mudança lingüística é levada a cabo*”. Refere-se à investigação da natureza e da extensão do encaixe da mudança na estrutura lingüística e nas relações sociais interdependentes. A variação e/ou mudança não ocorre de forma isolada, portanto uma alteração na estrutura da língua irá refletir nas demais estruturas como uma *rede de ligações*. Ressalta-se que a Sociolingüística analisa o problema do encaixamento em duas partes: **a.** o encaixamento na estrutura lingüística e **b.** o encaixamento na estrutura social. Lucchesi (2004, p.175) traça os seguintes pontos, apresentados por Weinreich, Herzog e Labov na estrutura lingüística:

- (i) o sistema lingüístico em que a mudança deve ser encaixada não se situa no indivíduo (i.é. no idioleto), e sim na comunidade de fala;
- (ii) esse sistema caracteriza-se por sua heterogeneidade estruturada, e é funcionalmente diferenciado dentro da comunidade de fala;
- (iii) as variáveis intrínsecas a esse sistema definem-se pela covariação com elementos lingüísticos e extra-lingüísticos;
- (iv) o processo de mudança lingüística raramente é um movimento de um sistema inteiro a outro, e sim o movimento de um conjunto limitado de variáveis de um sistema que altera gradualmente seus valores modais de um pólo a outro;

(v) não obstante, as variantes de cada variável podem ser contínuas ou discretas (as variáveis, contudo, mantêm uma gama contínua de valores que refletem as frequências observadas na atividade lingüística concreta);

(vi) a variação inerente à estrutura lingüística deve ser vista como parte integrante da competência lingüística dos membros da comunidade.

Já na estrutura social, o problema do encaixamento mensura o grau da correlação entre os elementos sociais e a mudança, evidenciando o reconhecimento de que não é possível dissociar as questões lingüísticas das questões sociais, uma vez que as estruturas lingüísticas não seriam aptas, sozinhas, a dar respostas ao processo de mudança. O exame do encaixamento social remete a algumas exigências para a precisão do efeito dos fatores externos sobre o sistema lingüístico: maior quantidade de dados lingüísticos; conhecimento acerca dos parâmetros sociais e estilísticos que envolvem o contexto de fala; apreensão das relações que se estabelecem entre os padrões socioculturais e ideológicos e o fenômeno lingüístico em análise.

O **problema da avaliação** busca “*encontrar os correlatos subjetivos (ou latentes) das mudanças objetivas (ou manifestas) que foram observadas*”. São averiguações de quais atitudes são despertadas entre os falantes em relação a uma determinada mudança, já que “*o nível de consciência social é uma propriedade importante da mudança lingüística [...]*” (WEINREICH; LABOV; HERZOG, 2006 [1968], p. 124). Dessa maneira, procura responder em qual medida a avaliação subjetiva pode afetar o processo da mudança lingüística. Segundo estudos empíricos, a consciência e a avaliação da mudança manifestam-se em diferentes estágios: em um primeiro momento, não há uma consciência social da mudança, mas, em seguida, surgem padrões estilísticos e de estratificação social. A consciência social é um estágio mais avançado da mudança, de tal modo, as variantes tornam-se identificáveis socialmente e passam, por isso mesmo, a identificar o falante. Dessa forma, a valoração social das variantes e a avaliação subjetiva são capazes de implementar ou limitar um processo de mudança.

Mais especificamente, é necessário ressaltar que os modelos lingüísticos estão sujeitos à avaliação positiva ou negativa, apesar disso, esse juízo de valor não possui fundamentação lingüística, já que cientificamente todas as manifestações lingüísticas são legítimas na medida em que exercem sua função precípua de comunicar. A avaliação decorre da associação que se estabelece entre a posição que o falante ocupa na

esfera social e a variante linguística, por ele, empregada. Dessa maneira, a estigmatização de uma variante linguística é consequência de preconceitos de classe no seio da sociedade.

A avaliação orienta a mudança linguística em dois sentidos: a *orientação para o prestígio* (*prestígio aberto* ou *descoberto*), quando os modelos linguísticos empregados pelas classes socialmente mais favorecidas são tomados como modelo a ser seguido; e a *orientação para a identidade* (*prestígio encoberto*) quando atitudes linguísticas que não gozam de prestígio social assumem valor sociocultural definidor da identidade de grupo – o indivíduo utiliza essas variantes com o desígnio de fortalecer seus laços identitários com o seu grupo social. Consoante as formas sociais e a consciência dos falantes diante das formas linguísticas em variação, a mudança pode ser definida *como mudança vinda de cima*, condicionadas por fatores dotados de prestígio social em um nível pleno de consciência e que, normalmente, surge de eventos de fala mais monitorados; ou *mudança vinda de baixo*, estabelecida no vernáculo e se implementa abaixo do nível de consciência social.

Uma questão pertinente a ser salientada dentro do problema da avaliação são os fatores socioeconômicos e estilísticos que atuam no processo da mudança linguística. O estudo dos fenômenos denominados como *indicadores*, *marcadores* e *estereótipos* são primordiais para ampliar a compreensão do nível de consciência do falante ao avaliar a língua. Estabelecem-se da seguinte forma:

(i) *Indicadores* – acontece quando a variação estilística não interfere de modo significativo no uso das variantes pelas classes sociais, o falante avalia a língua de maneira menos consciente. Existe variação social, mas pouca variação estilística. Um exemplo de indicador linguístico é a realização do pronome sujeito no português brasileiro;

(ii) *Marcadores* – ocorre quando a variação estilística interfere no uso das variantes pelas diversas classes sociais. Os falantes avaliam a variável de forma mais consciente, desta maneira, uma variante ganha prestígio sem que a(s) outra(s) seja(m) abertamente rejeitada(s). A variação entre *nós* e *a gente* seria um exemplo de marcador, com o *nós* sendo a variante de prestígio sem que haja em muitos contextos uma avaliação negativa do uso do *a gente*.

(iii) *Estereótipos* – acontece quando ocorre uma avaliação social abertamente negativa de alguns usos linguísticos. Existe um aumento no nível de consciência dos falantes e a variante padrão é usada frequentemente nas situações mais formais por todas as classes sociais. A falta de concordância verbal constitui assim um estereótipo.

Sendo assim, é evidente que algumas variantes linguísticas usufruem de maior prestígio social do que outras. Geralmente, a variante usada pelo segmento social que detém maior poder econômico adquire *status* positivo, enquanto uma variante pertencente a um grupo social desprovido de poder econômico é normalmente estigmatizada.

Há ainda o problema da restrição e o problema da implementação. O primeiro, também conhecido como “**problema dos fatores condicionantes**”, diz respeito à busca de fatores que fazem possíveis ou impossíveis as mudanças, demarcando sua produção. Esse problema está relacionado aos condicionamentos linguísticos e extralinguísticos que intervêm na trajetória da mudança de acordo com o que o sistema permite. Entretanto, como bem assinala Lucchesi (2004), “essa questão conduz novamente a teoria da mudança à idéia de que as mudanças seguem princípios gerais, ou mesmo universais, como, por exemplo, *todas as mudanças conduzem à simplificação e à generalização de regras da gramática* (cf. Labov, 1982:26-7)”. Labov (1982), desta maneira, termina por reconhecer os equívocos provenientes de uma relação de tendências gerais das mudanças e propõe que o problema das restrições, seja amalgamado com o problema do encaixamento.

O **problema da implementação** procura explicar o porquê de uma mudança linguística ter se implementado em determinado espaço de tempo e em determinado lugar. Segundo Weinreich, Labov e Herzog (2006 [1968] p. 124) “*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*”. De acordo com a Teoria da Evolução, a mudança linguística inicia-se no momento em que uma das variantes que compõe “o campo de batalha” (TARALLO, 2005) adquire vantagem e se distribui gradualmente dentro da comunidade de fala através do tempo e do espaço. A mudança se processa primeiramente de forma lenta e a variante apresenta baixa frequência de uso. Mais tarde, o processo de mudança ganha velocidade e, posteriormente, desacelera e, por fim, completa-se a mudança quando a

forma linguística se torna categórica. Algumas pesquisas têm comprovado esse princípio teórico<sup>3</sup>.

Conforme Lucchesi (2004, p. 179), a percepção do objeto de estudo da linguística, na visão do modelo alvitrado por Weinreich, Labov e Herzog (2006 [1968]), baseia-se em “*uma organização complexa de mecanismos de causa e efeito, que envolve fatores sociais e lingüísticos*” que termina conflitando com a concepção de língua como um fenômeno sócio-histórico e cultural. Segundo o autor, uma solução para resolver essa contradição, seria conceber o objeto de estudo da linguística como um conjunto *complexo de relações de determinação*, configurado a partir da interação entre o processo de estruturação da língua e a realidade sócio-histórica em que ela está inserida.

Definir o fenômeno da variação e da mudança linguística não é tarefa das mais fáceis, provavelmente, porque esse processo abarca grande quantidade de fenômenos que não operam isoladamente. A proposta de análise sistemática da mudança linguística e, por conseguinte, da própria língua, na visão da Sociolinguística é a de procurar identificar, de forma precisa, o ponto em que se verifica a covariação entre fatores lingüísticos e sociais, não o momento em que a mudança surge, mas o momento de sua propagação na estrutura social da comunidade de fala.

Uma forma de explicar os processos de variação e mudança, encontrada pela Sociolinguística, é distinguir um estado de variação estável do processo de mudança em curso.

### **1.1. 3 Mudança em tempo aparente e em tempo real**

Um dos propósitos da análise Sociolinguística, apresentados nos estudos de Labov e, já anteriormente, esboçados nas investigações de Gauchat, consiste em superar a ideia de que a mudança linguística não podia ser observada diretamente, mas somente após a sua consumação. O pioneirismo dos estudos de Gauchat reconhecido por Labov contribuiu para o propósito da análise da Sociolinguística de reconhecer, a partir da

---

<sup>3</sup> Como a pesquisa realizada por Oliveira e Silva (1982 *apud* NARO, 2004, p. 1982) acerca do uso do artigo definido diante de pronomes possessivos na língua portuguesa.

observação de determinada comunidade de fala, num dado momento da variação, com base nas diferenças no comportamento linguístico de gerações distintas, um processo de mudança em curso.

As teorias estruturais da língua faziam a oposição entre um estudo sincrônico de outro diacrônico. Entretanto, a investigação da Sociolinguística Variacionista contesta esta incompatibilidade assumindo duas representações – o estudo em *tempo aparente* e o estudo em *tempo real*.

O estudo da mudança numa análise sincrônica abriu espaço para a superação da dicotomia saussuriana entre sincronia e diacronia. Para a Sociolinguística, a mudança pode ser diretamente estudada através do processo de variação, pois a variação é, em determinado momento, um reflexo da mudança. Sendo assim, é possível observar processos de mudança em progresso.

Em um recorte sincrônico apreende-se a mudança linguística em tempo aparente – o que se verifica, na realidade, é uma projeção diacrônica tanto para o passado quanto para o futuro da variação sincrônica observada a partir da variável faixa etária e da sua relação com o processo de variação e mudança. Por outro lado, em uma abordagem diacrônica, em tempo real, com a observação da comunidade de fala em dois momentos distintos, pode-se observar aspectos da mudança que não ficam claros na abordagem em tempo aparente:

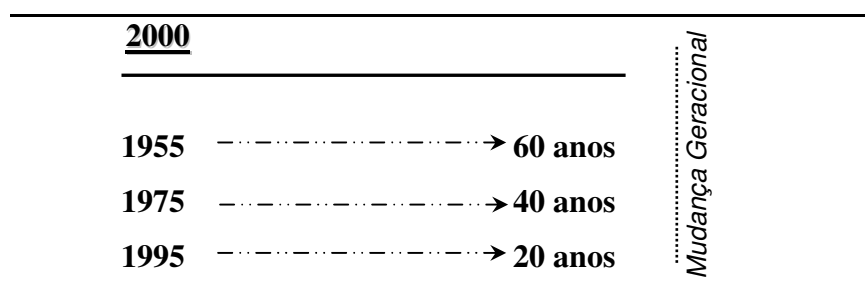
O estudo da mudança em tempo real (de curta ou longa duração) permite recobrir aspectos, que não seriam detectados pelo estudo em tempo aparente, distinguindo mudanças que se produzem de forma gradual em toda comunidade linguística daquelas que podem caracterizar a trajetória do comportamento linguístico individual ao longo da sua vida. (PAIVA E DUARTE, 2003, p. 14).

Contudo, a abordagem aparente é aquela que tem predominado nas pesquisas da Sociolinguística Variacionista. Essa abordagem consiste em escolher uma comunidade e confrontar a fala dos indivíduos mais velhos com a fala dos indivíduos mais jovens, assumindo que as diferenças encontradas refletem estágios diferentes de uma mudança linguística. De acordo com a hipótese clássica, “*o estado atual da língua de um falante adulto reflete o estado da língua adquirida quando o falante tinha aproximadamente 15 anos de idade*” (NARO, 2004, p. 15). Assim, quando o indivíduo chega à adolescência, fixa sua gramática e, por conseguinte, sua conduta linguística permanece constante.



[...] a validade do [tempo aparente] depende crucialmente da hipótese de que a fala das pessoas de 40 anos hoje reflete diretamente a fala das pessoas de 20 anos há 20 anos atrás e pode, portanto, ser comparada com a fala das pessoas de 20 anos de hoje, para uma pesquisa da difusão da mudança lingüística. As discrepâncias entre a fala das pessoas de 40 e 20 anos são atribuídas ao progresso da inovação lingüística nos vinte anos que separam os dois grupos (CHAMBERS e TRUDGILL, 1980, p.165).

Podemos inferir que uma fala gravada no ano de 2000 de um indivíduo de 60 anos representaria o estado de língua adquirida em 1955.



QUADRO 1: Gradação etária segundo a hipótese clássica.

A hipótese clássica apresenta alguns problemas, pois ainda não se tem comprovações de que o indivíduo mantém seu comportamento estável a partir da adolescência. Segundo Lucchesi (2001), a confirmação dessa estabilidade é discutível e pouco provável de ser confirmada, já que dependerá da variável em estudo, pois algumas estruturas são mais flexíveis que outras.

O estudo da mudança em tempo real está relacionado à mudança lingüística num período do tempo – o aspecto diacrônico da língua. Segundo Labov (2004), há dois métodos de se fazer observações em tempo real: no estudo de tendência e no estudo de painel.

No estudo de tendência (*trendy study*), temos a comparação de grupos distintos de falantes, enquanto que no estudo de painel (*panel study*), verifica-se o comportamento dos mesmos indivíduos em diferentes períodos do tempo. De acordo com Labov (1994), a combinação de observações em tempo aparente e em tempo real é a melhor maneira de analisar se estamos diante de uma mudança em curso. Assim, podemos formar a seguinte correlação:

	<i>a. Estabilidade</i>	<i>b. Gradação por idade</i>	<i>c. Mudança geracional</i>	<i>d. Mudança na comunidade</i>
<i>1. Indivíduo</i>	Estável	Instável	Estável	Instável
<i>2. Comunidade</i>	Estável	Estável	Instável	Instável

QUADRO 2: Relações entre tempo aparente x tempo real.

Na relação estabelecida em 1a e 2a o indivíduo e a comunidade não apresentam alterações durante o período de tempo observado, o que corresponde a um padrão estável. O padrão instável do indivíduo em 1b e o padrão estável na comunidade em 2b significa dizer que o indivíduo muda seu comportamento linguístico ao longo da vida, mas a comunidade continua estável – corresponde a gradação etária, como bem salienta Labov (1994, p. 84 *apud* MONTEIRO, 2000): “*Se os indivíduos modificarem seus hábitos linguísticos através de suas vidas, mas a comunidade como um todo não muda, caracteriza-se então um padrão próprio de uma faixa etária*”. A relação e 1c e 2c indica que estamos diante de uma mudança geracional – o indivíduo se mantém constante em sua conduta linguística, mas a comunidade muda com as novas gerações. 1d e 2d acontece quando tanto o indivíduo quanto a comunidade mudam o comportamento linguístico ou adquirem novos elementos linguísticos juntos.

Todas as situações observadas acima permitem que se façam asseverações mais seguras, no entanto, não promovem interpretações objetivas sobre a mudança. É importante salientar que além de entender e sistematizar as mudanças que incidem na estrutura da língua, do conhecimento da constituição histórica e funcionamento interno desta, faz-se necessário instituir uma conexão com as complexas relações intrínsecas na sociedade.

A apreciação variacionista proporciona um valioso subsídio para um conhecimento mais aprofundado e fundamentado empiricamente com relação à língua e, conseqüentemente, envolve a compreensão dos processos de variação e mudança linguística, na medida em que descreve, numa perspectiva social, padrões coletivos de comportamento linguístico decorrentes da covariação sistemática entre fatores linguísticos e sociais.

Há de se destacar que, o português falado no Brasil é um campo fértil para os estudos sociolinguísticos. As relações de contato entre povos que divergiam linguística e socialmente no período colonial nos leva a afirmar que a representação do português brasileiro tem sua origem no percurso de sua formação. Posto isso, a próxima seção traz

a reflexão acerca do processo de aquisição precária de uma língua segunda em situações adversas por indivíduos adultos, situação muito comum nos primeiros séculos da formação da sociedade brasileira, em virtude do amplo contato de colonos de origem europeia com índios aculturados e africanos escravizados. O conceito de *transmissão linguística irregular* (LUCCHESI, 2003) que será apresentado na próxima seção fundamentará a análise do comportamento da variável artigo definido nos SNs em comunidades do português popular no interior do estado da Bahia, que se apresenta nesta dissertação.

## 1.2 A TRANSMISSÃO LINGUÍSTICA IRREGULAR

A história das línguas humanas está repleta de contatos linguísticos nas mais diferentes condições sócio-históricas. Para a situação que nos interessa neste estudo, traçaremos as circunstâncias de contato entre línguas em um tipo de sistema agrícola baseado na monocultura de exportação mediante a utilização de latifúndios e de mão-de-obra escrava – a *plantation*. É de fundamental importância entender como o contato massivo entre populações de culturas distintas e que falam línguas ininteligíveis entre si pôde desencadear o que chamamos de **língua pidgin** e **língua crioula**, a fim de compreender o porquê do português no Brasil ter desencadeado, não uma língua crioula, mas uma **variedade em sua vertente popular**.

Em sua acepção primeira, é preciso destacar, que em um contexto de multilinguístico, a necessidade de uma comunicação emergencial torna-se imprescindível e termina por ocasionar adaptações na língua de seus interlocutores. Essas adaptações são comuns em contextos prototípicos de *plantation* e desencadeiam reduções e/ou simplificações e, até mesmo, a reestruturação gramatical da língua alvo.

Segundo Lucchesi (2000, p. 99), a variedade dessa língua segunda, formada precariamente, apresenta uma forte erosão em sua estrutura gramatical em função dos seguintes fatores:

- (i) o difícil acesso dos falantes das outras línguas aos modelos da língua alvo, sobretudo nas situações em que os falantes dessa língua alvo são numericamente muito inferiores aos falantes das outras línguas;

- (ii) o fato de os falantes dessas outras línguas serem, em sua grande maioria, adultos, não havendo, pois, acesso aos dispositivos da *faculté du langage*, que atuam naturalmente no processo de aquisição da língua materna;
- (iii) a ausência de uma ação normatizadora, ou seja, de uma norma ideal que oriente e restrinja o processo de aquisição/nativização, já que esse processo tem como objetivo fundamentalmente a comunicação emergencial com os falantes da língua alvo.

Num primeiro momento, o código emergente comum entre sobrepujados e dominadores tem como característica principal possuir um vocabulário bastante restrito relacionado à troca de informações básicas no sistema de *plantation*, que giram em torno de dar e exercer ordens. Denomina-se **jargão** a esse código emergencial de vocabulário compactado em que se predomina o léxico do grupo dominante e as regras gramaticais dos escravizados.

O jargão funcionava porque os escravos eram misturados etnicamente e se viam forçados a adquirir uma segunda língua (L2) para manter as relações, ainda que em alguns casos, mínimas entre si. Contudo, posteriormente, seus descendentes, que adquiriam como língua materna esse código emergencial, ampliavam a sua estrutura gramatical. Portanto, se num primeiro momento, há erosão gramatical da língua alvo, com a **nativização** e a **socialização** ocorre a reestruturação da gramática, ou seja, há o preenchimento de lacunas estruturais decorrente da aquisição precária da língua dominante.

Assim sendo, com a continuidade da situação de contato e a nativização dessa segunda língua que se mostra mais viável dentro das relações sociais da nova comunidade de fala, teremos o que a terminologia do campo designa como **crioulização**. A língua crioula resulta, portanto, da nativização do jargão ou do **pidgin**, este último também resultante do desenvolvimento gramatical do jargão, só que com base nos mecanismos gramaticais da(s) língua(s) nativa(s) dos falantes adultos do(s) grupo(s) dominado(s). Assim, a língua pidgin é caracterizada pela ampliação comunicativa e maior reestruturação gramatical em relação ao jargão.

Na pidginização não acontece a nativização da variedade linguística no contexto de contato. Quanto maior a reestruturação linguística, sem a nativização, maior será a possibilidade de diferentes graus de pidgins, assim temos o pré-pidgin, pidgin ou pidgin expandido. Um exemplo, embora raro, de pidgin expandido foi o que se sucedeu em Papua Nova Guiné, conhecido popularmente como *tok pisin*, uma das línguas nacionais dessa ilha localizada na Oceania.

A nativização, desta maneira, é condição essencial para a formação das línguas crioulas típicas. Na crioulação a criança não vai ter como modelo a língua materna de seus pais, nem a língua alvo, do dominador. Nesta situação a tarefa da criança consistirá em organizar uma gramática plena, preenchendo as lacunas que existem na estrutura gramatical de qualquer das fases (jargão, pré-pidgin, pidgin ou pidgin expandido) que ela teve como exemplo.

A principal característica de línguas pidgins e línguas crioulas é que ambas têm a tendência de ter uma gramática diferente da gramática da língua alvo. Praticamente toda estrutura gramatical da língua alvo falada como segunda língua é reduzida e, somente com o tempo, os indivíduos vão gramaticalizando o código emergencial, estabelecendo convenções sociais para a formulação de frases, morfemas, advérbios, orações, etc.

No nascimento de línguas crioulas observamos: (i) perda de morfemas de gênero, pessoa e número; (ii) raramente se mantêm a morfologia derivacional e a flexão pronominal de gênero e caso; e (iii) a perda da maioria das preposições. Verificamos ainda a reconstituição de elementos que se perderam no contato inicial que se dá, geralmente, pelo processo de gramaticalização. De todo modo, é importante salientar que a reestruturação gramatical está baseada numa lógica de funcionalidade e, provavelmente, por essa razão, existem itens lexicais mais conservados, enquanto que os itens gramaticais são os mais suprimidos e/ou substituídos (LUCCHESI, 2003).

O processo de constituição de línguas crioulas nos fornece elementos precisos para estabelecer a diferença entre o que vem a ser uma **transmissão linguística regular** e uma **transmissão linguística irregular**. No primeiro caso, a língua que as crianças têm como modelo para a aquisição é a língua materna de seus pais e não uma língua segunda aprendida em situação adversa. No segundo caso temos o oposto: é a língua adquirida em situação adversa, a L2, que servirá de modelo para as crianças.

Sendo assim, para entender estes mecanismos, é imperioso estabelecer que a maior ou menor interferência das línguas dos falantes dominados na formação de uma nova língua depende muito do contexto histórico em que se deram. Conforme salienta Lucchesi (2000, p. 109):

O fator que determina inicialmente o processo de transmissão linguística irregular é a intensidade da erosão gramatical que se dá no momento inicial do contato, ou seja, no momento em que os falantes do substrato fixam o código emergencial de comunicação verbal a partir dos recursos léxico-gramaticais da língua alvo.

Desta forma, o amplo e complexo processo que se desenrola no Brasil no período de implantação e consolidação do escravismo, a par do crescente fluxo de africanos de várias etnias, retirados de seu território sob intensa sujeição, assiste a um processo de aquisição linguística que “*constitui um contínuo de níveis diferenciados de socialização/nativização de uma língua segunda, adquirida massivamente, de forma mais ou menos imperfeita, em contextos sócio-históricos específicos*”. (LUCCHESI, 2000). Mas por que não se solidificou no Brasil uma língua crioula?

O período inicial do contato reproduzia-se distintas vezes nos quatro séculos de escravidão no amplo território brasileiro, caracterizando, deste modo, um processo de transmissão linguística irregular mais leve do que nos casos de pidginização e criouliização. Além disso, operaram-se no Brasil colonial diversos motivos que frearam o desenvolvimento de uma língua crioula típica, como assinala Lucchesi (2009, p. 70)<sup>4</sup>:

- i. *a proporção entre a população de origem africana e branca, que propiciou um maior acesso à língua alvo do que o observado nas situações típicas de criouliização;*
- ii. *a ausência de vida social e familiar entre as populações de escravos, provocada pelas condições sub-humanas de sua exploração, pela alta taxa de mortalidade e pelos sucessivos deslocamentos;*
- iii. *o uso de línguas francas africanas como instrumento de interação dos escravos segregados e foragidos;*
- iv. *o incentivo à proficiência do português;*
- v. *a maior integração social dos escravos urbanos, domésticos e das zonas mineradoras;*
- vi. *a miscigenação racial.*

Adotamos como hipótese explicativa para as variações observadas no PB têm em sua gênese uma aprendizagem imperfeita da língua pelos africanos: “*Um processo de transmissão irregular de L2 para L1 em que a L2 foi alterada devido a problemas de*

<sup>4</sup> No decorrer do capítulo 2, abordaremos mais amplamente os fatores inibidores de um processo generalizado de criouliização no Brasil, listados pelo autor.

*acesso à língua alvo (isto é, a língua do grupo dominante) e, possivelmente, à influência das línguas maternas dos falantes desta L2”* (BAXTER; LUCCHESI, 1997, p. 74). A variação de uso do artigo definido em SNs em uma comunidade rural afro-brasileira isolada do interior do estado da Bahia será analisada nesta dissertação com base nessa hipótese que considera a relação histórica de contato um fator proeminente na formação dessa variedade do português brasileiro.

### 1.3 METODOLOGIA

#### 1.3.1 Seleção e caracterização do *corpus*

Este trabalho está vinculado ao Projeto *Vertentes do Português Popular do Estado da Bahia*, coordenado há nove anos pelo Professor Doutor Dante Lucchesi. As pesquisas advindas do Projeto *Vertentes* têm contribuído para o esclarecimento acerca da língua portuguesa no Brasil e fornecido afirmações consistentes sobre a importância do contato massivo entre línguas africanas na formação do português brasileiro.

Objetivando a análise do uso variável do artigo definido nos SNs, o universo de nossa amostra é constituído pela fala vernácula da comunidade de Helvécia. Esta localidade, representante do português mais marcado etnicamente, pertence ao *corpus* do português afro-brasileiro do interior do estado da Bahia. São consideradas comunidades rurais afro-brasileiras pelo Projeto *Vertentes* as comunidades originadas de antigos quilombos e que seguem, segundo Lucchesi (2009, p. 75), os seguintes parâmetros:

- (i) são compostas majoritariamente por descendentes diretos de escravos africanos que se fixaram em localidades remotas do interior do país e de difícil acesso;
- (ii) mantiveram-se em relativo isolamento até a segunda metade do século XX;
- (iii) sua principal atividade econômica é a agricultura de subsistência.

Selecionaram-se, portanto, 12 inquéritos de *Helvécia*, seguindo o critério para a seleção dos informantes estabelecido pelo Projeto Vertentes:

- a. ter nascido e residir na comunidade;
- b. informantes com pouca ou nenhuma escolaridade;
- c. faixas etárias distintas, sendo instituídas quatro faixas: I (20-40 anos); II (41-60 anos); III (61 a 80 anos) e IV (mais de 80 anos);
- d. ambos os sexos.
- e. informantes que saíram da comunidade ou que nunca se ausentaram.

A tabela 1 mostra a divisão dos informantes por sexo e por faixa etária.

TABELA 1: Perfil social dos inquéritos selecionados para a realização da pesquisa – *corpus* do português afro-brasileiro.

Fatores sociais Localidade	Sexo	Faixa etária				Total
		20 a 40	41 a 60	61 a 80	+ de 80	
<i>Helvécia</i>	M	2	2	1	1	12
	F	2	2	1	1	

Para a coleta das amostras de fala, os pesquisadores do projeto Vertentes realizaram entrevistas que possam representar o vernáculo mais espontâneo do informante. Desta maneira, procurou-se suavizar o *paradoxo do observador*<sup>5</sup>, seguindo os procedimentos estabelecidos por Labov (1972) e adaptados à realidade das comunidades, como: (i) adequação, por parte dos inquiridores, das marcas do dialeto local, diminuindo a distância entre os interlocutores; (ii) não revelar que o motivo da entrevista é o estudo da fala da comunidade, mas se apresentar como alguém que deseja “conhecer a vida do lugar”, a história da comunidade; (iii) escolher temas que fazem parte das experiências de vida dos informantes, estimulando assim uma narração com maior nível de naturalidade. Os temas das entrevistas, que possuem aproximadamente a

<sup>5</sup> Problema metodológico referente aos meios empregados na recolha da fala e que acabam por interferir, às vezes, nos próprios dados.



duração de uma hora, ancoram-se na observação feita pelo pesquisador sobre a história da comunidade e giram em torno dos costumes antigos, da infância, das doenças mais frequentes e da morte de familiares, as festas locais, a lavoura, trabalho, relatos da vida conjugal, religião, política, etc.

Todas as entrevistas que constituem os *corpora* do projeto Vertentes foram digitalizadas com equipamentos e programas de áudio capazes de eliminar qualquer ruído ou barulho que pudessem interferir na inteligibilidade das mesmas. As entrevistas editadas foram transcritas grafematicamente, seguindo a Chave de Transcrição do Projeto Vertentes e conservando as especificidades da fala dos informantes. Ressalta-se que, para maior segurança das informações, todas as entrevistas passam ainda por mais duas etapas: (i) uma revisão, a fim de verificar possíveis erros de transcrição; e (ii) uma revisão de leitura que termina dando os últimos ajustes de pontuação para maior adequação de um texto de origem oral para um texto escrito.

### **1.3.2 Procedimentos da pesquisa**

Com o propósito de buscar amostras naturais de fala vernácula com maior espontaneidade, desconsideraram-se as 10 primeiras páginas transcritas de cada entrevista. Além disso, com o desígnio de se ter maior confiabilidade dos dados coletados cumpriu-se alguns critérios metodológicos desde a coleta de dados até a análise dos resultados, listados a seguir:

1. Selecionado o *corpus*, passou-se ao levantamento exaustivo das ocorrências das variáveis dependentes;
2. Audição de todos os 12 inquéritos para averiguar a veracidade dos dados levantados, evitando que uma possível falha de transcrição fosse considerada;
3. Leitura de textos especializados sobre o fenômeno em foco durante todo o processo da pesquisa;

4. Estabelecimento das variáveis linguísticas que condicionariam a presença ou a ausência do artigo definido em SNs e posterior elaboração de uma chave de codificação;
5. Consignação de uma cota de 50 a 55 ocorrências por inquérito devido à abundância de ocorrências válidas;
6. Codificação dos dados levantados;
7. Processo de refinamento dos dados, estabelecendo, em seguida, os casos que seriam descartados;
8. Quantificação dos dados com a aplicação do pacote de programas VARBRUL, com o objetivo de estabelecer os grupos de fatores significativos;
9. Análise e interpretação dos resultados obtidos.

### 1.3.3 Tratamento estatístico dos dados

O VARBRUL é um pacote de programas estatístico concebido com a finalidade de se observar fenômenos variáveis. Sendo assim, foram implementados modelos matemáticos que procuram dar esse tratamento a dados linguísticos que co-ocorrem. Os programas existentes municiam frequências dos diversos grupos de fatores em termos percentuais, selecionando os grupos com importância estatística. Além disso, o pacote de programas VARBRUL provê os pesos relativos sobre a variável analisada numa escala de 0 a 99. Os valores acima de .50 favorecem a realização da variante estudada, enquanto os valores abaixo de .50 desfavorecem, mostrando sua neutralidade quando obtém .50 de frequência.

O primeiro programa que compõe o pacote estatístico computacional é o **CHECKTOK**, que tem como função comparar os símbolos digitados na cadeia de codificação do **arquivo de dados** com o **arquivo de especificação** de fatores. É neste programa que o pesquisador decide a ordem em que os grupos de fatores deverão

aparecer, além do conjunto de símbolos válidos para cada grupo. É através do **editor de texto** que o pesquisador cria os arquivos de dados e de especificação, ou seja, a codificação feita através de uma chave pré-estabelecida precisa ser digitada nesse programa específico.

O arquivo de **especificação de fatores** é usado como entrada opcional para o CHEKTOK. Possui as seguintes informações: número de grupos de fatores, caractere que deverá completar as cadeias de codificação e os fatores válidos e inválidos.

O **CHEKTOK** é o responsável pela correção dos dados codificados manualmente, gerando um arquivo de dados devidamente corrigido; “verifica os erros grosseiros de codificação”. Os arquivos avaliados originados pelo CHECKTOK serão empregados como entrada para o passo seguinte.

O **READTOK** é o programa que reúne todas as cadeias de codificação para o processamento quantitativo. Se não houver erros, como vimos anteriormente, o CHEKTOK cria um arquivo corrigido (Ex: conc.cor). A função do READTOK, então, é ler todas as ocorrências de um ou mais arquivos de dados e escrevê-las em um arquivo de ocorrências sem o contexto e o parêntese da primeira coluna. O primeiro arquivo pedido pelo READTOK é o arquivo de saída.

O arquivo de ocorrências criado pelo READTOK é utilizado como entrada para o próximo programa, o MAKECELL/MAKE3000, e, posteriormente, pelo COUNTUP.

O **MAKECELL** recebe o arquivo gerado pelo READTOK e prepara os dados a serem executados para os programas IVARB, MVARB, TVARB e CROSSTAB, agrupando os dados em células. Para isso, é criado um **arquivo de condições**, em que é especificada a variável dependente e estabelecidos os parâmetros de aceitação, o reagrupamento e a exclusão de dados. O arquivo gera as frequências em termos percentuais.

Existem alguns tipos de problemas relacionados à inconformidade dos resultados. Um *knockout* ocorre quando um fator dentro de um grupo tem um comportamento não-variável em relação às variantes em questão; o *singleton* indica a ocorrência de um único fator dentro de um grupo de fatores. Esse arquivo só poderá ser usado como entrada para o **CROSSTAB**.

O **IVARB** é o programa que analisa os pesos relativos em forma binária (variáveis dependentes binárias), mostrando a interferência específica de cada fator e isolando outros. O **TVARB** e o **MVARB** obtêm os mesmos cálculos do **IVARB**, porém

diferenciam-se quanto à quantidade de variantes: o primeiro para três variantes e o segundo quatro ou cinco variantes.

O **TSORT** e o **TEXTSORT** são programas de busca e de re-ordenação dos dados. O primeiro efetua a procura em função da cadeia de codificação (aqui se pode ter em um único arquivo todas as ocorrências que possuam determinado símbolo em determinada coluna); já o segundo faz a procura em função de informações digitadas (possibilita a procura de cadeias de contexto, exemplo: caso o analista almeje ter todos os casos do verbo *estar* em sua forma *estão*).

O **CROSS 3000** é o programa responsável pelos cruzamentos das percentagens atribuídas a dois grupos de fatores.

Por fim, o **COUNTUP** que fornece a frequência geral dos dados por variável e por fator dentro de cada variável – oferece uma visão global da distribuição dos dados.

O pacote de programas estatístico computacional é uma ferramenta importante, pois permite calcular probabilidades de frequência da variante adotada como referência para análise e o efeito relativo de cada fator no fenômeno variável de forma precisa. Contudo, como bem salientam Callou, Omena e Silva (1991, p.20), “*os números não valem por si mas funcionam como ponto de referência para a interpretação*”. Guy (2007) afirma que:

O Varbrul é uma ferramenta poderosa e extremamente útil para a análise da variação linguística. Como acontece com qualquer ferramenta, sua utilidade é acentuada por uma compreensão de suas operações e de suas limitações. Mas não se deve nunca perder de vista o fato de que, na análise final, o Varbrul é apenas um recurso (embora sofisticado) para a manipulação dos dados. Não discerne padrões, não faz generalizações, nem explica achados.

Desta maneira, é essencial que o pesquisador tenha um conhecimento profundo dos condicionamentos linguísticos e extralinguísticos do fenômeno analisado, privilegiando, deste modo, o qualitativo em detrimento do quantitativo.

#### 1.4 A VARIÁVEL DEPENDENTE

O estudo desta dissertação volta-se para o uso variável do artigo definido em sintagmas nominais (SNs). O fenômeno que escolhemos despertou nossa atenção a

partir de um fato linguístico característico da comunidade de fala aqui estudada: a falta de artigo definido em contexto de referencial definido. Uma construção como *Io sabe Ø dia do ano*<sup>6</sup>, em que o substantivo *dia* é um referente definido, ou seja, está inventariado ao identificável e ao traço (+ específico), e não está acompanhado do artigo definido, também é verificada, 20 anos após a primeira abordagem, realizada por Ferreira (1961), em diferentes contextos no vernáculo do português afro-brasileiro. O estudo mais particularizado deste fenômeno motivou-nos a aprofundar nossa observação, a fim de trazer novas evidências empíricas para relacionar o contato linguístico de diferentes etnias no Brasil colonial e a atual realidade do PB, inserido na hipótese explicativa da *transmissão linguística irregular*.

Sendo assim, nossa variável dependente é binária e avalia a variação no uso do artigo definido em SNs.

### 1. Presença de artigo definido

(1) ... *a mãe de Ana Maria* morreu (HV-12)

### 2. Ausência de artigo definido

(2) *É mais perigoso, porque Ø mulé* pá ganhá neném... (HV04)

#### 1.4.1 Casos desconsiderados na pesquisa

Como mencionado no item 1.1.1, os dados recolhidos e codificados passaram por um processo de refinamento, com o objetivo de descartar dados ou por serem categóricos ou por não caberem nesta pesquisa. Faz-se necessário uma observação minuciosa dos dados para adotar essa medida, pois a não exclusão de alguns contextos poderiam enviesar os resultados. Os casos descartados, apresentados a seguir, foram excluídos da quantificação e da análise.

<sup>6</sup> Exemplo retirado de Ferreira, em 1961, de uma informante de Helvécia-BA, do sexo feminino, com 75 anos.

‣ ***Antropônimos, topônimos e possessivos***

Não obstante haja variação em relação ao uso do artigo definido diante de antropônimos, topônimos e possessivos não constitui o objetivo de nossa pesquisa analisá-los. Também foram rejeitados casos de artigo definido diante de apelidos, nomes de empresas, siglas, bem como de pronomes de tratamento e nos casos de nomeação de santos católicos como *Nossa Senhora*, *Nosso Senhor*, etc.

- (3) *Lá no... no Rio? É ... perigoso Rio de Janeiro.* (HV 07)
- (4) *... que você, sem sua mãe, 'cê não tem nada na vida.* (HV 07)
- (5) *Foi a minha irmã que crio.* (HV 07)
- (6) *... “Miúda não tá aí não?”* (HV 07)
- (7) *É, Oswaldo é gente boa.* (HV 04)
- (8) *O Nosso Senhor sabe.* (HV19)
- (9) *Graças a Nossa Senhora.* (HV 19)

‣ ***Expressões cristalizadas***

Foram desconsiderados os casos em que há expressões cristalizadas, já que em grande parte delas não há presença do artigo.

- (10) *E vô colocá fogo nele, hoje pra mim roçá o lugá.* (HV02)

‣ ***Construções com SNs preposicionados***

- (11) *Eu saí da escola* (HV01)
- (12) *Quano dava de noite qu'eu sentava... pronto.* (HV 01)

‣ ***Sentenças interrompidas ou incompletas***

- (13) *...que o.... povo... povo.... fura aqui... injeção...* (HV10)

‣ *Casos com a preposição em*

(14) Nem precisa conzinhá **em casa** nem nada, né, aí gente vai! (HV02)

(15) É... aqui **em casa** mesmo é difícil fazê. (HV01)

‣ *Vocativos*

(16) **Menina**, volte aqui! (CZ 06)

‣ *Partitivos e crases*

(17) *Polícia só anda **toda hora** lá e eu tá com medo!* (HV19)

‣ *Situações com o verbo ter em que a ausência de artigo atribui quantificação*

(18) Maria **tem medo**.

#### 1.4.2 Variáveis linguísticas e extra-linguísticas analisadas

Após a leitura de textos especializados sobre a utilização do artigo definido em sintagmas nominais, durante todo o processo da pesquisa, elaboramos uma chave de codificação com os fatores que, possivelmente, condicionariam a realização das variáveis dependentes. Perfazem 14 grupos de fatores<sup>7</sup> submetidos ao programa de pacotes VARBRUL, a saber:

‣ *Variáveis linguísticas*

1. Nível de referencialidade;
2. Número do SN;
3. Gênero do núcleo do SN;

<sup>7</sup> Cada grupo de fator será elucidado, juntamente com os resultados obtidos e interpretados, no capítulo 4. A exceção será a variável comunidade, explanada no capítulo 2.

4. Animacidade do núcleo do SN;
5. Tipo de núcleo do SN (se concreto ou abstrato);
6. Presença de modificadores pós-nucleares;
7. Contabilidade do núcleo do SN;
8. Determinante na menção anterior;
9. Estatuto informacional;
10. Função sintática do SN.

‣ ***Variáveis extra-linguísticas***

11. Faixa etária;
12. Sexo;
13. Estada fora da comunidade;
14. Nível de escolaridade.



## 2 HISTÓRIA SOCIOLINGUÍSTICA DO BRASIL

Como salientado no capítulo anterior, nos processos de variação e mudança, além de existirem interferências internas à língua, ocorrem fatores sócio-históricos que favorecem de modo proeminente a caracterização de uma comunidade linguística.

O português do Brasil (PB) apresenta certas especificidades em função de razões sócio-históricas distintas daquelas encontradas em outras terras colonizadas pelos portugueses. Essas características linguísticas do PB têm sido frequentemente estudadas, contribuindo de maneira significativa para o conhecimento da sua formação.

Com o intuito de demonstrar como se delineou a realidade linguística do Brasil, partiremos de alguns eventos históricos cruciais – desde a colonização até o português brasileiro contemporâneo – de modo a favorecer uma melhor compreensão sobre o contato linguístico entre índios, negros e brancos e, assim, elucidar o fenômeno do uso artigo definido em sintagmas nominais (SNs) e a hipótese da *transmissão linguística irregular*, em que se pauta este trabalho.

### 2.1 AS ORIGENS DO PORTUGUÊS BRASILEIRO

No século XVI, o Brasil foi descoberto por uma frota portuguesa de exploração do Atlântico e, em meados desse século, começou a ocupação e a colonização do litoral, o que mais tarde tornou o Brasil grande colônia portuguesa.

Conjectura-se que viviam cerca de seis milhões de índios nas terras recém descobertas pelos navegantes lusitanos e, de acordo com Rodrigues (1996), eram faladas, aproximadamente, 1.175 línguas indígenas<sup>8</sup> pertencentes a troncos linguísticos díspares. Implica-se, desta forma, que já havia um cenário multilinguístico em terras brasileiras antes mesmo da colonização portuguesa. Atualmente, remanescem 180 línguas indígenas pertencentes a dois troncos linguísticos – Tupi e Macro-Jê –

---

<sup>8</sup> Franchetto e Leite (2006) fazem referência a 1.275 línguas, espalhadas ao longo do vasto território brasileiro na época da descoberta do continente sul-americano.

ressaltando-se que com maior relevância o Tupi<sup>9</sup> manteve principal posição no contato com o idioma português. A relação das línguas indígenas com a língua portuguesa constitui de fato um tópico bastante complexo na história da sociedade brasileira, e uma das questões mais intrigantes no âmbito dessa relação refere-se à política da **língua geral**.

A utilização das línguas gerais foi a maneira que os portugueses encontraram para manter uma comunicação com os autóctones sobrepujados. A facilidade dessa comunicação decorria da relativa uniformidade que havia entre as línguas indígenas da costa, que, em geral, pertenciam ao tronco Tupi. O conceito de língua geral, no entanto, ainda passa por algumas contradições em meio a estudiosos do assunto, contudo, a proposta de Lucchesi (2000) alvitra as seguintes situações linguísticas:

- (i) uma koiné empregada na comunicação entre as tribos de línguas do tronco tupi da costa brasileira;
- (ii) versão como língua franca usada no intercuro dos colonizadores portugueses e indígenas;
- (iii) versão nativizada predominante nos núcleos populacionais mestiços que se estabeleceram no período inicial da colonização; e
- (iv) versão “gramaticalizada” pelos jesuítas sob o modelo do português e utilizada largamente na catequese, até de tribos de língua não tupi – chamados de tapuias, que significa ‘bárbaro’, em tupi.

A propósito dos interesses dos colonizadores, convém observar que pertence aos missionários jesuítas à figura dos principais estudiosos da língua indígena na época, pois, para esses missionários da fé católica, a língua era instrumento fundamental para a catequese. Logo instalaram “um tupi jesuítico” (MATTOS E SILVA, 2004), que posteriormente serviu de base para a língua geral da costa. As línguas gerais passaram a ser um veículo amplamente usado na intercomunicação entre índios, negros e brancos e chegaram a ter *status* de oponentes à hegemonia do português em território brasileiro.

Podemos diferenciar, basicamente, as línguas gerais de acordo com o eixo circunstancial da miscigenação em *língua geral paulista* (LGP) e *língua geral amazônica* (LGA) (RODRIGUES, 2006). A LGP difundiu-se no sudeste e foi falada na região de São Paulo até o início do século XX, contudo, não há um conhecimento mais aprofundado dessa língua geral de base Tupi originada mais ao Sul do Brasil. Segundo

<sup>9</sup> O tronco Tupi juntamente com o guarani (grupo de dialetos aparentados) formou o tupi-guarani, transição efetuada ao longo da área fronteira de São Paulo e do atual estado do Paraná (RODRIGUES, 1996).

Mussa (1991), a LGP entrou em declínio com a “decadência da atividade bandeirante e ao influxo imigratório português”. A LGA, também conhecida como *nheengatu* (fala boa) difundiu-se nos estados do Maranhão, Pará e Amazonas e, ainda hoje, há sinais dessa língua de base Tupinambá no médio e alto Rio Negro (Taylor, 1985 *apud* Noll, 2008).

De acordo com Rodrigues (1996), havia uma língua geral de base guarani que se estendia do oeste do Paraná até o Paraguai (hoje em dia, é a base do guarani falado neste país ao lado do espanhol), entretanto, no século XVII a região brasileira onde se falava essa língua geral foi devastada pelos bandeirantes. É oportuno ressaltar também que, há resquícios de uma língua geral no sul da Bahia (LOBO, MACHADO FILHO e MATTOS E SILVA, 2006), não obstante, estudos ainda são necessários para uma maior compreensão da difusão dessa língua no município de Ilhéus, na Bahia, e a confirmação de que realmente se trata de uma língua geral em termos conceituais.

Pesquisadores aventam suposições para saber ao certo o porquê das línguas gerais entrarem em declínio e, nesse sentido, deve-se falar em um conjunto de fatores, já que investigações mais sistematizadas sobre o assunto dão a perceber que diversas situações coadunaram para este fim. A primeira implicação a ser considerada seria a dizimação dos povos indígenas por vários motivos: apresamento, escravização, transmissão de doenças e epidemias, diminuição das áreas de plantio e caça e, por fim, o desgaste da diversidade cultural indígena frente à cultura dos colonizadores, os quais objetivavam um “progresso” pautado na cultura da Metrópole e da “cristandade universal”.

As guerras ao mesmo tempo eram frequentes. E aqui salientamos a cultura que algumas tribos indígenas praticavam: o *cunhadismo*. A prática do cunhadismo possibilitava o laço consanguíneo entre índias e brancos e, assim, o parentesco com todos os membros da tribo nativa. Para Darcy Ribeiro (2006), essa “instituição social” originou uma “numerosa camada de gente mestiça”, que era acessível a qualquer europeu que desembarcasse nas adjacências de aldeias indígenas – quando os laços eram feitos com franceses, holandeses, ingleses e alemães, as guerras eram inevitáveis e o extermínio de nativos também, já que Portugal sentia-se ameaçado ao perder territórios importantes para o sustentáculo do seu comércio.

Outro fator que possivelmente limitou a disseminação das línguas gerais foi a lei promovida pelo Marquês de Pombal em 1757, proibindo sua utilização e estabelecendo, exclusivamente, o português como língua oficial da colônia. Além disso, o sistema de

catequese dos jesuítas incorporando a “alfabetização catequética” em língua portuguesa, somados ao insucesso da escravização do índio terminou freando a dispersão das línguas gerais e, conseqüentemente, a possibilidade das línguas indígenas receberem a posição de principais influenciadoras do que hoje conhecemos como português popular brasileiro<sup>10</sup>. Podemos, no entanto, fazer um balanço, ainda provisório, no que concerne ao enriquecimento do PB pelo contato de pelo menos cinco séculos com as línguas indígenas: parte considerável da toponímia e da topografia brasileira, e um extenso léxico referente à cultura material, à alimentação, à fauna e à flora.

Por volta de 1550, outro elemento étnico ingressará na sociedade brasileira, deixando marcas de fundamental importância para a diversidade linguística do Brasil: o africano. Introduzir a escravidão ao sistema colonial de comércio transatlântico, a partir do século XVI, tornou-se essencial e lucrativo diante da crescente prosperidade dos engenhos de cana-de-açúcar no “Novo Mundo”. O tráfico negreiro foi a solução encontrada pelos donos de engenho, a fim de incrementar a produção açucareira – os escravos passaram a ser *as mãos e os pés do senhor do engenho*<sup>11</sup>. Iniciava-se, assim, a importação de um grande número de africanos capturados e escravizados em sua terra natal. Lucchesi (2008) descreve a introdução do cativo africano em terras brasileiras da seguinte forma:

O genocídio das populações indígenas e a sua resistência à integração à sociedade colonial, sobretudo nos centros urbanos, fizeram com que o negro africano se tornasse o esteio da mão-de-obra da colonização do Brasil. Inserindo-se nos diversos setores da atividade produtiva, mas concentrando-se na base da pirâmide social, os africanos e seus descendentes constituem um elemento fundamental na formação da sociedade brasileira e na expansão da língua portuguesa para todo o território brasileiro.

O sistema de escravização de africanos prolongou-se por longos três séculos e estima-se que, aproximadamente, quatro milhões de indivíduos aportaram em terras brasileiras, naturais de duas áreas da África: (i) a costa ocidental – atual Senegal até Angola, e (ii) da costa oriental – atual Moçambique até a Etiópia. Quanto ao cabedal de línguas africanas trazidas para o Brasil, Margarida Petter (2006), toma por base um número aproximado de 200 a 300 línguas, sendo a procedência, majoritariamente, de

<sup>10</sup> Houaiss (1985), “(...) com o etnocídio ocorria o glotocídio, pois línguas são fatos culturais que demandam homens e culturas”.

<sup>11</sup> Ver *Cultura e opulência no Brasil* (ANTONIL, 1982).

dois pontos da costa da África: a **Costa da Mina**, no litoral da Guiné – línguas do grupo linguístico Kwa, e o **porto de Luanda**, em Angola – línguas do grupo linguístico banto (LUCCHESI, 2000). Ressalta-se que na região da Bahia predominou os falantes das línguas kwa, o iorubá, também conhecido como nagô.



- *Século 16* - Saíam de Serra Leoa, Senegal, Guiné, Guiné-Bissau e Gâmbia e chegavam a Pará, Maranhão e Bahia.
- *Século 17* - Saíam de Congo, Gabão, Angola, Gana, Guiné Equatorial e Camarões e chegavam a Bahia, Pernambuco e Alagoas (por 300 anos foi a principal rota).
- *Fim do século 17 e século 18* - Saíam de Congo e Angola e chegavam a Bahia, Rio de Janeiro e São Paulo.
- *Século 19* - predominância de africanos vindos de Angola e Moçambique.

FIGURA 1: Rota dos escravos do século XVI ao século XIX.

FONTE: [epoca.globo.com/especiais/.../esp990920a.htm](http://epoca.globo.com/especiais/.../esp990920a.htm)

Para inviabilizar culturalmente o negro e conservar as diferenças étnicas intertribais (contendas comuns no continente africano e amplamente usado pelos traficantes negreiros) o sistema colonial do tráfico manipulava estrategicamente a seleção de escravos. A esse respeito, A. Houaiss (1985) afirma que:

(...) pelo tipo de escolha a que eram submetidos desde os portos negros até sua localização como mão-de-obra no Brasil, os negros

foram selecionados negativamente, afim de que não se adensassem em um ponto qualquer, étnica, cultural e lingüisticamente.

A política da conhecida “seleção negativa” potencializava a égide do tráfico negreiro: controlar e reter possíveis ameaças da ordem escravista<sup>12</sup>. Os escravizados com características linguísticas e étnicas comuns eram separados e distribuídos geograficamente em pontos diferentes e, eis aqui, um fator decisivo na trajetória linguística das línguas africanas para a formação do português brasileiro, nas palavras de Nina Rodrigues (1977 [1890-1905] *apud* BONVINI, 2008) – “(...) *ao desembarcar no Brasil, o negro novo (= recém-chegado) era obrigado a aprender o português, para falar com os senhores brancos, com os mestiços e os negros crioulos*”. A necessidade de aprender o português para o intercuro entre os cativos na senzala e os colonizadores (senhores, feitores e capatazes) logo ditou os trâmites de uma proficiência precária da língua, baseada na captação de ordens.

Lucchesi (2008) faz uma síntese elucidativa a respeito da situação da língua portuguesa no Brasil, do século XVI até meados do século XVIII, período caracterizado pelo “multilinguismo generalizado”:

(...) o português era apenas uma das línguas faladas pelos cerca de 300 mil habitantes da América portuguesa. Nas províncias mais periféricas, como São Paulo e Maranhão, a língua geral de base tupi predominava. Nas províncias que não impulsionavam o projeto colonial brasileiro, Pernambuco e Bahia, a massa da mão-de-obra escrava lançava mão de línguas francas africanas, como o quimbundo, para se comunicarem entre si. No interior, ao redor dos engenhos, ou nos quilombos, línguas francas africanas conviveriam com variedades pidginizadas ou crioulistadas do português.

A partir do final do século XVII, alguns desdobramentos relacionados à economia colonial presidiram singulares consequências no cenário sociolinguístico do Brasil, como bem assinala Lucchesi (2008):

- a. A descoberta das primeiras minas de ouro e, conseqüentemente, a “primeira onda lusófona” acarretará durante todo o século XVIII o influxo cada vez maior de portugueses “*e, com eles, mais de 1 milhão de escravos africanos*”;

<sup>12</sup> A Revolta dos Malês, na Bahia em 1835, é um exemplo oposto à seleção negativa: foi uma insurreição de escravos com características étnicas e linguísticas comuns.

- b. Também em razão do ciclo do ouro, teremos a interiorização do Brasil e o *primeiro surto de urbanização*;
- c. O Rio de Janeiro passa a ter importância na sociedade colonial brasileira graças ao sistema de tráfico negreiro;
- d. Há uma progressiva redução das línguas francas africanas, notava-se a diferença entre escravo doméstico/urbano e o escravo de mineração<sup>13</sup>;
- e. Advento de uma norma linguística brasileira “(...) *em termos de um falar corrente entre os brancos nascidos no Brasil*”;
- f. Em 1808, outro grande contingente de portugueses se transfere para a colônia devido à tomada da metrópole pelas tropas francesas – segunda “onda lusófona”;
- g. As línguas gerais limitavam-se aos territórios inexplorados (Amazônia e centro-oeste do Brasil);
- h. O fim oficial do tráfico negreiro, em 1850, enfraquece a possibilidade de formação de um pidgin e a permanência da **transmissão linguística irregular**;
- i. Os principais agentes linguísticos – afro-descendentes e mestiços – propagam o português popular do Brasil;

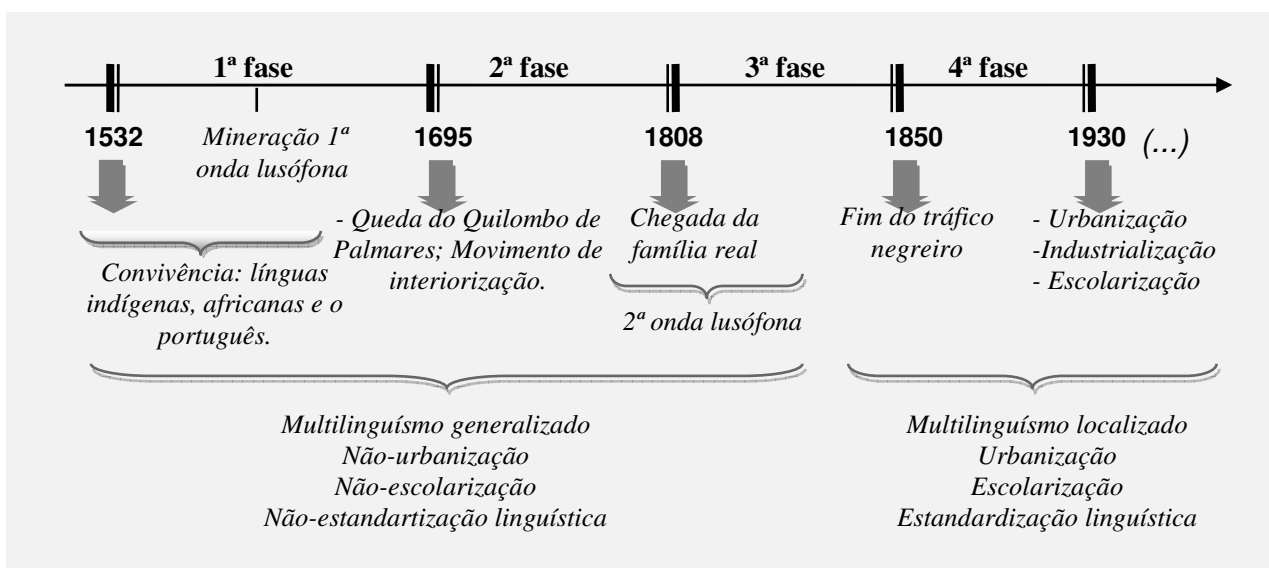


FIGURA 2: Periodização da história linguística do Brasil, segundo dados depreendidos de Lobo (2003) e Lucchesi (2008).

<sup>13</sup> Lucchesi (2008) cita Yeda Pessoa de Castro em “Os falares africanos na interação social dos primeiros séculos” como a detentora da explicação sobre a diferença entre dois tipos de cativo africano inseridos no sistema escravocrata: o escravo urbano e o escravo da mineração.

A figura 2 é uma proposta de ilustrar, com eventos da sócio-história, a formação da atual configuração linguística brasileira e, neste sentido, um subsídio valioso é fazer uma comparação com dados demográficos do Brasil. Para Mattos e Silva (2004), informações da história demográfica apontam “*pistas interessantes e significativas para a história das línguas*”, embora, a autora faça ressalva em não concordar com a afirmação de que “*a história das línguas humanas é essencialmente uma história demográfica*” (MUSSA, 1991 *apud* MATTOS E SILVA, 2004). Desta maneira, deve-se levar em conta que não há precisão de dados numéricos referentes à história étnica demográfica do Brasil em tempos mais remotos, mas, como bem salientam Reis e Silva (2009, p. 14), “*é necessário relativizar o problema das fontes, não imaginando nunca que ele possa impedir os avanços*”. É partindo deste princípio, que o trabalho de A. Mussa (1991) tem sido recorrente em estudos que analisam o percurso dos agentes que constituíram a história linguística brasileira.

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

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

Fonte: A. Mussa (1991, p. 163).

A tabela de A. Mussa quando confrontada com dados da periodização esboçam qualitativamente a trajetória dos personagens que confluíram em todo o complexo mosaico que caracteriza a realidade linguística brasileira.

Ao correlacionar a figura 2 com a tabela 2, percebemos que nos séculos XVI e XVII os africanos e, posteriormente, os afro-descendentes correspondem a cerca de 70% da população do Brasil, período em que temos o empreendimento da lavoura canavieira e, posteriormente, o movimento para o interior do país conduzido pelo novo ciclo econômico – a atividade da mineração. Este é um dado importante, na medida em que restringiu a possibilidade de se generalizar uma língua crioula no Brasil.



Segundo Bickerton (1981), a proporção numérica entre dominadores e sobrepujados é uma questão determinante para a formação de uma língua crioula – o autor estabelece que a parcela de falantes dominados não deva ultrapassar 20%. Do mesmo modo, Lucchesi (2009) afirma que: “O nível de criouliização é inversamente proporcional à facilidade de acesso aos modelos da língua alvo”. No Brasil, em nenhum momento a população de dominantes foi inferior a 30% e, entre os séculos XVII e XVIII, 60% de negros e de seus descendentes já tinham o português como língua materna. Somente em situações particulares, o número de escravos excedia o número de falantes da língua alvo: isso ocorria em algumas senzalas do interior e em alguns quilombos.

Ainda sobre a questão da não formação de uma língua crioula, vale ressaltar que a definição de escravidão está amplamente marcada pela “dessocialização, processo em que o indivíduo é capturado e apartado de sua comunidade nativa” (ALENCASTRO, 2000, p. 144). A vida de um escravo era desprovida de qualquer tipo de convívio social ou familiar e esta situação corroborou, expressivamente, para a inibição de um crioulo típico no Brasil de maneira genérica. Além disso, as altas taxas de mortalidade infantil e os contínuos deslocamentos relacionados às zonas de atividade vigente em determinada época inviabilizavam a nativização da língua-alvo – fator incisivo na gênese de línguas crioulas.

Esse período, caracterizado pelo *multilinguismo generalizado*, em que a convivência simultânea entre línguas distintas, desprovida de uma ação normatizadora, é também discutido por Mattos e Silva (2004, p. 106), que chega à seguinte constatação sobre os principais difusores do português popular brasileiro:

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, decorrentes das vicissitudes da vida econômica de seus senhores e da economia brasileira; os diversificados e múltiplos papéis por eles desempenhados na sociedade colonial 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.

Em finais do século XVIII, há um decréscimo na participação dos africanos e dos negros brasileiros, 2% e 13%, respectivamente. O movimento de interiorização do Brasil com o ciclo de ouro e diamantes de forma elevada, seguido pela cultura do algodão, arroz e a colheita de especiarias ofereceu condições a **mestiçagem** – contributo fundamental que também inibiu a probabilidade de uma real efetivação de línguas pidgin e crioula. De 10% no século XVI, os mestiços passam a quase metade da população brasileira no século XVIII.

O aspecto da mestiçagem apresentou outro olhar, ainda que gerido de forma ávida pelo espectro da escravidão, por parte da elite colonial, principalmente no que concerne a valoração social: sobreveio uma hierarquia nas esferas da sociedade brasileira, estimulando a diferença entre africano, negro e mestiço. Com o fim do tráfico, em 1850, e, mais tarde, com a abolição dos escravos, as diferenças ficavam cada vez mais intensas, possibilitando que alguns mestiços conseguissem ascender econômico e socialmente. Como não podia ser diferente, transcorria pela linha enviesada dessas diferenças, durante todo o período colonial, **a língua**, como nota Serafim da Silva Neto (1951 [1963, p. 88-89]):

Dos princípios da colonização até 1808, e daí por diante com intensidade cada vez maior, se notava a dualidade linguística entre a nata social, viveiro de brancos e mestiços que ascenderam, e a plebe, descendente dos índios, negros e mestiços da colônia.

As últimas décadas escravistas estabeleciam um forte contraste linguístico que acarretou acontecimentos decisivos que se materializaram no que Lucchesi (1994) denominou como: **polarização sociolinguística do Brasil**. E aqui, não podemos deixar de lado a chegada de imigrantes europeus e asiáticos entre os anos de 1920 e 1930.

De acordo com Lucchesi (2009, p. 55) a vinda de mais de 3 milhões de imigrantes deu espaço a uma nova etapa no percurso sociolinguístico brasileiro:

Esses indivíduos, em sua maioria, ingressaram na pirâmide social brasileira, dirigindo-se para o trabalho braçal do campo. Nestas circunstâncias, o modelo mais acessível de que dispunham para a aquisição do português era o proveniente dos capatazes e dos trabalhadores braçais locais, que, em sua maioria, eram ex-escravos africanos ou seus descendentes endógenos e mestiços; ou seja, o português que esses imigrantes – italianos, japoneses, alemães, etc – aprenderam, ao chegar ao Brasil, era o português popular, com as profundas mudanças decorrentes do processo de transmissão linguística irregular por que este havia passado.

A inserção dos imigrantes terminou proporcionando outras diferenças linguísticas no cenário da sociedade brasileira: como possuíam uma herança cultural institucional, acabavam ingressando com mais rapidez nas camadas da elite, “*levando para o seio da norma culta algumas estruturas da matiz popular que haviam adquirido em seu contato inicial com o português*” (LUCCHESI, 2009). A convivência com as línguas dos imigrantes, inicialmente, não preocupou o estado brasileiro, contudo, no século XX, algumas intervenções foram tomadas para silenciar os *idiomas estrangeiros* – decretos, leis e prisões eram frequentes, principalmente no Sul do país.

Outro singular fator contribuiu com o processo de nivelamento das duas grandes vertentes do português brasileiro, graças ao progresso no sistema rodoviário e ferroviário: a urbanização. A partir do final do século XIX, o Brasil passa de país rural para país urbano. Os crescentes centros citadinos abriram caminho para o aumento do trabalho assalariado e, com isso, certa melhoria nas condições civis dos indivíduos pertencentes à base da pirâmide.

Conectado ao ambiente urbano, teremos o surgimento dos meios de comunicação em massa (rádio, TV), que propiciaram maior interação entre as variedades rurais e variedades cultas do português e, entre 1890 e 1920, um aumento significativo na escolarização brasileira. Todos esses aspectos reunidos tornaram manifesto um desejo patente de “uniformização do português”, embora, haja ainda no Brasil uma ineficiência alarmante no ensino da norma padrão nas escolas. Mattos e Silva (2004) assim nos situa sobre essa pretensa estandardização do português brasileiro:

Não se pode aceitar mais hoje a tese da unidade do português no Brasil, uma das teses de Serafim da Silva Neto (1950, 1960) e outros. Fundava-se esse autor na possibilidade de intercomunicação em todo território brasileiro, fato que é verdadeiro, mas que recobre uma polarização de caráter social, entrevista por Paul Teyssier (1982:79), e que a sociolinguística no Brasil vem demonstrando.

Diante do exposto, não há espaço para deixar de lado a polarização do português brasileiro, marcado pelo “*multilinguismo generalizado nos dois primeiros séculos de colonização*” e, hoje, caracterizado pelo “*multilinguismo localizado*” (MATTOS E SILVA, 2004, 2006). Portanto, para entender o caráter múltiplo da língua portuguesa no Brasil, devemos examiná-la como “heterogênea, plural e polarizada” (MATTOS E

SILVA, 2006; LUCCHESI, 2006), sendo possível determinar a coexistência de várias normas: de um lado, as **normas populares ou vernáculas** e, do outro, as **normas cultas** e, no horizonte, “**para, ou pára a norma padrão**” (MATTOS E SILVA, 2006, p. 230).

### 2.1.1 Reflexos da polarização linguística: convivência de normas x ensino

Ao pensarmos sobre língua contextualizando-a a aspectos políticos e sociais, logo observamos refletidas em sua avaliação as relações de poder e dominação ligadas às transmissões culturais de um povo. Além de “veicular informação”, a língua torna impressa a posição social do indivíduo e, é através dela, que numa sociedade de classes, comumente de tradição escrita, se instala, muitas vezes, a dicotomia: incluído e excluído socialmente. Não há como negar que a palavra reflete o poder acumulado pelo falante, sendo impossível pensar a construção de um indivíduo sem perceber os efeitos de seu ato linguístico dentro da sociedade da qual ele faz parte (GNERRE, 1998). Neste sentido, é pertinente a constatação de Beger e Luckmann (1985) quando afirmam que “[...] *não há pensamento humano que seja imune às influências ideologizantes de seu contexto social*”.

No Brasil, a relação entre língua e poder socioeconômico é bastante característico devido a sua conjuntura sócio-histórica e, de fato, os grupos sociais são identificados pelas configurações linguísticas que lhes são de comum caráter. Em vista disto, um conceito que detém lócus privilegiado sobre a questão da polarização linguística brasileira e que perpassa “um *agregado de valores socioculturais articulados*” é, precisamente, o conceito de **norma linguística**.

A teorização de Coseriu (1979 [1921]) sobre o conceito de *norma* é considerada por muitos linguistas como a que melhor estabeleceu, no contexto estruturalista, a ambiguidade da antinomia de Saussure entre *língua* e *fala*. Coseriu institui a separação entre o que é funcional do que é social, definindo diferenças entre **sistema**, **norma** e **fala**. Desta maneira, segundo o autor, a *norma* “es la realización ‘colectiva’ del sistema, que contiene el sistema mismo y, además, los elementos funcionalmente ‘non-pertinentes’, pero normales en el hablar de una comunidad”. A *norma* caracterizada por Coseriu é um grau intermediário de abstração entre o *sistema* (+abstrato) e a *fala*

(+concreto) e exerce papel fundamental na atuação do sistema dentro de uma coletividade.

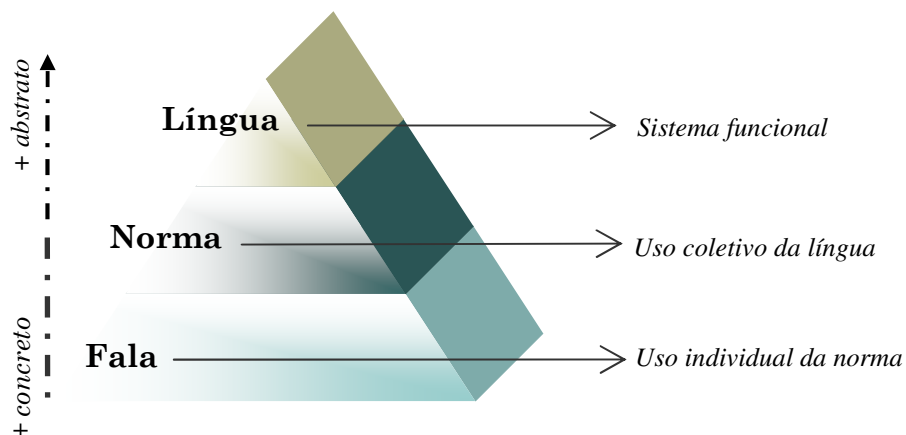


FIGURA 3: *Língua, norma e fala*, segundo a teorização de Coseriu (1979 [1921]).

Segundo Lucchesi (1994), a teorização de Coseriu sobre a definição de norma linguística “esbarra em alguns problemas teóricos cruciais”: (i) sustentar empiricamente a concepção de sistema funcional diverso do sistema normal e (ii) não explicar a relação entre variações sistemáticas e mudança linguística. Assim, Lucchesi (1994), e posteriormente, Lobo e Lucchesi (1998) retomam o conceito de *norma* inserindo-o no contexto brasileiro, por meio do qual assinalam diferenças entre o que intitularam de **normas cultas** e **normas populares brasileiras**. Sem ignorar a polissemia do termo, os autores deixam claro que, mesmo sendo uma formulação estruturalista, o conceito de *norma* pode ser amoldado no âmbito sociolinguístico, resolvendo os problemas teóricos deixados de lado por Coseriu. Portanto, apóiam-se na identificação de “*sistemas de avaliação social da variação linguística, associados a processos independentes de mudança lingüística*” (LUCCHESI, 2006, p. 90).

Lucchesi (1994, p.27), assim, distingue a categorização da noção de norma: não de apenas uma *norma* exclusiva, mas de **normas** (no plural):

Apesar de conceber o português do Brasil como um sistema heterogêneo e variável, defendo que ele não pode ser analisado como um sistema unitário, mas um sistema plural, mais precisamente um *diassistema*, constituído por pelo menos dois subsistemas, também eles heterogêneos e variáveis, que defini aqui como normas.

Segundo o autor, podemos sistematizar três conceitos de *norma*: (i) normas cultas; (ii) normas vernáculas e (iii) norma padrão. As **normas cultas** são típicas dos segmentos mais escolarizados da sociedade, com acesso aos principais espaços da cidadania; as **normas populares**, dizem respeito aos padrões de comportamento linguístico da maior parte da população brasileira, que não tem o mesmo ingresso aos direitos elementares de cidadão. A **norma padrão**, por sua vez, tem sido frequentemente confundida com *norma culta*, todavia são duas entidades sociolinguísticas bastante diferentes. Enquanto esta última, como já mencionado, refere-se às variedades efetivamente usadas pelos falantes urbanos escolarizados, a primeira é uma norma idealizada, abstrata e acumulária “*as formas contidas e prescritas pelas gramáticas normativas*”. É crucial, desta forma, para entender o entrelaçamento que se encontra subjacente entre o linguístico e o social no Brasil, distinguir o que no senso comum já se cristalizou, erroneamente, como sinônimos: padrão e culto.

A confusão desses adjetivos está arraigada no conceito tradicional, alicerçado na escrita, formando um modelo distanciado da realidade de usos da língua. O próprio Serafim da Silva Neto, filólogo brasileiro que fundamentava seu pensamento crítico na acepção de que “A língua constitui, sem dúvida possível, o mais importante traço de cultura”, já na década de 50, reconhecia a “dualidade lingüística” do português brasileiro. Entretanto, na visão do autor, havia uma superioridade cultural portuguesa quando confrontada com a cultura brasileira, e propunha em seu pensamento a valorização da modalidade escrita:

Mas como superior produto de intercomunicação, refinado pelos esforços de gerações sucessivas, existe a *língua escrita*, poderoso reflexo de civilização e mantenedor da unidade. Ela é, ao mesmo tempo, disciplina e seleção – não pode apresentar-se com a gama infinita da língua falada. (IELPB, 1986 [1950], p. 19-20)

O pensamento de Serafim da Silva Neto não ficou no passado, pois, ainda hoje, o ensino da *norma padrão* é disseminado, protegido e preconizado como detentor de único modelo correto, eficiente e adequado pelos puristas da língua, embora, “[...] o mesmo professor que ensina essa gramática não consegue observá-la em sua própria fala nem mesmo na comunicação dentro de seu grupo profissional” (RODRIGUES, 2007, p. 18). Partindo dessa contradição, os linguistas conhecedores da “realidade multifacetada” de que se reveste a língua portuguesa no Brasil apóiam um ensino de

língua baseada em “uma pedagogia voltada para o *todo* da língua e não para algumas de suas formas” (SILVA, 1996, p. 37).

Infelizmente, a noção de pluralidade de normas não é reconhecida de forma geral no extenso território brasileiro, embora haja maior acesso às pesquisas promovidas pelos cursos de Letras em todo país. Até mesmo os Parâmetros Curriculares Nacionais de 1998 já dispõe em seu texto uma compreensão do ensino da língua materna mais próxima da realidade brasileira quando dão ênfase à não-estigmatização das variedades linguísticas populares e inserem um conceito mais amplo – o de “língua de cultura” – no lugar de *norma culta* ou *língua culta* como detentora de *status* elementar do ensino de língua portuguesa<sup>14</sup>. Contudo, as instituições de ensino continuam a impor uma abstração, um padrão linguístico inalcançável.

É notável, a evidente ausência de harmonia entre a descrição dos compêndios gramaticais e os fatos linguísticos em curso, que unicamente promove entre os jovens vindos de classes desprestigiadas economicamente o sentimento de estarem aprendendo uma língua “nova” que mais sugere ser uma língua estrangeira (cf. SILVA, 1996). Consideram-na uma língua difícil ou até mesmo dizem, equivocadamente, que não têm conhecimento de sua própria língua materna. Essa falta de reciprocidade faz com que haja uma preocupação com as políticas de ensino de língua portuguesa, já que não se sabe, muitas vezes, quais estratégias o professor deve usar para garantir um ensino eficiente. Mattos e Silva, em artigo publicado no Jornal *a Tarde* em março de 1986, defende seu ponto de vista sobre a questão do método de ensino de língua portuguesa nas escolas da seguinte forma:

Fugindo-se das posições exclusivas e extremistas que defendem o purismo lingüístico, ou seja, a estrita observância de uma norma ideal prescrita por incertos sabedores, ou o populismo lingüístico que renega qualquer norma e admite qualquer uso lingüístico adequado a qualquer usuário, em qualquer situação de comunicação, deve-se defender que seja dada ao falante – e será esse o trabalho ao ensinar-se o falante nativo – a possibilidade de conhecer, para poder escolher conscientemente, as formas de uso de sua língua entre as diversas formas de se manifestar que ela pode oferecer. Está implícito, portanto, o poder do indivíduo, sujeito de sua fala, de decidir sobre a

<sup>14</sup> Silva (2006, p. 31), no entanto, faz uma crítica contundente as diretrizes e bases dos Parâmetros Curriculares Nacionais de 1998: “As expectativas positivas criadas na introdução (...) caem por terra quando lemos os tópicos 7 e 8 do documento que, respectivamente, tratam das orientações metodológicas e dos programas de ensino. Analisando-os, fica claro que a *tabula rasa* da pedagogia de tradição secular, em relação à diversidade lingüística, se mantém no documento oficial de 1986, apesar do diagnóstico que faz da realidade social e de suas conseqüências para a realidade lingüística”





FIGURA 4: Distribuição das variedades linguísticas do PB, conforme o *continuum* de urbanização.

FONTE: Bortoni-Ricardo (2006, p. 52).

Bortoni-Ricardo deixa claro que não há “fronteiras rígidas que separem os falares rurais, rurbanos ou urbanos” e afirma que, a propósito de suas fronteiras, deve-se falar em fluidez e sobreposições de variedades linguísticas. Desta forma, podemos afirmar que os deslocamentos do campo para a cidade, fortemente integrados com a construção de rodovias e ferrovias, aliados à proliferação dos meios de comunicação e à posterior democratização da educação, terminam por promover desdobramentos efetivos na variedade de língua dos grupos urbanos em direção ao vernáculo rural e vice-versa. A norma culta, retentora de uma tendência a absorver as características do arquétipo do português europeu, passa a ter possibilidades ativas de introduzir em seu vernáculo construções típicas das normas populares. Do mesmo modo, o estilo de vida e hábito urbano, almejado pela maioria da população, fez com que houvesse aproximação real de unidades linguísticas peculiares do padrão culto nas esferas das normas populares, sobretudo, em situações formais.

Assim, para alicerçar a problematização das feições centrais sobre valorizar uma variante linguística em detrimento de outra no Brasil, deve-se fazer um retorno ao passado colonial, a fim de reconhecer o universo de possibilidades linguísticas surgidas a partir do contato entre o português europeu e os diversos grupos étnicos e sociais desse período. Somente com o recuo às origens formadoras da identidade brasileira, alcançaremos as respostas do porquê de se impor um conjunto de valores linguísticos ancorado no discurso preconceituoso do que é melhor/pior, certo/errado, bom/ruim, de prestígio/sem prestígio e fixado, exclusivamente, na tradição escrita.

## 2.2 A COMUNIDADE DO PORTUGUÊS AFRO-BRASILEIRO ESTUDADA

### 2.2.1 O português afro-brasileiro do interior do estado da Bahia

Várias foram as formas de resistência encontradas pelos negros quando alocados em países do continente europeu e americano como escravos. A luta árdua e cotidiana consistia inegavelmente pelo estabelecimento da manutenção do seu passado histórico e

pela oposição à opressão do regime escravocrata. Entretanto, ganha força e conotação de símbolo e auge da resistência o **quilombo** – *“toda habitação de negros fugidos que passem de cinco, em parte desprovida, ainda que não tenham ranchos levantados nem se achem pilões neles”* (Conselho Ultramarino Português de 1740).

Segundo Reis e Gomes (2008), as comunidades quilombolas formadas no Brasil tiveram procedências que variavam de acordo com a relação instituída pela ordem escravista e com as circunstâncias da demografia do território colonial. Assim, embora considerados pelas autoridades portuguesas como redutos de escravos fugidos e que deveriam ser desmembrados a qualquer custo, os quilombos em algumas regiões e em situações especiais colaboravam no abastecimento de produtos e alimentos de comunidades relativamente afastadas dos centros mais urbanizados. Salles e Soares (2005, p.) sintetizam da seguinte maneira a importância do quilombo como instituição afro-brasileira:

[...] os quilombos foram parte significativa e expressiva na formação de experiências na formação de experiências de sociabilidade, práticas culturais, religiosas, políticas, geração de identidades, ocupação territorial que estiveram na origem da configuração de uma identidade brasileira, mais precisamente afro-brasileira. Mesmo quando sua história e sua memória tenham sido, em grande parte, silenciadas.

Conforme historiadores especializados no assunto, podemos dividir o quilombo em duas modalidades: (i) os que se formaram no período colonial como focos de resistência, constituídos de negros foragidos (como o Quilombo de Palmares) e (ii) os que se formaram após a abolição, constituídos de alforriados que viviam à margem da sociedade, desassistidos pela mesma sociedade que os “libertou”. A modalidade que nos interessa é a última, pois a comunidade aqui analisada foi formada por ex-escravos que migraram para áreas desabitadas e/ou que permaneceram nas terras abandonadas por colonos europeus.

Hoje em dia, são reconhecidas oficialmente pela Fundação Cultural Palmares (instituição pública ligada ao Ministério da Cultura, estabelecida pela Lei Federal n. 7668, de 22 de agosto de 1988, aprovada pelo Decreto n. 418, de 10 de janeiro de 1992) 159 comunidades quilombolas. Algumas, no entanto, ainda estão passando por um processo de reconhecimento, outras sequer estão na lista de espera.

### 2.2.1.1 Helvécia

*"Um falar crioulo que deve ter sido geral, já que em 1961 dele subsistiam ainda vestígios"* (FERREIRA, 1984)

A região de Helvécia localiza-se no extremo sul do Estado da Bahia, no Município de Nova Viçosa. Foi fundada em 1818 e, quarenta anos mais tarde, era constituída de algumas dezenas de estâncias com 200 colonos, em sua maior parte, imigrantes alemães e suíços, além de franceses e empregados brasileiros. Quanto ao número de negros, fala-se em, aproximadamente, 2000 homens que serviam como mão-de-obra escrava na cultura do café.

A então conhecida Colônia Leopoldina no Brasil teve anos prósperos com o empreendimento cafeeiro, chegando a ser a responsável por aproximadamente 90% da produção de café no início de 1850. Somente em fins da década de 1880 a cultura entra em declínio, por duas razões principais: (i) a abolição da escravatura em 13 de maio de 1888 que terminou e a expansão da cultura do café para regiões mais férteis da Província, de modo que, no final da década de 1880, o Recôncavo já havia se tornado o maior centro cafeeiro da Bahia fazendo com que os colonos europeus abdicassem da região. Em seguida, de forma gradativa, estabeleceu-se na região ex-escravos que passam a viver isolados. O complicado acesso foi fator imprescindível para o conservadorismo de elementos crioulizantes na comunidade de fala de Helvécia. Segundo Lucchesi (2009, p. 92), as circunstâncias em que se achava Helvécia nesse período eram favoráveis às situações prototípicas de crioulição: primeiro, devido à proporção numérica entre escravos e brancos de dez para um, que terminava por inviabilizar o acesso aos modelos da língua alvo; segundo, a probabilidade de os empregados assalariados constituírem de falantes do português; e terceiro, a incidência de escravos crioulos vindos de outras áreas em que tanto senhores quanto trabalhadores eram falantes da língua alvo.

As primeiras observações linguísticas na comunidade foram feitas por Ferreira, em 1961, nas pesquisas de campo do *Atlas prévio dos falares baianos* (APFB). Não havia ferramentas de áudio disponíveis e tudo foi registrado em punho pelos

pesquisadores. Ferreira relacionou peculiaridades interessantes que apontam a existência de um processo diferenciado do português rural falado no Brasil, demonstrado traços característicos de uma possível criouliização em tempos pretéritos no sistema linguístico dessa comunidade. A esse respeito, nos diz a pesquisadora (1984):

Fatos que são peculiares de Helvécia, que ocorrem com uma certa constância, e que não registramos em nenhuma outra localidade ou então não foram registradas ocasionalmente. Esses fatos nos fizeram decidir que sim, que em Helvécia há, ou melhor, havia em 1961, remanescentes de um falar crioulo.

Os traços linguísticos encontrados por Ferreira (1984), que apontam evidências de um estágio pretérito de criouliização, foram: (1) variação na concordância de gênero no centro de SNs; (2) simplificação da morfologia flexional do verbo e, por fim, o objeto de estudo desta dissertação, (3) uso variável do artigo definido.

- |  |
|--|
| <p>a. <i>Éla é mûitu sa'ïdu</i><br/> b. <i>Io 'sabi</i><br/> c. <i>Io koêi'se Ø pai di Ugu</i></p> |
|--|

Trinta e três anos se passam quando os pesquisadores Alan Baxter e Lucchesi fazem as primeiras gravações das amostras vernáculas da comunidade de Helvécia com o apoio metodológico da Sociolinguística, compondo o *corpus* do português afro-brasileiro do Projeto *Vertentes*. O estudo sistemático do *corpus* permitiu a constatação de que realmente havia uma variação na fala dos moradores que “teriam feito parte do repertório gramatical do antigo crioulo de Helvécia” (LUCCHESI, 2009, p. 93). Além de observarem os mesmos traços linguísticos assinalados por Ferreira (1961), ainda que com menor frequência, outros traços foram reconhecidos e fazem parte dos trabalhos de iniciação científica, dissertações de mestrado e teses de doutorado veiculados ao Projeto *Vertentes*.

### 3 O ARTIGO DEFINIDO EM PORTUGUÊS

#### 3.1. FORMAÇÃO E DEFINIÇÃO TRADICIONAL DO ARTIGO DEFINIDO

O Latim Clássico era uma língua que se caracterizava pelas marcas internas, admitindo desinências e casos específicos para cada declinação. Segundo Mattos e Silva (1993, p. 23), o surgimento do artigo definido é proveniente do “complexo sistema de demonstrativos latinos, *illle*, na forma de acusativo, como determinador por excelência dos nomes substantivos, além de marcador básico do gênero”. Entende-se, deste modo, que a mudança ocorreu devido à necessidade de se suprir os mecanismos funcionais perdidos na passagem do Latim Clássico para o Latim Vulgar. Para Teyssier (2001, p. 20), esse processo ocorreu do seguinte modo:

As quatro formas saídas do acusativo, diferenciadas em número e em gênero – *illum, illam, illos, illas* –, dão inicialmente *lo, la, los, las*, em virtude da aférese sofrida pelo seu emprego proclítico. Como estes artigos vinham freqüentemente precedidos de palavras terminadas de vogal – ex: vejo lo cavalo, vende la casa – o “l” desapareceu à semelhança de todos os “l” da língua em que se achavam em preposição intervocálica, com o que se chegou às formas *o, a, os, as*.

Assim sendo, tornam-se claros os motivos da proximidade semântica entre o artigo definido e o pronome demonstrativo quanto à determinação do substantivo em todas as línguas românicas. Segundo Dubois (1973), os artigos somente se distinguem dos demonstrativos na sintaxe, particularmente na pronominalização.

Na maior parte das gramáticas tradicionais, o artigo definido é a classe de palavras responsável pela consignação do substantivo, ancorando sua definição de cunho funcional, no domínio da sentença. O artigo indefinido, por sua vez, estabelece a indeterminação do substantivo, ou seja, do ser em relação à sua classe. Desta maneira, a analogia entre o sistema de artigos da língua portuguesa dá-se, sobretudo, no princípio

de definitude, que termina por conduzir o comportamento morfossintático e semântico dessa subcategoria dos determinantes.

Cunha e Cintra (1985, p. 210) conceituam o artigo definido como sinal de *notoriedade*, de conhecimento prévio dos falantes sobre o elemento referido, caracterizando-o como o intróito do substantivo para indicar-lhe gênero e número. Ao comparar o artigo definido com o artigo indefinido e o demonstrativo, os autores explicam que o valor determinativo perfaz uma linha sucessiva de exatidão do substantivo, como nos exemplos abaixo:

- (1) Encontrou *uma mulher bonita*, com *um vestido vermelho*.
- (2) Encontrou *a mulher bonita*, com *o vestido vermelho*.
- (3) Encontrou *esta mulher bonita*, com *este vestido vermelho*.

Logo, percebemos que os substantivos *mulher* e *vestido* alteram sua precisão à medida que mudamos o emprego do artigo indefinido para o artigo definido e, por fim, para o demonstrativo. Em (1) temos a indicação da espécie/classe *mulher* e *vestido*; em (2) o artigo definido restringe os substantivos, individualizando-os e em (3) notamos uma maior abreviação do sentido dos substantivos *mulher* e *vestido*, já que os demonstrativos os localizam no tempo e no espaço.

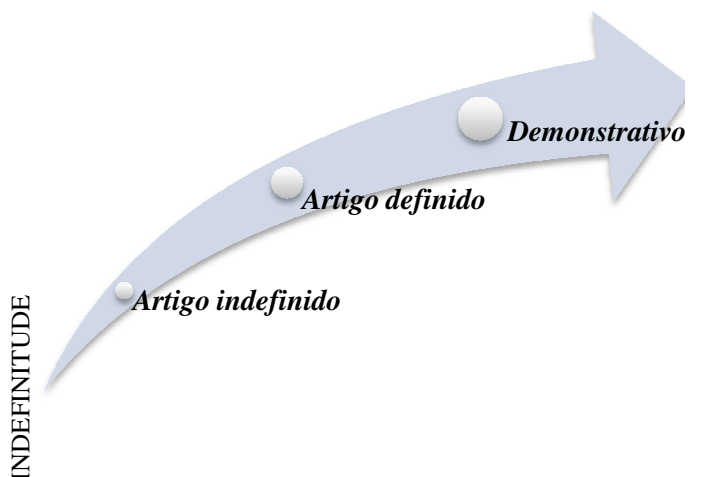


FIGURA 5: *Continuum* de definitude entre o demonstrativo e os artigos definido e indefinido.

Cunha e Cintra (2001) também abordam o emprego do artigo definido para *expressir a totalidade específica de um gênero*, principalmente no singular, com valor

genérico: “*O guarani fez-se aliado do espanhol*”. Os autores ressaltam ainda que esse tipo de valor genérico é mais distintivo nos provérbios e, frequentemente, observado em substantivos abstratos.

De forma equivalente, a *Nova Gramática do Português Contemporâneo* (ROCHA LIMA, 2003, p. 93) assume que o artigo é a classe que se integra ao substantivo “para indicar que se trata de um ser claramente determinado entre outros da mesma espécie – que o ouvinte ou o leitor já sabem quem é, pelas circunstâncias que cercam a enunciação da frase”. Contudo, o gramático faz uma ressalva: “o artigo definido também se usa com referência à espécie inteira” como em: “*O limão é uma fruta ácida* (Isto é: todo limão)”. O autor não explica o porquê dessa possibilidade na língua.

Na abordagem de Bechara (2001, p. 153), temos o valor *atualizador* do artigo como contributo importante a se levar em conta, já que a partir deste decorrem os demais valores contextuais:

O artigo definido identifica o objeto designado pelo nome a que se liga, delimitando-o, extraíndo-o de entre os objetos de mesma classe como aquele que já foi (ou será imediatamente) conhecido do ouvinte quer através do discurso (que dele faz menção) quer pela “*dêixis*” (que o mostra, ordenando-o espacial e temporalmente), quer pelo contexto idiomático, no qual a palavra é quando não ulteriormente determinada, nome de conceito ou de toda uma classe de objetos.

Na Gramática *Houaiss da Língua Portuguesa* (AZEREDO, 2008, p. 246), no tópico *determinação identificadora e remissiva: o artigo*, temos uma explicação, ainda que sucinta, da dimensão discursiva no uso desse determinante. Segundo o autor, o artigo definido “representa qualquer unidade conceitual – coisa, ideia, ser – como parte do conhecimento prévio do interlocutor”. O conhecimento antecedente é caracterizado por dois tipos de referente: (1) partilhado pela comunidade no mais amplo sentido possível, de uso geral pela sociedade e (2) restrito a uma situação particular, “conhecidos nos limites de situações particulares de interlocução”.

(4) Newton descobriu a lei da gravidade;

(5) O campeonato começará amanhã.

Embora haja nas gramáticas mais modernas o reconhecimento da importância do artigo definido no discurso e na interação, é perceptível que o foco ainda permanece no aspecto da definitude dos substantivos em função do artigo que os acompanha. Assim, o que notamos nos compêndios gramaticais é que o uso do artigo está pautado em seu reconhecimento como artefato à parte, colaborador da informação veiculada pelo SN.

No entanto, diversos exemplos, até mesmo os encontrados nas gramáticas normativas, confirmam que o simples emprego do artigo definido no SN não é auto-suficiente para garantir a peculiaridade da definitude. Os gramáticos tradicionalistas terminam utilizando a intuição, listando uma infinidade de usos e exceções, sem oferecer, contudo, nenhuma explicação dos diversos contextos que escapam às regras. Sendo assim, anulam, sistematicamente, a variabilidade da língua e, no caso do artigo definido, por não ser uma variável desprestigiada, a falta de esclarecimento é latente.

A carência de conformidade entre os critérios para estabelecer esse valor semântico e uma nomenclatura que facilite sua identificação implica no obscurecimento dos reais motivos que direcionam a identificação ou a não-identificação de um nome. A seleção do falante em empregar ou não o artigo está associada não somente ao critério funcional, mas aos critérios semânticos e pragmáticos do fenômeno, e, aqui, vale mencionar a observação de Kato (1974, p. 31) sobre o assunto: “(...) a abordagem puramente sintática tem dado apenas soluções parciais para o problema do artigo”.

Dessa maneira, apreender como funcionam as regras de utilização ou não utilização do artigo definido passa pelo entendimento das *oposições definido-indefinido* e *genérico-particular*, dando atenção especial ao componente semântico e pragmático.

### 3.2 AS FUNÇÕES DO ARTIGO DEFINIDO

Kato (1974), ao tratar do artigo definido, segue a semântica gerativa e defende a hipótese de que este determinante “*tem valor de um quantificador, cujo escopo são todos os elementos de um conjunto referencial*”. O emprego do artigo definido no português, segundo a autora, segue as seguintes definições:

1. Indicador do elemento que é definido pelo conceito dentro de certo perfil ou limite imaginário (de referente preciso).



(6) *Albert Einstein é o mais memorável físico do mundo.*

2. Aplicável em pessoa ou coisa que é única em gênero (entidade única).

(7) *A lua é o satélite da Terra.*

3. Empregado perante um nome para indicar que a ideia expressada pelo nome já foi citada e para referir a declaração novamente;

(8) *Achei um lugar lindo para passar as férias. O lugar é famoso por suas praias paradisíacas e comida típica.*

4. Obrigatório em alguns topônimos (nomes próprios);

(9) *Os Estados Unidos da América são uma república constitucional federal composta por cinquenta estados e um distrito federal.*

5. Definidor de pessoa ou coisa exercendo o genitivo, o advérbio, o sintagma preposicionado ou oração relativa.

(10) *A menina dos olhos; o caminho lá fora; o livro em cima da mesa; o caderno que entregarei amanhã.*

De acordo com Kato (1974, p. 135), uma das condições para se ter artigo definido advém *da possibilidade de identificação de um particular pelo interlocutor, ou melhor, a pressuposição do elocutor de que o interlocutor identificará o seu referente.* Essa afirmação pode ser comprovada com o recurso aos conceitos de *dêixis* e *anáfora*.

Ao situarmos a *dêixis* dentro da abordagem de referencialidade (cf. Lahud, 1979 *apud* Negri, 1986), tomamos por conceito todo enunciado que se revela numa circunstância definida no espaço e no tempo. Neste sentido, o ouvinte intui a identidade do referente no momento da consolidação de determinada ideia durante o enunciado.

São considerados comumente **elementos dêiticos** os *pronomes* e os *demonstrativos*, embora outros elementos operem igualmente a mesma função<sup>15</sup>.

Assim sendo, podemos afirmar que o artigo definido age como elemento dêitico, sobretudo quando intercambiável ao demonstrativo (cf. Kato, 1974, p. 136):

(11) **O** gato gosta de carne.

**Esse/Aquele** gato gosta de carne.

O uso anafórico surge em uma ocorrência de SNs com o mesmo referente, sendo que a segunda menção apresenta forma pronominalizada, seja por componente lexical seja por outro elemento identificador:

(12) Conheci um linguista em Londres. **O linguista** é famoso.

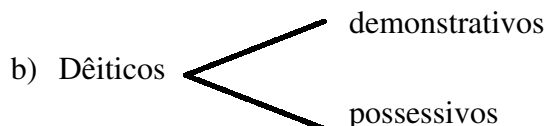
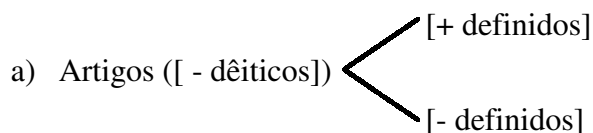
Ao comparar as línguas inglesa, portuguesa e a japonesa, Kato (1974) afirma haver uma mesma base nas três línguas com diferenças no *output* da regra que a autora nomeia de *regras transformacionais*<sup>16</sup>. Sua conclusão é a de que o artigo não pode ser postulado como uma categoria, já que em muitas línguas esse elemento é inexistente.

### 3.3 PROPRIEDADES SEMÂNTICAS E SINTÁTICAS DO ARTIGO DEFINIDO

Na *Gramática da Língua Portuguesa* (MATEUS *et al.* 2003), no tópico sobre os determinantes, as autoras afirmam que se trata de uma classe limitada de elementos que antecedem o nome, estabelecendo valores referenciais de individualização das expressões nominais. Pertencem a esse grupo os artigos, os demonstrativos e os possessivos, agrupados em relação às propriedades sintáticas e semânticas da seguinte maneira:

<sup>15</sup> Lyons (1968) acrescenta alguns adjuntos adverbiais (*agora, depois, aqui, ali*) elementos que atuam como dêiticos.

<sup>16</sup> As regras transformacionais são interpretações relativas às orações. Sobre as regras transformacionais que geram artigos, ver Kato (1974, p. 15).



Ressaltam a proximidade entre os artigos definidos e os demonstrativos, pois ocorrem em distribuição complementar e possuem conteúdo semântico definido, em oposição aos possessivos que co-ocorrem com os primeiros sem possuírem a definição como atributo inerente. Salientam, ainda, que o artigo definido não pode surgir de forma isolada, *quer por elipse no N quer por movimento do N e do complemento* (ao contrário dos indefinidos ou de quantificadores):

- (13) \* A Maria viu as.
- (14) A Maria viu umas/ várias/ algumas.
- (15) \*Livros de história, li os.
- (16) Livros de história, li um/uns.

A única maneira de ocorrer artigo definido sem o nome é em construções com complemento sob a forma de SP ou um modificador (adjetivo ou oração relativa):

- (17) Os [-] do andar de cima são meus amigos.
- (18) Os [-] simpáticos são meus amigos.
- (19) Os [-] que vivem no andar de cima são meus amigos.

### 3.3.1 Sobre a ausência de artigo definido

Segundo Mateus *et al.* (2003), há vários contextos em que um nome pode aparecer sem determinantes ou quantificadores. Destacam-se, nesse sentido, as posições não argumentais:

**1) Vocativos**

(20) Menina, vem cá!

(21) Pedro, vem cá!

**2) Exclamativas elípticas**

(22) Menina bonita!

**3) Posição predicativa**

(23) A Maria é professora

≠ A Maria é a professora

≠ A Maria é uma professora.

Existem casos em que a ausência torna-se um meio de quantificação. Isso ocorre particularmente em posições argumentais, em que há possibilidade de um nome surgir sem qualquer determinante ou quantificador:

**4) Verbo *ter***

(24) Tenho medo/ raiva/ fome.

**5) Nome massivo no singular**

(25) Faço pintura, escultura, cerâmica.

(26) Perdiam dinheiro ao jogo.

(27) Bebi água.

**6) Plurais simples com nomes contáveis**

(28) Sei línguas.

(29) Quis tomar comprimidos.

(30) Cadeiras desenhadas pelo Siza Vieira são caras.

### 3.4 A VARIAÇÃO NO USO DO ARTIGO DEFINIDO EM PORTUGUÊS

Grande parte dos trabalhos relacionados ao artigo definido dentro de uma perspectiva sociolinguística restringe a abordagem do fenômeno a dois contextos específicos: a variação no uso do artigo definido **diante de possessivos** e **diante de nomes próprios**. Destacam-se os trabalhos realizados por Giselle M. Silva e Dinah Callou. As autoras observaram que a presença ou a ausência de determinantes no português brasileiro não são tão eventuais como se poderia pensar, apesar de aparentemente não haver uma distribuição clara em relação à variação social (CALLOU e SILVA, 1996, 1997 e 2004; e SILVA, 1982, 1996 e 1998).

Segundo Callou e Silva (1996, p. 3), com base em dados da língua escrita e falada de Portugal (século XIV ao XX) e do Brasil (século XIX ao XX)<sup>17</sup>, a frequência do uso do artigo definido frente a possessivos se intensificou amplamente nos dois países, como podemos observar pelo gráfico abaixo:

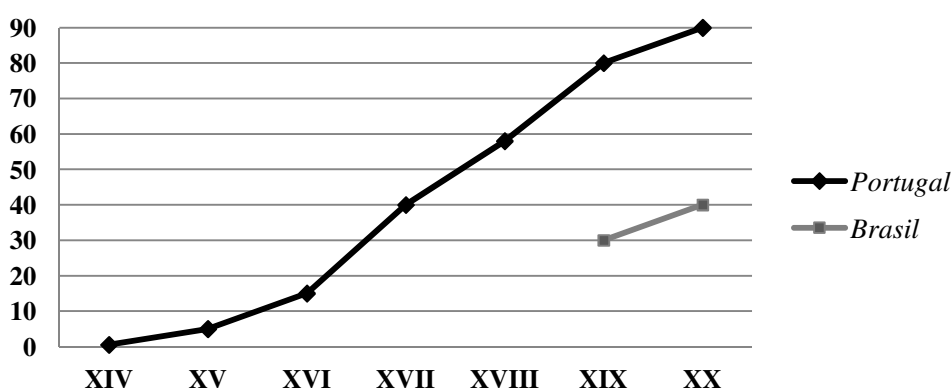


GRÁFICO 1: Frequência do artigo definido diante de possessivo, adaptado de Callou e Silva (1996, p. 3).

Percebemos, nos dados apresentados no gráfico 1, que a utilização do artigo definido diante de possessivos, em Portugal, é incipiente no século XIV, aumentando durante os séculos subsequentes, até alcançar a expressiva marca de 90% no século XX. Em relação ao uso do artigo diante de possessivos no Brasil, temos praticamente a

<sup>17</sup> As autoras analisam a língua escrita, textos de Portugal e Brasil, e a língua oral, falantes com curso universitário completo das cidades de Porto Alegre, São Paulo, Rio de Janeiro, Salvador e Recife, relacionados ao Projeto NURC – Norma Urbana Culta.

mesma frequência assinalada em Portugal nos séculos XVI e XVII, período que diz respeito ao início da colonização em terras brasileiras. Segundo as autoras, esse fato aponta para um caráter mais conservador do português do Brasil nesse aspecto da gramática.

No que se refere ao uso do artigo definido diante de nomes próprios, no português luso, há uma progressão no percentual de modo significativo somente a partir do século XVII, atingindo pouco mais de 25% no século XX (cf. Gráfico 2). As frequências do uso de artigo definido, nesse contexto, no português brasileiro são maiores entre os séculos XIX e XX, de 30% a 40%, respectivamente.

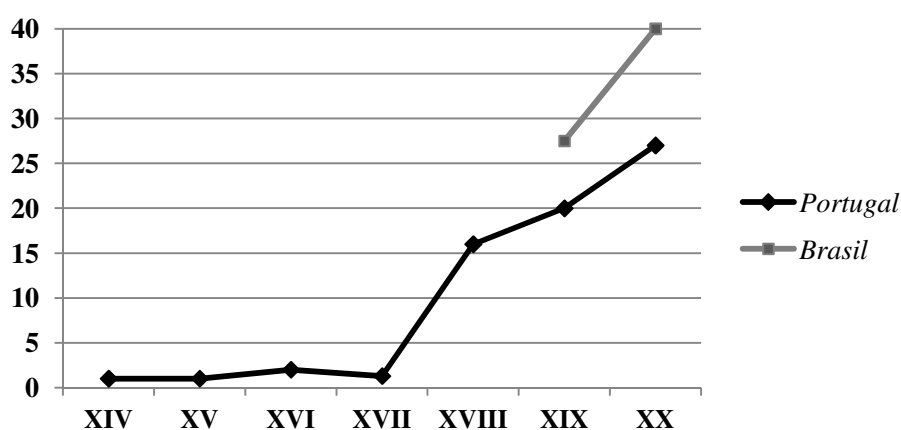


GRÁFICO 2: Frequência do artigo definido diante de nome próprio, adaptado de Callou e Silva (1996, p. 3).

Callou (1992), utilizando dados do português escrito em várias épocas, salienta que o uso do artigo definido frente a nome próprio de pessoa está estritamente ligado ao fenômeno da definitude e que, no século XX, sua ocorrência é favorecida por fatores, como: **situação discursiva**, **dimensão do vocábulo** e **função sintática**. A autora depreende que há um equívoco das gramáticas tradicionais em relacionar a utilização do artigo definido diante de antropônimos a fatores de familiaridade ou vínculo afetivo. Sobre essa incongruência, a autora argumenta da seguinte forma:

(...) se as gramáticas normativas brasileiras estão certas, ao insistirem, entre outros fatores, na familiaridade, intimidade e afetividade refletidas no uso do artigo com nome de pessoas (...) é justo pensar, imediatamente, no que bloquearia a demonstração desses sentimentos pelos portugueses antigos, visto que nos textos medievais o artigo

ocorre pouco. Também por que brasileiros de algumas regiões, como São Paulo e Rio exteriorizam a sua afetividade, enquanto outros, de Salvador (ou até de Niterói) são tão reprimidos, uma vez que não costumam usar artigo com nomes próprios, na fala cotidiana, mesmo que se refiram a pessoas muito próximas [afetividade] (...) Se a intensificação do uso se verifica ao longo da história do português, seria necessário descobrir os fatores condicionantes, fora do âmbito emocional, afetivo, tarefa que não parece fácil.

Segundo Callou e Silva (1997, p. 12), o licenciamento do uso do artigo não tem uma valoração social que o leve à estigmatização e unem os dois contextos analisados (frente a nomes próprios e frente a possessivos) a princípios comuns, reduzindo-os a um mesmo fenômeno. As autoras assumem que a utilização do artigo definido está atrelada não somente às circunstâncias da definitude em si, mas aos fatores de natureza semântica e pragmática, amplamente vinculada às questões da mudança linguística. As variáveis selecionadas pelo pacote de programas VARBRUL como favorecedoras no uso do artigo definido foram:

- i. *a presença de preposição* – constatou-se forte dependência do artigo em relação à presença de preposições, principalmente as preposições *de, em, para, a*;
- ii. *função sintática* – a posição de sujeito, tópico, objeto e predicativo são as funções mais atuantes no licenciamento do artigo definido; em relação ao nome próprio, a posição de tópico e adjunto adnominal são as que mais favorecem o uso do artigo;
- iii. *grau de familiaridade* – em relação aos possessivos, percebe-se que as relações humanas promovem em menor grau o uso do artigo quando comparado com os objetos; diante de nomes próprios, o grau de familiaridade só demonstrou relevância no Rio de Janeiro;
- iv. *região de origem* – as capitais do Sul e Sudeste empregam com maior frequência o artigo;
- v. *prosódia* – uso do artigo em vocábulos de até duas sílabas.

Callou (1992, p.33), ao analisar uso do artigo definido diante de nome próprio de pessoa, excetuando-se topônimos, descobre que “o uso do artigo diante de nomes próprios se iniciou em um contexto específico [art + det + NP], numa determinada

posição, a de sujeito, e numa determinada estrutura vocabular, isto é, em vocábulo de duas sílabas”. A autora faz as seguintes considerações:

- 1) a frequência de uso do artigo obedece a uma distribuição por região;
- 2) se confirma a ausência categórica de artigos em vocativos, ausência também referida em Silva (1987), quando trata do artigo frente a possessivos. Vale lembrar que esta ausência tem sido atribuída ao alto grau de especificidade dessa função. Poder-se-ia explicar o aumento de frequência em adjuntos adverbiais, de menor grau de especificidade, por raciocínio semelhante, o uso do artigo correspondendo, assim, a um processo compensatório de intensificação do traço [+específico];
- 3) o grau de familiaridade referido pelos gramáticos como condicionador do uso do artigo só se mostrou relevante nos dados do Rio de Janeiro, fato já assinalado no trabalho de Silva (1989), em relação aos falantes de menos grau de escolaridade;
- 4) embora esse uso tenha sido atribuído a uma característica da língua falada, e até mesmo à linguagem popular, verifica-se que na linguagem escrita (literária) do século XX, o uso do artigo nesse contexto apresenta resultados percentuais muito próximos: 34% na escrita e 36% na falada, estando sujeito aos mesmos condicionamentos.

Costa (1998), ao analisar o uso do artigo definido diante de possessivo e de nome próprio de pessoa, selecionou textos com estilos variados, com o intuito de perfazer o período arcaico da língua portuguesa (do século XIII a meados do século XVI). São eles: escritos selecionados por Clarinda Maia, Crônica de Dom Pedro, Diálogo da viciosa vergonha e Diálogo em louvor da nossa linguagem. Os resultados obtidos foram confrontados com os resultados sincrônicos provenientes dos trabalhos de Callou (1992), Silva (1996), Callou e Silva (1997). Desta comparação, Costa (1998) ressalta que verificou mais semelhanças do que diferenças:

TABELA 3: Semelhanças e diferenças no uso do artigo definido diante de nomes próprios e de possessivos – comparação diacrônica e sincrônica – depreendidos de Costa (1998).

	<b>USO DO ARTIGO DEFINIDO DIANTE DE POSSESSIVO</b>	<b>USO DO ARTIGO DEFINIDO DIANTE DE NOME PRÓPRIO</b>
<b>SEMELHANÇAS</b>	a) Favorecimento do uso do artigo definido em estruturas no singular; b) Preposições fomentam o uso do artigo, principalmente as que aceitam contrair-se; c) Os SNs estimulam menos o uso	a) O uso do artigo definido em linguagem escrita é muito baixo; b) O uso do artigo na função de sujeito era muito



	do artigo definido que os SPs; d) A relação humana (parentesco) fomenta menos o uso do artigo definido.	frequente até o século XVII; c) Presença da variável preposição mostrou-se irrelevante.
<b>DIFERENÇAS</b>	a) Os possessivos <i>seu(s)</i> e <i>sua(s)</i> alcançaram índices maiores de ocorrências com o artigo definido.	—

Outros estudos mais recentes retomaram o tema do uso do artigo diante de antropônimos. Mendes (2000) investigou a ausência de artigo definido diante de nomes próprios no português da cidade mineira de Barra Longa; Amaral (2003, 2007) analisou a ausência/presença de artigo definido diante de antropônimos em três localidades de Minas Gerais: Campanha (no Sul de Minas), Minas Novas (no Vale do Jequitinhonha) e Paracatu (no Noroeste do Estado); Almeida Mendes (2006) verificou a ausência/presença do artigo diante de antropônimos nas localidades mineiras de Matipó e Abre Campo; e Alves (2008) tratou da variação ausência/presença de artigo definido diante de antropônimos no português brasileiro, na fala de jovens mineiros de Barra Longa que residem na capital, Belo Horizonte. São pesquisas significativas, que analisaram o uso do artigo definido sob diferentes perspectivas teóricas e que, principalmente, fazem um interessante mapeamento geográfico da frequência do fenômeno.

É importante destacar que esses estudos distam do escopo deste trabalho, já que analisamos a variação no emprego do artigo apenas em SNs não preposicionados com referência definida, como em: *Botô Ø cabeça pra báxo*. Optamos por esse foco em função do paralelismo que existe entre uma variedade do português brasileiro, o português afro-brasileiro, e línguas crioulas de base lexical portuguesa, bem como outras variedades da língua portuguesa formadas em situação de contato, já que o objetivo deste estudo é buscar evidências empíricas que dêem suporte à relação entre o contato linguístico no período colonial do Brasil e a formação das variedades populares do português brasileiro.

Ressaltamos, no entanto, que a leitura dos trabalhos mencionados foi de fundamental importância na preparação das variáveis linguísticas empregadas aqui. Contudo, vamos focalizar a partir de agora a variação no uso do artigo em SNs de

referência definida em variedades do português formadas em situação de contato radical e massivo.

### 3.5 O ARTIGO DEFINIDO NAS LÍNGUAS CRIOULAS DE CABO VERDE E SÃO TOMÉ E PRÍNCIPE

Lucchesi (1993) faz uma análise sobre como funciona o sistema de artigos nos crioulos de base lexical portuguesa no Atlântico, mais especificamente, no arquipélago de Cabo Verde e na maior ilha do arquipélago de São Tomé e Príncipe, a Ilha de São Tomé. O autor procura testar os dois modelos teóricos que explicam o sistema de artigos em línguas crioulas: a **hipótese do bioprograma** (BICKERTON, 1984, 1988, 1989, 1992, 1999) e o modelo que atribui a formação das línguas crioulas à **influência das línguas lexificadoras** (JANSON, 1984).

De acordo com Bickerton (1984, 1988, 1989, 1992, 1999) há um *input* linguístico restrito na mudança de uma língua *pidgin*, bastante instável, para uma língua crioula típica. Segundo o autor, o bioprograma da linguagem existe em todos os seres humanos. No entanto, as crianças nascidas no âmbito de uma comunidade linguisticamente heterogênea, como as que ocorreram em áreas geográficas de colonização baseada na economia de *plantation*, carecem de um modelo linguístico completo para a aquisição da língua materna. Desse modo, o programa biológico atua de forma decisiva na estruturação gramática da língua nativa que assim se forma, o crioulo. A nova língua crioula tem, portanto, como principais agentes formadores as crianças, que fixam as estruturas linguísticas nas lacunas deixadas durante o processo de aquisição precária, transformando, assim, o *input*, antes limitado, em uma língua plena.

A fim de demonstrar que o cérebro humano estaria programado previamente para distinguir determinadas estruturas gramaticais, constatação que permite definir sua teorização como universalista, Bickerton analisou os artigos no crioulo inglês do Havaí. Afirma o autor que o sistema de artigos é uma evidência basilar do bioprograma da linguagem, considerando sua teoria a chave para a apreensão da formação das línguas crioulas e, por conseguinte, da aquisição da linguagem.

As conclusões de Bickerton baseiam-se na sua observação de que no *pidgin* inglês do Havaí os artigos apareciam esporadicamente e de forma imprevisível,

enquanto que, na variedade crioula que se forma a partir desse pidgin, o uso dos artigos é bastante regular e definido. O autor apresenta o seguinte sistema de artigos no crioulo inglês do Haváí (BICKERTON, *apud* LUCCHESI, 1993, p. 82):

- a. Um artigo definido *da*, usado para todos, e só para SNs de referência específica, conhecidos pelo ouvinte;
- b. Um artigo definido *wan*, usado para todos, e só para SNs de referência específica, desconhecidos pelo ouvinte – tipicamente, em caso de primeira menção;
- c. Todos os outros SNs não possuem artigo. Nesse grupo estão: SNs genéricos; SNs num escopo de negação – i.e., claramente não específicos – e casos em que, mesmo que um referente específico exista, a exata identidade desse referente, ou é desconhecida pelo falante, ou é irrelevante para o ponto em questão.

De acordo com Lucchesi (1993, p. 82), a característica mais proeminente nos resultados demonstrados por Bickerton concentra-se na distinção específico/não específico, que é marcada com nitidez e determinada pela oposição entre a ausência do artigo em SNs não específicos<sup>18</sup> e a presença do artigo definido e indefinido em SNs com referencial específico.

Segundo Bickerton, todas as línguas crioulas possuem afinidades gramaticais e, dessa maneira, compartilham o mesmo sistema de artigos: (i) um artigo definido utilizado em SN específico, pressuposto; (ii) um artigo indefinido utilizado em SN específico, não pressuposto e (iii) um morfema zero para SN não específico.

Já o modelo teórico proposto por Janson (1984, *apud* LUCCHESI, 1993, p. 63) rejeita veementemente a teoria do bioprograma, incluindo o sistema de artigos descrito por Bickerton como comum a todas as línguas crioulas. De acordo com Janson, o uso do artigo nessas línguas pode ser esclarecido através da mudança linguística sob a contínua influência das línguas europeias, ou seja, das línguas lexificadoras.

Para o autor, a distinção específico/não específico não encontra respaldo no sistema de artigos das línguas crioulas. Além disso, a questão das peculiaridades que diferenciam o inglês padrão do Haváí e o crioulo inglês do Haváí, no que diz respeito às inovações deste último, é o resultado de um novo sistema de referência do SN:

<sup>18</sup> A ausência de artigo com referência não específica é, para Bickerton (1981, p. 26), uma invenção do crioulo inglês do Haváí.

- a. O numeral para *um* assume a função de artigo indefinido, marcando todo o SN singular que introduz um novo referente;
- b. Um pronome demonstrativo é enfraquecido semanticamente em tal medida que, às vezes, só marca definição, i.e., a qualidade de um ser já conhecido.

As observações de Lucchesi (1993) a partir dos textos narrativos orais de contos populares de Cabo Verde e de São Tomé foram: (i) a inexistência de um dispositivo especial para marcar a distinção específico/não específico; (ii) o uso regular e sistemático do artigo indefinido *um* para marcar SNs específicos não conhecidos pelo ouvinte, tipicamente na primeira menção do SN específico e (iii) o uso de *kel*, como artigo definido de forma irregular, predominando a sua ausência.

No que concerne à proficiência do modelo teórico elaborado por Bickerton, Lucchesi chega à conclusão de que:

[...] o modelo proposto por Bickerton, no qual se destaca a distinção específico/não específico claramente marcada nas línguas crioulas, não se mostra adequado à explicação dos artigos no crioulo caboverdiano e no santomense. Pode-se mesmo questionar a validade de se tomar esta distinção enquanto instrumento de análise, na medida em que ela não dá conta de todas as situações observadas no estudo lingüístico de um texto. (LUCCHESI, 1993, p. 100)

Mais adiante, prossegue com a seguinte informação:

Segundo Bickerton, a distinção específico/não específico nas línguas crioulas se expressaria no plano morfossintático através da ausência do artigo nos SNs não específicos, contra a presença de um artigo nos SNs específicos. Contudo, em todos os textos analisados, só os SNs específicos não pressupostos contaram com a presença regular do artigo indefinido; tanto os SNVS não específicos como os SNs específicos pressupostos não apresentaram uma marca regular, predominando a ausência do artigo nesses SNs. Portanto, não se encontra no crioulo caboverdiano e no santomense o sistema de artigos que “seria virtualmente compartilhado por todos os crioulos” (LUCCHESI, 1993, p. 101)

Já ao tratar do modelo teórico baseado na mudança linguística governado pela influência de línguas lexificadoras, sugerido por Janson (1984), Lucchesi considera suas

afirmações mais “produtivas” e mais adequadas aos resultados encontrados no crioulo de Cabo Verde e de São Tomé. Segundo o autor, “*a explicação do abandono para as línguas crioulas do sistema de artigos nas línguas lexificadoras, como uma consequência da mudança disruptiva, que caracteriza o processo de formação dessas línguas*” é satisfatória, de modo que não encontra posteriores críticas no que diz respeito a essa rejeição e substituição durante o processo de nativização.

No entanto, Lucchesi (1993) adverte que Janson não responde a duas questões indispensáveis para o entendimento sobre o sistema de artigos nas línguas crioulas de Cabo Verde e São Tomé, tornando seu modelo impreciso: (i) a diferença do uso do demonstrativo como artigo em línguas crioulas e em línguas românicas e (ii) o emprego irregular do artigo definido. Desta forma, propõe uma explicação baseada nos princípios da *transparência semântica* e da *simplificação*, na tentativa de dar conta da função do artigo definido nessas línguas. O *princípio da transparência semântica* baseia-se na preferência das línguas crioulas em formação por elementos de significados mais imediatos e diretos; o *princípio da simplificação* baseia-se no cancelamento de redundâncias e ampliação da carga funcional de elementos gramaticais disponíveis. Sendo assim, o uso irregular do artigo definido deve-se à maximização da função do artigo indefinido no sistema de referência do SN em línguas crioulas, como descrito por Lucchesi (1993, p. 101):

Quando um falante introduz um SN marcado por um artigo indefinido em seu discurso, ele oferece ao ouvinte as seguintes informações:

- (i) trata-se de um SN específico na medida em que o artigo indefinido destaca um (uns) indivíduo(s) dos demais da sua espécie;
- (ii) a presença do artigo indefinido também informa ao ouvinte que esse SN não faz parte do seu conhecimento prévio; carrega, pois, nesse momento, a noção gramatical de ‘indefinição’.

Essa marcação do artigo indefinido na primeira menção de um SN propaga-se na cadeia discursiva, de modo que, nas menções seguintes desse SN, o ouvinte pode recuperar sua referência anterior independentemente da presença de artigo definido, podendo ocorrer a sua ausência. Assim, pode-se concluir que o *artigo definido não faz parte do núcleo da gramática da língua crioula*.

Numa outra abordagem do sistema de artigos do caboverdiano, Baptista (2007) mostra que a complexidade desse sistema reside no modo como os determinantes

explícitos (*overt determiners*) relacionam-se com os artigos nulos na marcação da especificidade e não-especificidade dos SNs. A autora enfoca uma característica particular dos determinantes explícitos no crioulo caboverdiano, que consiste no princípio econômico da marcação de plural – determinantes geralmente levam marca de plural enquanto os nomes não o fazem.

Em sua exposição, Baptista contrasta o sistema clássico de artigo do crioulo descrito por Bickerton (1981) com o sistema de artigo caboverdiano. Além disso, oferece um panorama geral das observações de Lucchesi (1993) sobre o sistema de artigo no crioulo de Cabo Verde, inserindo algumas modificações na análise de Lucchesi, com o propósito de salientar toda a complexidade de tal sistema referencial.

Ao contrário de Meintel (1975) e Silva (1985), cujos estudos não reconhecem completamente a existência de um sistema de artigos no crioulo caboverdiano, Baptista afirma que o caboverdiano tem um sistema referencial bastante complexo e considera que os morfemas nulos são uma parte essencial do referido sistema.

Baptista destaca a existência de dois tipos de determinantes utilizados pelos caboverdianos: (i) o artigo indefinido *un* (sing.)/ *uns* (pl.), que, além de introduzir um novo referente, figura muitas vezes como mais próximo de um quantificador do que de um artigo; e (ii) o artigo definido *kel* (sing.)/ *kes* (pl.), também empregado como demonstrativo. Ambos os artigos apresentam marca de número, mas não mostram nenhuma concordância de gênero, como ilustram os exemplos abaixo:

**a) Artigo indefinido:**

‣ [+ plural]; [- gênero]

(31) *Un skritor skrebe un libru pa un mudjer.*

*‘Um escritor escreveu um livro para uma mulher’.*

*Uns skritor skrebe uns libru pa uns mudjer.*

*‘Uns escritores escreveram uns livros para umas mulheres’.*

‣ [+ indefinido]

(32) *N odja un omi.*

*‘Eu vi um homem’.*

‣ [+quantificador]

- (33) *N odja **un** omi so.*  
 ‘Eu vi um homem somente’.

**b) Artigo definido:**

- (34) ***Kel** skritor da **kel** minina **kel** libru.*  
 ‘O escritor deu o livro à menina’.  
***Kes** minina da **kes** skritor **kes** kaneta.*  
 ‘As meninas deram as canetas aos escritores’.

Ainda ao tratar dos artigos indefinidos, Baptista menciona três mecanismos para marcação de plural nos SNs: (i) o próprio nome leva a marca de plural; (ii) o nome é precedido do quantificador indefinido *uns*; e (iii) em alguns contextos os SNs podem aparecer sem marca de número e sem determinante e ainda assim a ideia de plural indefinido é alcançada na interpretação. Abaixo, os respectivos exemplos:

**a) Marca de plural no nome:**

- (35) *N odja omis sintadu riba di kaza.*  
 ‘Eu vi homens sentados em cima da casa’.

**b) Precedido de quantificador indefinido *uns*:**

- (36) *N odja **uns** omi sintadu riba di kaza.*  
 ‘Eu vi uns homens sentados em cima da casa’.

**c) Sem marca de plural no nome e sem artigo:**

- (37) ***Omi** ten pe di katxor. (expressão idiomática)*  
 ‘Homens estão sempre fora (de casa)’.

De acordo com Baptista (2007, p. 18) a indefinitude e a definitude podem também ser expressas por um artigo nulo, como no exemplo abaixo:

a) **Artigo nulo**

- (38) *N odja Ø omis na taberna.*  
*‘Eu vi homens na taverna’ ou*  
*‘Eu vi os homens na taverna’.*

Baptista (2007, p. 28) discorda das afirmações de Lucchesi quanto ao crioulo caboverdiano em duas proposições. Primeiro, a autora não observou em seu estudo ser necessário marcar o SN com artigo indefinido para estabelecer referencia definida. Utiliza o seguinte exemplo para confirmar sua afirmação:

- (39) *Prezidenti papia di independensia di Kabu Verdi.*  
*‘O presidente falou da independência de Cabo Verde’.*

Segundo Baptista, não há a introdução do artigo na menção anterior, no entanto, tanto *presidente* e quanto *independência* são referenciais definidos.

A segunda divergência refere-se à omissão de Lucchesi a respeito da leitura específico e não-específico de morfemas nulos. De acordo com a autora, nomes no singular, não podem ter uma interpretação específica (indefinida), ou seja, somente nome com plural marcado sem artigo pode propiciar uma interpretação específico e indefinido baseado no contexto. A exceção para que ocorra interpretação de indefinidade com nomes no singular sem artigo acontece somente com expressões idiomáticas:

- (40) *El tem kurason di pedra.*  
*‘Ele tem um coração de pedra’.*

Baptista (2007, p. 30) sumariza assim o sistema referencial do crioulo de Cabo Verde:

Indefinido		Definido	
Singular	Plural	Singular	Plural
um/*Ø (exceto exp. idiom.)	uns/ø	kel/ø	kes/ø

QUADRO 3: Sistema referencial do crioulo de Cabo Verde, segundo Baptista (2007).



Dessa forma, as discordâncias com Lucchesi (1993) estão relacionadas aos SNs definidos na presente pesquisa como *definido pragmaticamente* e *definido anaforicamente*<sup>19</sup>. Provavelmente, Lucchesi não considerou essas categorias em seu trabalho nos crioulos de Cabo Verde e São Tomé, pois analisava textos narrativos orais de contos populares, em que nem sempre é possível uma interpretação pragmática das funções referenciais.

Sumariamente, segundo Baptista (2007), o crioulo caboverdiano faz um amplo uso de nomes nus (isto é, SNs sem artigo) e que essa característica decresce tanto no português europeu quanto no português brasileiro. Neste último, no entanto, existem especificidades compartilhadas com o crioulo de Cabo Verde no que se refere à possibilidade da não utilização do artigo.

### 3.6 O USO DO ARTIGO DEFINIDO NO PORTUGUÊS AFRO-BRASILEIRO

Há notáveis diferenças entre o sistema de artigo dos crioulos de Cabo Verde e São Tomé e Príncipe, de um lado, e o sistema de artigos do português europeu e do português brasileiro culto, de outro, conquanto se observem também diferenças entre esses dois últimos. Por outro lado, as variedades populares do português brasileiro, particularmente a fala de comunidades rurais afro-brasileiras isoladas (muitas delas oriundas de antigos quilombos), exibem diferenças no uso do artigo que afastam essas variedades do padrão culto e as aproximam das variedades crioulizadas do português. Esses paralelos entre o comportamento gramatical do chamado *português afro-brasileiro* (LUCCHESI; BAXTER; RIBEIRO, 2009) e as línguas crioulas de base lexical da África podem fornecer indícios valiosos para a compreensão do papel que o contato entre línguas desempenhou na formação histórica de variedades do português no Brasil. Alguns trabalhos anteriores, nos quais esta análise se baseia, procuraram explorar esses paralelos no que concerne sistemas de artigo.

Nessa linha, Baxter e Lopes (2004, 2009), analisando a presença variável do artigo definido em SNs de referência definida na comunidade rural afro-brasileira de Helvécia-BA, constataram que a ausência do artigo é mais frequente na fala de

---

<sup>19</sup> Os conceitos da variável referencialidade, usados nesta dissertação, serão explanados no capítulo 4.

indivíduos mais velhos e que o seu uso, considerado padrão, é predominante na fala de indivíduos que são profissionalmente ativos. Segundo os autores, esse comportamento aponta para a aquisição do uso do artigo em comunidades formadas em situação de contato entre línguas, onde, no passado, o uso do artigo teria sido mais raro.

As variáveis independentes analisadas foram: (i) faixa etária do informante; (ii) animacidade do substantivo; (iii) contabilidade do substantivo; (iv) referência do SN; (v) número do SN; (vi) presença de outros modificadores no SN; (vii) determinante no SN anafórico prévio e (viii) função sintática. No trabalho de 2009, Baxter e Lopes refazem a análise quantitativa no GOLDVARB-X, além, claro, de revisarem a codificação.

Foram considerados onze informantes em quatro faixas etárias: faixa 2 (35 a 45 anos); faixa 3 (60 a 70 anos); faixa 4 (+80 anos) e um informante com mais de 100 anos. O resultado obtido na única variável extralinguística analisada pelos autores apresentou, diacronicamente, o perfil aquisicional do fenômeno, com distribuição geracional bem marcada, como podemos observar no gráfico 3:

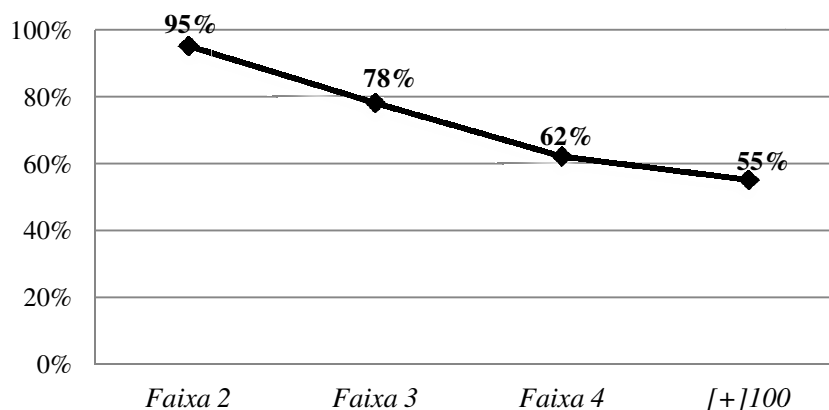


GRÁFICO 3: Frequência do artigo definido segundo a variável faixa etária, segundo Baxter e Lopes (2004, 2009).

Esse resultado confirma o condicionamento social identificado na análise de outros aspectos da gramática na mesma comunidade<sup>20</sup> e nos permite afirmar que o uso

<sup>20</sup> Para maiores informações sobre o resultado dos condicionamentos sociais em comunidades afro-brasileiras, ver LUCCHESI, Dante, BAXTER, Allan; RIBEIRO, Ilza (Orgs.). O português afro-brasileiro. EDUFBA: Salvador, 2009. p. 539.

do artigo definido no português do Brasil seria reduzido em comunidades mais diretamente afetadas pelo contato entre línguas.

Sobre os aspectos semânticos, destacam-se os resultados encontrados nas variáveis **referencialidade** e o **número do SN**. Na primeira, foram considerados dois tipos de referencialidade: [+único] – quando o nome tem conhecimento compartilhado pelos interlocutores, e [-único] – quando o nome não tem referencial identificável pelos interlocutores. O exame dessas propriedades revelou que SNs com referencial mais nitidamente identificável, como em “*lua tá traz do sol uma bandinha*”<sup>21</sup>, são desfavorecedores do uso do artigo definido. Quanto ao número do SN, singular ou plural, os resultados encontrados por Baxter e Lopes (2004, 2009) revelam que o SN plural favorece intensamente a utilização do artigo definido em 93% das ocorrências.

Ressalta-se que o estudo dos autores sobre o artigo definido considerou sete funções sintáticas. Os resultados cotejados assinalaram que o uso do artigo definido é amplamente favorecido pela função de *adjunto adverbial*, com .68 de peso relativo. Esse resultado é explicado pelos autores devido à presença da preposição que motivaria o uso do artigo definido. As funções de sujeito e de objeto direto não apresentaram uma assimetria na distribuição dos SNs sem determinante de referência definida nos dados observados por Baxter e Lopes (2004, 2009).

Já Ribeiro (2010), em comunicação realizada na UFBA, no VI encontro da Associação Brasileira de estudos Crioulos e Similares (ABECS), descreveu o sistema de definitude e de referencialidade na gramática de uma afro-descendente com mais de 100 anos de idade, membro da comunidade de Helvécia-BA. O objetivo da autora consistiu em comparar o sistema de determinação do português brasileiro com estudos realizados sobre o mesmo fenômeno em línguas crioulas<sup>22</sup> e demonstrar que houve no Brasil *inputs* diversificados no processo de formação das variedades populares do português brasileiro.

A autora descreve os determinantes definidos e indefinidos utilizando os seguintes conceitos de referencialidade e definitude, com base em Givón (1978), Lyons (1999), Laca (1999) e Baptista (2007):

<sup>21</sup> Exemplo extraído de Baxter e Lopes, (2009, p. 323)

<sup>22</sup> A autora compara seus resultados com os de Baptista (2007).

(1) **Referencialidade** – abrange a intenção do falante de indicar que a expressão nominal não é vazia em relação à identificabilidade ou à existência do referente em duas situações:

a) No universo particular do discurso:

(41) *Maria comprou **uma casa nova**.*

b) No uso de um nominal não-referencial:

(42) *Maria está procurando **uma casa nova** para comprar.*

(2) **Genérico** – abrange o referencial do falante a uma espécie, classe, tipos e fatos, excetuando-se objetos e membros individuais específicos da espécie ou classe. No português, o uso genérico pode ocorrer com:

a) Determinante indefinido:

(43) ***Um cão** tem quatro pernas.*

b) Determinante definido:

(44) ***O cão** tem quatro pernas.*

c) Determinante  $\emptyset$ :

(45) ***Cães** têm quatro pernas.*

(3) **Expressão definida** – relaciona-se ao identificável, ao que é familiar na esfera interacional entre o ouvinte e o falante.

(46) *Maria comprou **a casa que queria**.*

(4) **Expressão indefinida** – pode significar uma entidade específica ou familiar do falante; ou pode significar um membro arbitrário, não-específico, da classe descrita pela expressão nominal:

a) Determinante indefinido específico:

(47) *Eu comprei **um** carro.*

b) Determinante indefinido não-específico:

(48) *Passe-me **um** livro.*

(5) **Traço [+/- específico] com determinante definido** – a depender das leituras de *intenção* ou de *extensão*:

a) Intenção [- específico]:

(49) *Joana quer entregar o prêmio ao **vencedor** – assim, ela tem de esperar a competição terminar.*

b) Extensão [+ específico]:

(50) *Joana quer entregar o prêmio ao **vencedor** – mas ele não quer recebê-lo.*

Ribeiro (2010) deixa claro que as operações de referencialidade e definitude apresentadas dependem de *múltiplas pistas linguísticas*, ligadas à morfossintaxe, e *pistas conceptuais*, ligadas à pragmática e ao conhecimento de mundo. Sobre as informações encontradas pela autora, enfatizaremos o determinante definido, realizado ou não com o artigo definido.

Ao comparar os dados de Baptista (2007) sobre os nomes nus com os dados encontrados no inquérito da informante de Helvécia, seguindo os conceitos de referencialidade e definitude pré-estabelecidos, conclui que se aproximam bastante nas seguintes situações de uso: genérico, indefinido [+específico] plural, indefinido [-

específico] plural, indefinido [+específico] singular, indefinido [-específico] singular, definido [+específico] singular e definido [-específico] singular.

Outra particularidade relevante encontrada por Ribeiro é a de que o determinante  $\emptyset$  é notoriamente excluído das realizações plurais definidas, provavelmente porque informantes pertencentes às comunidades afro-brasileiras marcam morfologicamente o plural nos determinantes.

Ao identificar esses pontos de contato entre a gramática do português afro-brasileiro e a das línguas crioulas de base lexical portuguesa da África, essas análises indicam as pistas que devem ser perseguidas por uma análise variacionista que deseje contribuir com o esclarecimento do papel que o contato entre línguas desempenhou na formação dessas variedades do português brasileiro. E essas pistas serão perseguidas ao longo da análise que constitui o cerne desta dissertação.

## 4 ANÁLISE DOS DADOS

Apresentaremos, neste capítulo, os resultados da análise variacionista do emprego do artigo definido na comunidade afro-brasileira de Helvécia, os quais foram submetidos ao pacote de programas estatístico VARBRUL. Descrições e interpretações das variáveis linguísticas e extralinguísticas controladas nesta análise serão comparadas, quando necessário, com o que já foi obtido em análises anteriores.

É preciso ressaltar que nem sempre o programa matemático distribui um resultado estável e equilibrado (problema analítico). Desta maneira, fazer outras rodadas do Programa VARBRUL com configurações alternativas dos grupos de fatores é um recurso válido para avaliar a possibilidade de desequilíbrio na distribuição dos dados (GUY; ZILLES, 2008, p. 177). Dentro dessa perspectiva, confrontamos o comportamento das variáveis independentes com relação à variável dependente, fazendo algumas rodadas e adequações, a fim de eliminar sobreposições e/ou fatores categóricos. Além disso, fizemos alguns cruzamentos, imprescindíveis na observação de interferências entre dois ou mais grupos de fatores, além de amalgamarmos algumas variantes<sup>23</sup>.

Assim sendo, serão apresentados, neste capítulo, o número de dados, as percentagens, os cruzamentos realizados e os pesos relativos com os respectivos níveis de significância de cada fator favorecedor ou que mereça atenção especial devido sua importância no contexto social.

### 4.1 VARIÁVEL DEPENDENTE

Como já mencionado no capítulo 1, na seção 1.4, trabalhamos com uma variável binária, que consiste em analisar a variação ausência/presença de artigo definido em SNs na fala de informantes da comunidade remanescente quilombola situada no distrito de Helvécia-BA. Nosso objetivo é demonstrar a existência de sistematicidade do fenômeno em foco e contribuir com outros estudos que revelaram indícios de como o

---

<sup>23</sup> Todos os ajustes aplicados serão explicados no decorrer do presente capítulo.

contato entre línguas forneceu elementos significativos para a atual representação do português brasileiro.

Totalizamos 634 ocorrências, exaustivamente levantadas, que passaram por três rodadas no programa quantitativo VARBRUL, em decorrência dos procedimentos de cruzamentos dos fatores sobrepostos e de amalgamações. É importante lembrar que o elemento linguístico variável em questão desencadeou ampla preocupação em sua coleta, bem como os contextos de sua ocorrência, devido à sua fragilidade – é um fenômeno fácil de perder ou de ser acrescentado nas transcrições grafemáticas, apesar de todo aparato (incluindo programas de áudio e de transcrição) e da mão-de-obra capacitada que o Projeto *Vertentes* possui e vem aprimorando nos últimos anos.

Abaixo, ilustrados em (19) e (20), exemplos da variação analisada, excluindo-se todos os casos descartados e elucidados no item 1.3.1, sendo  $\emptyset$  o símbolo escolhido para estabelecer a ausência do artigo definido em SNs:

(19) *Quande **a maré enche, enche tudo!*** (HV-19)

(20) *Inda mais você encontrá  $\emptyset$  **vestido!*** (HV-19)

Os exemplos foram retirados da fala vernácula da mesma informante com 103 anos de idade quando foi gravada a entrevista na década de noventa. Salientamos que os SNs em destaque são constituídos de nomes com referente definido, identificável pelo ouvinte e pelo falante. Em (19) observamos o emprego do artigo definido e em (20) a ausência deste determinante.

A análise quantitativa dos dados gerados pelo pacote de programas VARBRUL, revelou que a presença de artigo definido em SNs corresponde a 59% em detrimento de sua ausência que corresponde a 41% dos casos, como se pode ver no gráfico a seguir.



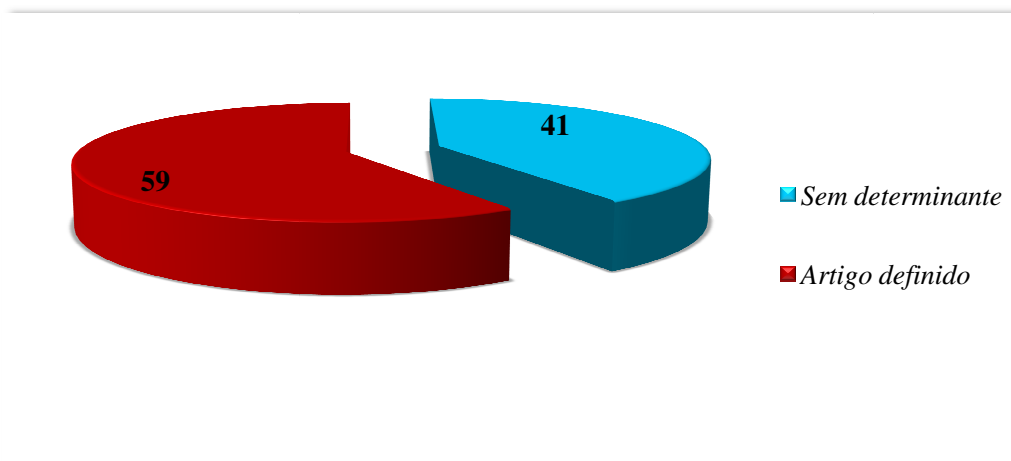


GRÁFICO 4: A presença do artigo definido em SNs na comunidade de Helvécia-BA.

Percebemos, pelos dados computados, a existência de um número considerável de ocorrências com SNs sem o artigo (*nomes nus*), embora seja maior o emprego do determinante definido. Esse resultado será analisado em função da hipótese de que o artigo definido é a variante inovadora, proveniente da influência do padrão urbano, e de que a ausência do artigo em SNs é uma característica proeminente da relação de contato linguístico por que passaram as variedades populares do português brasileiro durante a sua formação. Não podemos esquecer que outros estudos sociolinguísticos também verificaram na comunidade de Helvécia a manutenção de traços peculiares motivados por um contato entre línguas mais intenso (LUCCHESI; BAXTER; RIBEIRO, 2009).

Assim, o resultado ratifica trabalhos anteriores (RIBEIRO, 2010; BAXTER; LOPES, 2004, 2009) e corrobora com evidências empíricas a real possibilidade de um estágio de crioulização do português brasileiro em tempos de sua colonização. O sistema de marcação de definitude, dessa maneira, é um traço linguístico que merece aprofundamento, pois provavelmente faz parte do conjunto de mudanças ocorridas na morfologia e na sintaxe do português do Brasil em decorrência do contato linguístico.

## 4.2 VARIÁVEIS LINGUÍSTICAS SIGNIFICANTES

### 4.2.1 Referencialidade

A variável em questão foi a mais complexa e a que mais exigiu leitura dos inquiridos, a fim de buscar pistas do contexto linguístico e, assim, estabelecer o nível de referencialidade. Mas, seria impossível analisar o artigo definido sem considerar sua propriedade mais peculiar.

Em nossas observações, ficou bastante claro que o índice de definitude e de referencialidade não estão atrelados a presença do artigo definido como postula a tradição. Absolutamente, como bem afirma Lyons (1999, p. 52), é impreciso considerar que o papel do artigo seja simplesmente definir o SN. Os nomes apresentam características semânticas que, quando relacionadas ao contexto, indicam a variabilidade do emprego do artigo definido. Desta maneira, o estudo da variável referencialidade busca identificar quais propriedades ancoram o processo variável em estudo.

As definições assumidas por linguistas e estudiosos<sup>24</sup> sobre o assunto foram essenciais para a obtenção das noções de referencialidade e definitude empregadas aqui. Ressaltamos, no entanto, que os conceitos aplicados tiveram origem nas observações dos fatos linguísticos ao longo desta análise<sup>25</sup>.

A escolha em definir parâmetros de referencialidade sem adotar um autor específico para direcionar a investigação do fenômeno teve dois motivos: (i) haver na literatura um amplo número de conceitos sobre referencialidade/definitude e igualmente, inúmeras divergências acerca dessas definições e (ii) as acepções existentes nem sempre supriam as necessidades para uma averiguação mais precisa do fenômeno variável dentro da comunidade analisada. Sendo assim, foram feitas algumas adaptações dos modelos existentes.

Asseguramos que fazer apreciações sobre a referencialidade sem uma análise pormenorizada do contexto em que se deu a interação verbal causa equívoco e generalizações que não demonstram a realidade da determinação referencial. É necessário, ainda, ressaltar, a existência de exemplos de complicado exame, o que nos fez descartar algumas ocorrências por possuírem interpretação intrincada no diagnóstico de referencialidade. Mesmo com indícios linguísticos e contextuais, o processamento de algumas informações suscitou imprecisões ou a possibilidade de mais de uma leitura.

<sup>24</sup> Givón (1978), Lyons (1999), Laca (1999), Baptista (2007) e Ribeiro (2010).

<sup>25</sup> As definições foram estabelecidas sob a supervisão do professor Dante Lucchesi e pela contribuição significativa da professora Ilza Ribeiro.

Para gerenciar o nível de referencialidade, instituímos cinco condições de referência, a fim de atingir um resultado mais esclarecedor quanto à sua influência. Seguem abaixo as definições assumidas e seus respectivos exemplos:

a) **Definido anaforicamente** – quando o informante retoma o SN em uma oração imediatamente anterior usando o mesmo significante. Relacionado ao identificável, uso associativo.

(21) *O pai dele tomém é confortado, o pai dele tem um gadinho aí.* (HV- 02)

(22) *O pezim do moleque é tamaim... Mas o pezinho do moleque era des' tamaim assim...* (HV- 08)

b) **Definido pragmaticamente** – quando há possibilidade de determinar o SN a partir de especificações do contexto interacional. Nos exemplos que se seguem, podemos identificar o sintagma nominal através do conhecimento compartilhado pelos interlocutores durante a narrativa.

(23) *Depois que eu vi caso mau mermo, os médico num deu a duração pra tirá.*  
(HV- 01)

(24) *O olho dele é do tamanho de um farol!* (HV- 08)

c) **Genérico** – SNs que evidenciam a característica de serem elementos representativos de um determinado grupo, espécie ou classe. Encaixam-se também nessa situação objetos e fatos.

(25) *Hum! Os menino d'agora tudo achô coisa bom.* (HV- 13)

(26) *A gente serrava madêra.* (HV- 13)

(27) *Ele invade a terra alhêa.* (HV- 20)

d) **Existencial** – SNs cujo nível de referencialidade é genérico, mas estão inseridos em uma narrativa específica, em que o falante situa temporalmente os eventos narrados.

(28) *Nós rompeu capuêra virge...* (HV- 19)

(29) *Voltô, comprô casa em Têxera!* (HV- 20)

e) **Entidade única** – nomes com referente único, nitidamente compartilhado pelos interlocutores em decorrência do conhecimento de mundo.

(30) *Passei carnaval lá...* (HV- 03)

(31) *...lua tá 'trás do sol uma bandinha.* (HV- 20)

A tabela 4 retrata a quantificação dos resultados, excluindo-se a variante entidade única por apresentar poucas ocorrências:

TABELA 4: A presença do artigo definido em SNs, segundo a variável *nível de referencialidade*.

<b>Referencialidade</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Definido anaforicamente</i>	46/62	<b>74</b>	<b>.73</b>
<i>Definido pragmaticamente</i>	183/248	<b>74</b>	<b>.63</b>
<i>Existencial</i>	40/83	48	.51
<i>Genérico</i>	98/227	43	.29
<b>TOTAL</b>	<b>367/620</b>	<b>59</b>	<b>---</b>

*Log likelihood = -299.972*      *Significância = .037*      *[Rodada 2]*

A seleção da variável nível de referencialidade foi catalogada como estatisticamente significativa pelo programa quantitativo. Podemos perceber pelos resultados que o maior uso do artigo definido está relacionado ao nível de referencialidade *definido anaforicamente*, com .73 de peso relativo, e *definido pragmaticamente* com .63 de peso relativo. Ambos com percentual de 74%, confirmando que o conhecimento compartilhado dos interlocutores favorece o emprego do artigo definido nos SNs.

Por outro lado, os níveis *existencial* e *genérico*, sobretudo o último, desfavorecem o uso do artigo definido com .51 e .29, respectivamente.

#### 4.2.2 Função sintática do SN

Foram considerados dois fatores na variável *função sintática do SN*, a fim de averiguar que contexto é mais suscetível à realização do artigo definido: *sujeito* e *objeto direto*. Abaixo seguem dados extraídos do *corpus* analisado:

##### a) *Sujeito*

(32) *Rapai, Ø lacraia dá sim, mas é bem difícil...* (HV- 02)

(33) *O sangue já tinha esfriado.* (HV- 03)

##### b) *Objeto direto*

(34) *Depois que ele lá agessô a perna, ficô tudo certo.* (HV- 03)

(35) [...] *mas quero tomá Ø remédo lá.* (HV- 22)

A tabela 5 apresenta o resultado desta análise:

TABELA 5: A presença do artigo definido em SNs, segundo a variável *função sintática*.

Função sintática SN	Nº. de ocorrências/total	%	Peso relativo
<i>Sujeito</i>	195/262	74	.68
<i>Objeto Direto</i>	172/358	48	.37
<b>TOTAL</b>	367/620	59	---

*Log likelihood = -299.972*      *Significância = .037*      *[Rodada 2]*

A diferença sintática com relação à ausência ou à presença de artigo definido é bastante acentuada tanto em termos percentuais quanto em termos de peso relativo. Podemos dizer que os resultados ratificam nossa hipótese inicial de que os argumentos externos favorecem mais o uso do artigo nos SNs. No caso analisado, a função de *sujeito* com 74% das ocorrências e .68 de peso relativo. O argumento interno, objeto direto, inibe a utilização do artigo com .37 de peso relativo e 48% dos casos.

Verifica-se, assim, que a **posição argumental** atua favorecendo o uso do artigo definido em SNs periféricos ao sintagma verbal (SV). No item 4.4 serão apresentados os cruzamentos da variável *função sintática* com a variável *referencialidade*.

#### 4.2.3 Número do SN

Essa variável tem atuado amplamente como condicionante na realização do artigo definido, pois são eles que marcam o plural do SN, já que nesse dialeto não se verifica concordância nominal. Alexandre e Hagemeijer (2004 *apud* BAXTER; LOPES, 2004), sobre o crioulo de São Tomé, e Baptista (2007), sobre o crioulo de Cabo Verde, também observaram a mesma tendência.

##### a) *Plural*

(36) *Não! Como pode tomá os remédio e andá na sereno!* (HV-19)

(37) *Eu passo os óio assim, ela tá em pé assim.* (HV- 19)

##### b) *Singular*

(38) *...rala no motô, o motô é a óleo, a gente compra o óleo.* (HV- 01)

(39) *...ó, o calo que me deu aqui, ó.* (HV- 01)

(40) *Ó, Ø perna também ficô. Oh, meu Deus!* (HV- 03)

Podemos visualizar os resultados da variável número do SN na tabela 6:

TABELA 6: A presença do artigo definido em SNs, segundo a variável *número do SN*.

Número do SN	Nº. de ocorrências/total	%	Peso relativo
<i>Plural</i>	46/51	<b>90</b>	<b>.90</b>
<i>Singular</i>	321/569	56	.45
<b>TOTAL</b>	367/620	59	---

*Log likelihood = -299.972*      *Significância = .037*      *[Rodada 2]*

O resultado da tabela 6 ratifica trabalhos anteriores realizados na mesma comunidade (BAXTER e LOPES, 2004, 2009; RIBEIRO, 2009) e trabalhos realizados em línguas crioulas de base portuguesa (LUCCHESI, 1993; BAPTISTA, 2007), apresentando o plural do SN como fator altamente favorecedor da presença do artigo com peso relativo de **.90**. Já nos SNs no singular, observa-se o uso do artigo em um nível próximo ao da neutralidade (56% frente a uma frequência geral de 59%). O resultado do plural está certamente associado a não utilização de marca de plural nos nomes, uma peculiaridade da comunidade de Helvécia. Geralmente os informantes marcam o morfema ‘s’ de plural apenas no determinante. Portanto, a realização do artigo é crucial para a inserção da marca gramatical de plural.

#### 4.2.4. Determinante na menção anterior

Este fator tem demonstrado um comportamento relevante no condicionamento de diversos fenômenos linguísticos variáveis, sendo denominado **paralelismo discursivo** (SCHERRE; NARO, 1993). Segundo esse princípio, o falante tenderia a repetir sua opção formal dentro de uma mesma cadeia discursiva.

A análise dessa variável indica se o antecedente mais próximo influencia a escolha de uma das variantes no SN seguinte correferente, verificando, assim, se a repetição da mesma forma produzida anteriormente gera um efeito sequencial. Ao mesmo tempo, examina como se processa a interação entre o artigo definido ou a sua ausência com outros determinantes utilizados previamente.

Foram considerados cinco fatores relacionados à correferência prévia, listados a seguir:

a) *Artigo definido na ocorrência anterior*

(41) *Eu disse: o gesso já tá bom de tirá. Aí tirô **o gesso grande**, aí botô aqui, ó!*  
(HV- 03)

b) *Demonstrativo na ocorrência anterior*

(42) *Depois o moleque vêi de nôte... bateu esse tambô, **o tambô** aí embaixo...*  
(HV- 08)

c) *Sem menção anterior* – quando não há antecedente com o mesmo elemento.

(43) *... $\emptyset$  **grávida** pá ganhá neném em casa dá trabalho.* (HV- 04)

d) *Artigo indefinido na ocorrência anterior*

(43) *Um menino que ia passano na hora, depois descobriu que era **o menino de Rosa**.* (HV- 04)

e) *Sem determinante na ocorrência anterior*

(44) *Sempre a gente compra  $\emptyset$  **remédio** em farmácia,  $\emptyset$  **remédio casêro mesmo**.* (HV- 04)

Ressaltamos que o fator *possessivo na ocorrência anterior*, inicialmente computado (apenas três ocorrências no total), foi retirado da análise final. A tabela 7 aponta os resultados estatísticos para o uso do artigo definido, conforme a utilização do determinante na menção anterior:



TABELA 7: A presença do artigo definido em SNs, segundo a variável *determinante na menção anterior*.

<b>Det. na menção anterior</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Artigo definido</i>	70/89	<b>79</b>	.55
<i>Demonstrativo</i>	3/4	75	.56
<i>Sem menção anterior</i>	276/451	61	.56
<i>Artigo indefinido</i>	3/6	50	<b>.63</b>
<i>Sem determinante</i>	12/67	18	.12
<b>TOTAL</b>	<b>364/617</b>	<b>59</b>	<b>---</b>

*Log likelihood = -299.972*      *Significância = .037*      *[Rodada 2]*

O resultado apontado na tabela acima vai ao encontro da hipótese inicialmente estabelecida de que a presença do artigo definido no contexto precedente favorece a reiteração do mesmo na ocorrência subsequente. Isto se revela pelo percentual de 79% das ocorrências, seguido do demonstrativo. No estudo realizado por Lucchesi (1993, p. 81) a referência anafórica engatilhada pelo artigo definido prepondera a retomada do mesmo determinante, mas também a sua ausência. Quando o SN anterior é encetado pelo artigo indefinido, o autor registra que há o favorecimento do artigo definido no SN posterior. Essa é uma das discordâncias mencionadas por Baptista (2007).

A afirmação no estudo de Lucchesi (1993) não definiu de forma explícita a diferença entre *definido pragmaticamente* e *definido anaforicamente*. Por outro lado, ao visualizarmos a tabela 4, observamos que há um efeito sequencial registrado pelo peso relativo de .63 nas interações com o artigo indefinido na menção anterior, condicionando a realização do artigo definido no SN posterior. No entanto, o pequeno número de ocorrências com artigo indefinido na menção anterior dificulta conclusões definitivas.

#### 4.2.5 Contabilidade do núcleo do SN

Com a análise da variável contabilidade do núcleo do SN, pretendemos verificar se há uma relação entre a noção de massa e a empregabilidade do artigo definido. Segundo Müller e Paraguassu-Martins (2007, p.79) “*a distinção contável-massivo enquanto um fenômeno lexical deve ser analisada à parte da distinção contável-massivo enquanto um fenômeno contextual*”. Optamos por considerar a abordagem

contextual, que considera a contabilidade como um fenômeno linguístico adjunto ao sintagma (ALLAN, 1980, *apud* MÜLLER; PARAGUASSU-MARTINS, 2007, p. 74). Nos exemplos (a) *João comprou vinho* e (b) *João comprou dois bons vinhos*, o nome *vinho* pode ser usado tanto como incontável como em (a) quanto como contável em (b). Salientamos, no entanto, que para diagnosticar a contabilidade não se faz imprescindível a presença de um quantificador no SN.

Sendo assim, a sentença (45) mostra um *nome contável* – conceitualização individualizada; e a sentença (46) um *nome incontável* – com conceitualização não-individualizada.

a) *Contável*

(45) *Ali, você ligou o motô, lá 'cê vai e faz.* (HV- 01)

(46) *Ele sortô Ø tartaruga, mas Ø tartaruga morreu.* (HV- 08)

(47) *...eu passo os dedos assim, ela tá em pé assim.* (HV- 19)

b) *Incontável*

(48) *Quand' eu ia jogá Ø veneno no tomate...* (HV- 08)

(49) *Quano a terra tá molhada é mais fácil.* (HV- 04)

(50) *...a gente pranta Ø amendoins no mês.* (HV- 12)

Chegam-se aos seguintes resultados apresentados na tabela 8:

TABELA 8: A presença do artigo definido em SNs, segundo a variável *contabilidade do núcleo do SN*.

<b>Contabilidade</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Contável</i>	334/545	<b>61</b>	<b>.52</b>
<i>Incontável/massivo</i>	33/75	44	.35
<b>TOTAL</b>	367/620	59	---

*Log likelihood = -299.972*      *Significância = .037*      *[Rodada 2]*

O núcleo do SN quando contável favorece o uso do artigo com .52 de peso relativo, enquanto nomes incontáveis desfavorece-o com .35. Da análise desse fator, percebemos também que não há uma relação entre a ausência de morfologia de número e nomes massivos e contáveis.

### 4.3 VARIÁVEIS EXTRALINGUÍSTICAS SIGNIFICANTES

#### 4.3.1 Faixa etária

A comparação da linguagem de indivíduos de faixas etárias distintas é capaz de revelar estágios diferentes de uma língua, refletindo uma mudança em curso ou uma variação estável (Cf. LABOV, 1972). Assim, o exame dessa variável é bastante pertinente para verificar um estágio de mudança.

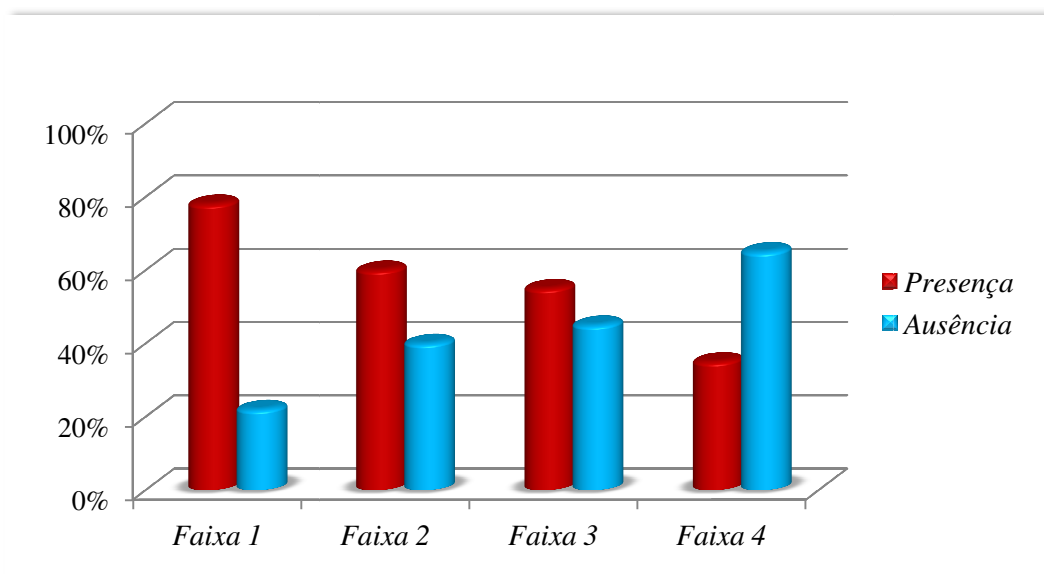
Baxter e Lopes (2004, 2009) evidenciaram na distribuição geracional um nítido perfil aquisicional do artigo, observando uma mudança em direção a variante padrão. Desta maneira, também nos orientamos pela mesma hipótese: o crescente aumento do uso do artigo definido na *Faixa I*, decrescendo na medida em que aumenta a idade dos informantes.

O estabelecimento do perfil diacrônico do fenômeno pode revelar se realmente houve no passado dessa comunidade um processo inicial de crioulização leve que teria servido de base para a aquisição do português como língua primeira. Chegam-se aos seguintes resultados mostrados na tabela 9 e, para melhor visualização, segue o gráfico 5:

TABELA 9: A presença do artigo definido em SNs, segundo a variável *faixa etária*.

Faixa etária	Nº. de ocorrências/total	%	Peso relativo
<i>Faixa I</i>	174/224	<b>78</b>	<b>.69</b>
<i>Faixa II</i>	91/152	60	.55
<i>Faixa III</i>	56/102	55	.35
<i>Faixa IV</i>	54/156	35	.28
<b>TOTAL</b>	375/634	59	---

*Log likelihood = -299.972*      *Significância = .037*      *[Rodada 2]*

GRÁFICO 5: Presença/ausência do artigo definido em SNs com relação à *faixa etária*.

De imediato, observa-se o incremento do uso do artigo definido por falantes mais jovens com 78% das ocorrências e .69 de peso relativo. Os dados confirmam as expectativas de que falantes mais velhos, *Faixa IV*, empregam menos o artigo definido (peso relativo de .28). Assim, de acordo com dados desta pesquisa, quanto mais jovem o falante, maior a incidência do emprego da variante padrão, atestando a existência de um contínuo aquisicional do artigo dentro da comunidade de Helvécia.

O quadro bastante claro dos resultados confirma nossa hipótese inicial e corrobora com trabalhos anteriores de que os mais jovens por razões situacionais, como o de estarem mais inseridos no mercado de trabalho e terem maior acesso à instituição escolar formal, lideram o processo de implementação da variante padrão. Os mais velhos, por sua vez, conservam os vestígios que *testemunham os processos de reestruturação morfossintática* pretéritos nas variedades populares do português brasileiro (BAXTER; LOPES, 2004, 2009). Em outras palavras, pode-se postular que,

na formação do dialeto, o uso do artigo definido caiu sensivelmente, em decorrência do processo de transmissão linguística irregular desencadeado na situação de contato entre línguas; porém, a partir de meados do século XX, o uso do artigo tem sido incrementado, em função de influências externas à comunidade. A análise dos resultados quantitativos das demais variáveis sociais pode confirmar, ou não, essa interpretação.

#### **4.3.2 Estada fora da comunidade**

A proposição dessa variável social é uma decorrência da visão de que estaria em curso um processo de nivelamento linguístico das variedades populares do interior do país a partir da influência e do contato com a chamada norma urbana culta, aferindo, assim, como o contato com outros dialetos pode interferir no uso do artigo definido na comunidade estudada.

Geralmente, os indivíduos pertencentes às comunidades populares tendem a procurar melhores condições de vida em cidades mais urbanizadas. A limitada medida de planejamento do poder público nessas comunidades suscita a carência de emprego e uma ineficiente assistência médica – setores essenciais na vida de qualquer população seja ela rural ou urbana.

Para nosso estudo, nos interessa saber o quanto o afastamento por pelo menos seis meses da comunidade influenciou na aplicação ou não da variante padrão. Nesse sentido, a razão para sair da comunidade também se relaciona às visitas familiares ou mesmo a passeios. A hipótese é a de que os indivíduos que, por qualquer uma dessas razões, viveram pelo menos seis meses fora da comunidade sofreram mais influência externa, incrementando o uso do artigo.

A tabela 10 assinala uma diferença entre os falantes que viveram algum tempo fora da comunidade e os que nunca se ausentaram:

TABELA 10: A presença do artigo definido em SNs, segundo a variável *estada fora da comunidade*.

<b>Estada fora da comunidade</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Sim</i>	179/268	<b>67</b>	<b>.56</b>
<i>Não</i>	188/352	53	.45
<b>TOTAL</b>	<b>367/620</b>	<b>59</b>	<b>---</b>

*Log likelihood = -299.972*      *Significância = .037*      *[Rodada 2]*

Nota-se, pela observação dos resultados da tabela 10, que a hipótese inicial é confirmada – os indivíduos que saíram lideram o processo de implementação da variante padrão com .56 de peso relativo. Os falantes que nunca se ausentaram, com .45, comprovam a menor disposição às influências linguísticas de indivíduos pertencentes aos centros urbanos.

#### 4.3.3 Escolaridade

O estudo da variável escolaridade tem revelado, em numerosos trabalhos sociolinguísticos, o empenho da escola em substituir as formas linguísticas que estão em discordância com o padrão gramatical. Alguns fenômenos linguísticos do português brasileiro – variantes desprestigiadas – estão mais expostos a essa coerção gramatical realizadas pelas instituições escolares. O que se espera é, então, que indivíduos mais escolarizados revelem usos mais próximos da norma gramatical de prestígio.

Em nossa amostra, nenhum informante tem alto grau de escolaridade, já que pretendemos observar as marcas mais acentuadas deixadas pelo contato entre línguas no passado da comunidade de Helvécia. Assim, consideramos dois tipos de informante quanto à escolarização: (i) os *analfabetos* – falantes sem nenhum contato com a escola e (ii) os *semi-analfabetos* – falantes que chegaram a cursar, no máximo, as séries iniciais. Em geral, os informantes semi-analfabetos sabem assinar apenas o próprio nome, e, em grande parte, o contato com a escola ocorreu na infância.

No que se refere ao uso do artigo definido em SNs, a variável escolaridade não foi selecionada na rodada inicial. Isso ocorreu porque esta variável não é estratificada, ou seja, não está distribuída equitativamente pelas células sociais. Assim, acreditamos que a variável escolarização está imbricada com a variável idade, pois, em geral, os

falantes da *Faixa 1* são um pouco mais escolarizados, enquanto os falantes das demais faixas são analfabetos, refletindo portanto o perfil da comunidade: os jovens tendem a frequentar a escola, enquanto os mais velhos não tiveram acesso ao ensino regular formal.

O VARBRUL, portanto, interpretou o fato como uma sobreposição de fatores, uma vez que a variável faixa etária já havia sido interpretada como relevante. Numa terceira rodada, retiramos a variável faixa etária. Sem essa sobreposição, a variável escolaridade foi selecionada com nível de significância 0,41. A análise dos dados demonstrou, assim, que há uma atuação direta do fator escolaridade na escolha da variante de prestígio, como evidencia a tabela 11:

TABELA 11: A presença do artigo definido em SNs, segundo a variável *escolaridade*.

<b>Escolaridade</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Semi-analfabeto</i>	162/222	<b>73</b>	<b>.64</b>
<i>Analfabeto</i>	205/398	52	.42
<b>TOTAL</b>	367/620	59	---

*Log likelihood = -300.0099*      *Significância = .041*      *[Rodada 3]*

Na amostra estudada, o peso relativo de .64 e 73% das ocorrências para o uso do artigo definido pelos informantes semi-analfabetos comprova a sensibilidade do fenômeno em relação à valoração social e, conseqüentemente, da sua aquisição no contexto escolar.

#### 4. 4 CRUZAMENTO SIGNIFICATIVO

Foram realizados vários cruzamentos, principalmente com a variável linguística *referencialidade*. Dos grupos estabelecidos como possíveis condicionantes da regra, o VARBRUL selecionou apenas um: *referencialidade x função sintática*.

##### 4.4.1 Referencialidade x função sintática

As variáveis *referencialidade* e *função sintática* foram consideradas muito significativas para a aplicação da variante padrão. A fim de investigar melhor a relação entre estas duas variáveis foi realizado o cruzamento destes dois fatores, conforme apresenta a tabela 12 e o gráfico 6:

TABELA 12: A variação do artigo definido em SNs, segundo o cruzamento das variáveis *referencialidade* x *função sintática*.

Nível de Ref.	Existencial		Def. prag.		Genérico		Def. anaf.	
	N/total	%	N/total	%	N/total	%	N/total	%
<b>F. Sintática</b>								
<i>Objeto direto</i>	32/73	<b>44</b>	89/128	<b>70</b>	34/130	<b>26</b>	16/27	<b>63</b>
<i>Sujeito</i>	8/10	<b>80</b>	94/120	78	64/97	<b>66</b>	29/35	83
<b>TOTAL</b>	40/83	48	183/248	74	98/227	43	46/62	74

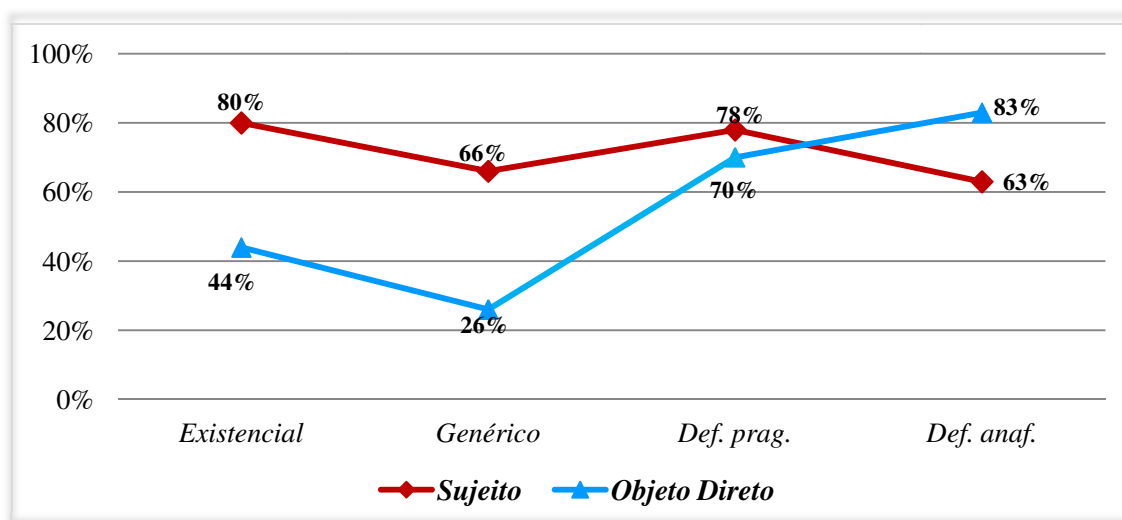


GRÁFICO 6: Cruzamento das variáveis *referencialidade* x *função sintática*.

A representação gráfica evidencia bem que a função sintática só interfere nos SNs não referenciais. Quando o SN é definido, tanto anaforicamente quanto pragmaticamente, a função sintática não interfere muito, ficando o uso do artigo com uma frequência entre 70% e 83% do total de ocorrências.

Os resultados são mais expressivos quando o SN é genérico: o uso do artigo sobe quase três vezes na posição de sujeito, 66% das ocorrências, em relação ao objeto direto, 26%. E o valor quase dobra quando o SN tem referente existencial, apresentando 80% dos casos na posição de sujeito e 44% dos casos na posição de objeto direto.



#### 4.5 OUTRAS VARIÁVEIS ANALISADAS

A única variável social que não foi selecionada pelo VARBRUL foi o *gênero do informante*. As variáveis linguísticas *tipo de núcleo do SN* (concreto/abstrato), *modificadores pós nucleares* (quantificador, sintagma preposicionado, nenhum modificador, adjetivo, partícula locativa e oração relativa), *animacidade do núcleo do SN* ([+humano] x [-humano]) e *estatuto informacional* (dado, novo e contrastivo) também não foram selecionadas como estatisticamente relevantes. Nas seções seguintes, vamos apresentar os resultados obtidos junto a essas variáveis. No geral, o que se observa é uma pequena diferença nos valores de cada fator dentro do grupo. Dessa forma, pelo que nos dizem os dados quantitativos, essas variáveis não teriam interferência sobre o uso do artigo; o que nem sempre está de acordo com os resultados de estudos anteriores.

##### 4.5.1 Gênero do núcleo do SN

Estudos sociolinguísticos na comunidade de Helvécia têm registrado o sexo masculino como aquele que lidera os processos de mudança linguística em direção ao padrão urbano, principalmente nas variáveis muito marcadas (LUCCHESI, 2009). Isso se deve ao fato de os homens terem mais contato com o mundo exterior. Já as mulheres, por sua vez, exibem atitudes linguísticas mais conservadoras, pois possuem uma vida social mais recatada, exercendo suas atividades no ambiente doméstico.

A variável *uso do artigo definido no SN* é pouco marcada e não foi observada antes na comunidade de Helvécia. Podemos ver os resultados na tabela 13 abaixo:

TABELA 13: A presença do artigo definido em SNs, segundo a variável *gênero do núcleo do SN*.

<b>Gênero do núcleo do SN</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Masculino</i>	223/366	<b>61</b>	
<i>Feminino</i>	144/254	57	
<b>TOTAL</b>	367/620	59	---

Como podemos observar, os resultados são inconclusivos, estão muito próximos do percentual total e não permitem considerações a respeito da influência da variável gênero sobre o fenômeno em foco.

#### 4.5.2 Tipo de núcleo do SN

Observamos a possibilidade de haver interferência na escolha do informante, se o núcleo do SN fosse *concreto* – indicador de seres materiais ou espirituais, reais ou fictícios; e núcleo do SN *abstrato* – indicador de seres que dependem de outros para se manifestar ou existir (atributos, estados, qualidades e ações, derivados de um conceito original).

##### a) *Concreto*

(55) *O motô é a óleo, a gente compra o óleo.* (HV- 01)

##### b) *Abstrato*

(56) *É, 'báxa o preço, a gente leva tudo: mandacumã, batata...* (HV- 01)

Os resultados obtidos estão na tabela 14 que se segue:

TABELA 14: A presença do artigo definido em SNs, segundo a variável *animacidade do núcleo do SN*.

<b>Tipo de núcleo</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Concreto</i>	300/506	59	
<i>Abstrato</i>	67/114	59	
<b>TOTAL</b>	367/620	59	---

Como podemos observar, a variável não exerce influência para o condicionamento do fenômeno.

### 4.5.3 Presença de modificadores pós-nucleares

Esta variável considera a influência de posteriores modificadores no SN. Os exemplos que se seguem apresentam os fatores analisados:

a) *Quantificador*

(57) *Comprava os remédio tudo.* (HV- 07)

b) *Sintagma preposicionado*

(58) *...é difíci o dono da tenda cobrá, né?* (HV- 01)

c) *Nenhum modificador*

(59) *Aí, ele foi, aí ele pegô **a bacia**.* (HV- 03)

d) *Adjetivo*

(60) *Aí tirô o gesso grande, aí botô aqui, ó!* (HV- 03)

e) *Partícula locativa*

(61) *Os pessoal lá fora afrito!* (HV- 03)

f) *Oração relativa*

(62) *A casa que eu morava é lá.* (HV- 12)

Os resultados da variável *presença de modificadores pós-nucleares* estão dispostos na tabela 15:

TABELA 15: A presença do artigo definido em SNs, segundo a variável *modificadores pós-nucleares*.

<b>Modificadores pós-nucleares</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Quantificador</i>	7/9	<b>78</b>	
<i>Sprep</i>	40/84	48	
<i>Nenhum</i>	263/434	61	
<i>Adjetivo</i>	14/26	54	
<i>Partícula locativa</i>	13/26	50	
<i>Oração relativa</i>	20/31	65	
<b>TOTAL</b>	357/610	59	---

Há certa convergência no emprego do artigo definido quando os modificadores manifestam o traço definidor. Temos neste caso os quantificadores (78%) e as orações relativas (65%). Segundo Baxter e Lopes (2009, p. 326), o comportamento do fator *partícula locativa* (advérbios locativos) tem origem de sua atuação como demonstrativo.

#### 4.5.4 Animacidade do núcleo do SN

A variável animacidade do núcleo do SN, como já mencionado, não foi selecionada pelo programa VARBRUL, manifestando a inexistência de uma influência direta desse fator sobre a variação do fenômeno em questão.

Contudo, outros trabalhos (ALEXANDRE e HAGEMEIGER, 2007; BAXTER e LOPES, 2004 e 2009; BAPTISTA, 2007) demonstraram que essa variável semântica influencia de certa forma o uso do artigo definido quando o núcleo do SN é [+humano]. Abaixo, exemplos dos traços semânticos considerados:

a) [+ *humano*]

(51) *Os médico* num deu a duração pra tirá. (HV- 01)

(52) Mandô procurá *Ø empregada* pra cá e aí ela foi. (HV- 09)

b) [- humano]

(53) ...aí correu pa chamá *Ø carro*. (HV- 03)

(54) *O sangue* já tinha esfriado. (HV- 03)

Os resultados percentuais da distribuição para os fatores desse grupo encontram-se na tabela 16 que segue:

TABELA 16: A presença do artigo definido em SNs, segundo a variável *animacidade do núcleo do SN*.

Número do SN	Nº. de ocorrências/total	%	Peso relativo
[+humano]	111/173	<b>64</b>	
[-humano]	256/447	57	
<b>TOTAL</b>	367/620	59	---

Os resultados da tabela acima indicam que o traço [+ humano] favorece o emprego do artigo definido com 64% das realizações. Embora a percentagem não seja significativa, apresentando um resultado muito próximo do percentual total, resolvemos cruzá-la com a variável função sintática, no intuito de observar a dependência entre o SN sujeito e o traço [+humano] e entre SN objeto direto e o traço [-humano].

#### 4.5.5 Estatuto informacional

Esta variável refere-se ao *status* informacional do SN no nível do discurso. De modo que um SN *dado* veicula uma informação pressuposta, ou seja, que já foi mencionada no discurso, direta ou indiretamente, e faz parte do conhecimento compartilhado entre falante e ouvinte. O *status* informacional novo, em contrapartida, é uma informação recentemente inserida no discurso. O fator *status* informacional dado parece ser mais favorável para a ocorrência do artigo definido, embora não haja

obrigatoriedade do emprego deste na referência de um termo já conhecido. O *status* informacional contrastivo também foi analisado, neste caso, a negativa faz parte do discurso do informante. Em todos os casos, a fala do documentador foi levada em consideração.

a) **Dado**

Documentador: *E o trabalho lá pagava direito?*

(63) *Com abono de família, tirava mil e quiento. A turma falô: é, João sabido, você vai quebá a firma. Aí eu respondê: Se Ø firma num 'guentá comigo, eu... num`guenta com vocês! (HV- 12)*

b) **Novo**

Documentador: *E as mulheres plantavam?*

(64) *Samiava tudo aí; mulé que chegava... queria samiar. (HV- 07)*

c) **Contrastivo**

Documentador: *A menina é filha de sua cunhada?*

(65) *Não, a menina né filha dela não! (HV- 08)*

Almejava-se com esta variável que a informação conhecida pelos interlocutores favoreceria o emprego do artigo definido, tendo em vista a relação semântico-discursiva. Os resultados encontram-se na tabela 17 que segue:

TABELA 17: A presença do artigo definido em SNs, segundo a variável *estatuto informacional*.

<b>Estatuto informacional</b>	<b>Nº. de ocorrências/total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Dado</i>	197/322	<b>61</b>	
<i>Novo</i>	169/295	57	
<i>Contrastivo</i>	1/3	33	
<b>TOTAL</b>	<b>367/620</b>	59	---

Como os resultados estão muito próximos do percentual total, não há possibilidade de concluir de forma assente o comportamento desta variável.

## CONSIDERAÇÕES FINAIS

Neste trabalho, analisamos o uso variável do artigo definido em sintagmas nominais na fala de indivíduos pertencentes à comunidade de Helvécia localizada no extremo sul do estado da Bahia. As análises sobre as amostras da fala vernácula dos habitantes de Helvécia, muitas das quais reunidos no livro *O Português Afro-Brasileiro* (LUCCHESI; BAXTER; RIBEIRO, 2009), têm chamado a atenção para características morfossintáticas que apontam para um processo pretérito de tipo criouliizante. Nesse sentido, a variação no uso do artigo definido pode trazer novas evidências de um processo anterior de reestruturação gramatical, pelo qual teria passado o dialeto de Helvécia, fato já observado por Baxter e Lopes (2004, e 2009) e, mais recentemente, por Ribeiro (2010). Assim, a análise variacionista do uso do artigo definido desenvolvida nesta dissertação buscou trazer novas evidências que comprovem a hipótese da transmissão linguística irregular (BAXTER; LUCCHESI, 1997; LUCCHESI, 2003; LUCCHESI; BAXTER, 2009), por meio do confronto dos resultados obtidos com análises do sistema de artigos dos crioulos de Cabo Verde e São Tomé (Lucchesi, 2003 e Baptista, 2007). Constatou-se, então, que as semelhanças estão concentradas nos contextos em que são empregados sintagmas nominais sem o artigo (nomes nus).

Por outro lado, no plano sociolinguístico, a concatenação dos resultados das variáveis sociais aponta para uma mudança em progresso na comunidade em direção ao padrão urbano. Tal mudança estaria revertendo um processo anterior derivado do contato entre línguas, em que o uso do artigo definido teria sido muito restringido.

Já no plano do encaixamento do processo variável na estrutura linguística, as variáveis estabelecidas como favorecedoras no uso do artigo definido estão: referencialidade, função sintática, número do SN, determinante na menção anterior, contabilidade do núcleo do SN. O cruzamento entre as variáveis referencialidade e função sintática também foi selecionado como significativo pelo programa estatístico VARBRUL.

O resultado da apreciação da variável referencialidade demonstra que o uso do artigo definido é favorecido quando a referencialidade do SN é definida anaforicamente ou pragmaticamente, apresentando um peso relativo de .73 e .63, respectivamente. Ratifica, assim, resultados encontrados no crioulo cabo-verdiano e no santomense, em

que há maior possibilidade de ausência do artigo em SNs de referencialidade existencial ou genérica, sobretudo nesse último contexto.

A função sintática foi a segunda variável selecionada como significativa pelo VARBRUL, evidenciando sua interferência no condicionamento das variantes analisadas. A posição de sujeito favorece o uso do artigo definido em 74% dos casos, confirmando nossas expectativas de que os argumentos externos intervêm no emprego do artigo definido. Para tanto, cruzamos esta variável linguística com a variável semântica referencialidade, a fim de observar melhor a interferência de ambas na implementação do uso do artigo definido. O resultado do cruzamento revela que o nível de referencialidade genérico ascende significativamente na posição de sujeito em detrimento da função de objeto direto, deixando claro que a variável função sintática só interfere em SNs não referenciais.

Os resultados revelaram ainda que a variável número do SN condiciona a escolha das variantes, sendo o plural bastante favorável à presença do artigo, com peso relativo de .90. Esse resultado pode ser explicado pela estratégia de marcação de plural dos nomes nas comunidades mais isoladas: há uma tendência em assinalar morfologicamente os plurais na primeira posição do sintagma, sendo o determinante decisivo na introdução dessa marca de plural.

A análise da variável determinante na menção anterior, com base no princípio do paralelismo discursivo (SCHERRE; NARO, 1993), revelou que a presença do artigo definido no contexto anterior favorece sua reiteração na ocorrência subsequente, com 79% das ocorrências. Foi confirmada, igualmente, a nossa expectativa de que, quando o SN prévio é iniciado pelo artigo indefinido, maior seria a possibilidade de uso do artigo definido no SN posterior, com .63 de peso relativo.

A última variável linguística selecionada como significativa pelo programa computacional foi a contabilidade do núcleo do SN. Os resultados revelam que o núcleo contável é favorecedor do uso do artigo definido, com .52 de peso relativo.

Das variáveis sociais estabelecidas nesta pesquisa, destaca-se a variável faixa etária, apontando fortemente para um cenário de incremento do uso do artigo definido na comunidade. A escala crescente em direção ao emprego do artigo definido é demonstrada pela maior incidência do uso do artigo definido na fala dos mais jovens, com .69 de peso relativo. Os falantes mais idosos mantêm os vestígios de um processo anterior de reestruturação morfossintática que teria inibido o uso dessa partícula



gramatical, constatando-se um baixo nível de emprego na sua fala, expresso no peso relativo de apenas .28.

A análise da variável estada fora da comunidade revelou que os indivíduos que saíram da comunidade por pelo menos seis meses utilizam mais o artigo definido com .56 de peso relativo. O resultado confirma também que os falantes que nunca se ausentaram são menos vulneráveis às influências linguísticas dos centros urbanos.

Após reparar o problema de sobreposição de fatores, a variável escolaridade foi selecionada pelo VARBRUL como relevante estatisticamente. O resultado revelou a aquisição do artigo definido no contexto escolar, demonstrando que falantes mais escolarizados usam mais este determinante com .64 de peso relativo.

Portanto, o conjunto das variáveis sociais constituiu um cenário sociolinguístico, na comunidade de fala de Helvécia, de mudança em direção ao padrão urbano que reverte um processo anterior de reestruturação gramatical induzido pelo contato linguístico que teria inibido fortemente o uso do artigo, particularmente em SNs de referência genérica, o que constitui um paralelo notável com uma das características consideradas como mais típicas das línguas crioulas, desde o estudo clássico de Dereck Bikerton (1981): a ausência de artigo como marca dos SNs de referência genérica.

Porém, apesar dos resultados alcançados nesta análise, muito ainda há a ser feito, para entender como o contato entre línguas afetou o uso do artigo definido na formação das variedades populares do português brasileiro; a começar pela ampliação da base de dados, analisando as outras comunidades do português popular rural que fazem parte do acervo do Projeto Vertentes do Português Popular do Estado da Bahia.

## REFERÊNCIAS

ALENCASTRO, Luiz Felipe. **O trato dos viventes: formação do Brasil no Atlântico Sul**. São Paulo: Companhia das Letras, 2000.

ALMEIDA MENDES, Andréia. Ausência/presença de artigo definido diante de antropônimos nas localidades de Matipó e Abre Campo. In: **Anais da Semana de Atividades Acadêmicas Especiais – Ensino, Pesquisa e Extensão**. Belo Horizonte: Faculdade de Letras/UFMG, 2006.

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.

ALVES, Ana Paula Mendes. **Um estudo sociolingüístico da variação sintática ausência/presença de artigo definido diante de antropônimos na fala dos jovens moradores de Barra Longa/ MG que residem em Belo Horizonte**. Dissertação (Mestrado) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

AMARAL, Eduardo Tadeu Roque. **A ausência/presença de artigo definido diante de antropônimos em três localidades de Minas Gerais: Campanha, Minas Novas e Paracatu**. Dissertação (Mestrado) - Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2003.

AMARAL, Eduardo Tadeu Roque. **A importância do fator intimidade na variação ausência/presença de artigo definido diante de antropônimos**. In: Veredas on Line – ATEMÁTICA – 1/2007. p. 116-127.

ANTONIL, André João. **Cultura e opulência do Brasil**. Belo Horizonte: Itatiaia; São Paulo: Editora da Universidade de São Paulo, 1982.

BAPTISTA, Marlyse. On the syntax and semantics of DP in Cape Verdean Creole. In: M. Baptista & J. Guéron. (eds.). **Noun phrases in creole languages**. A multi-faceted approach. John Benjamins, 2007, p. 61-105.

BAXTER, Alan; LOPES, Norma. O artigo definido. In D. Lucchesi, A. Baxter & I. Ribeiro. (orgs.). **O português afro-brasileiro**. Salvador: EDUFBA, 2009.

BAXTER, Alan. Língua Pidgin e crioulas. In: FARIA, I.H., PEDRO, E.R., DUARTE, I., GOUVEIA, C. A. **Introdução à Lingüística Geral e Portuguesa**. Lisboa: Caminho, 1996.

BAXTER, A. N.; LUCCHESI, D. A relevância dos processos de pidginização e Crioulização na formação da língua portuguesa no Brasil. **Estudos lingüísticos e Literários**. Salvador: EDUFBA, 1997. n.19, p. 65-84.

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

BERGER, Peter; LUCKMANN, Thomas. **A construção social da realidade**: tratado de sociologia do conhecimento. Petrópolis: Vozes, 1985.

Bickerton, Derek. Relexification. **Journal of pidgins and creole languages** 3,2.277-282, 1988.

Bickerton, Derek. **Roots of Language**. Karoma Publishers, Inc. Ann Arbor, 1981.

BONVINI, Emile. Línguas africanas e português falado no Brasil. In: FIORIN, José Luiz; PETTER, Margarida (Orgs.). **África no Brasil**: a formação da língua portuguesa. São Paulo: Contexto, 2008.

BORTONI-RICARDO, Stella Maris. Um modelo para a análise sociolingüística do português do Brasil. In: BAGNO, Marcos. (Org.). **Lingüística da Norma**. São Paulo: Edições Loyola, 2002. p. 333-350.

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

CALLOU, Dinah Maria Insensee; LOPES, Célia Regina dos Santos. Contribuições da sociolingüística para o ensino e a pesquisa: a questão da variação e da mudança lingüística. **Revista do GELNE**, João Pessoa, v. 5, n. 1-2, p. 63-74, 2003.

CALVET, Louis- Jean. **Sociolingüística**: uma introdução crítica. Tradução de Marcos Marcionilo. São Paulo: Parábola, 2002.

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

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

CHAMBERS, Jack; TRUDGILL, Peter. **Dialectology**. New York: Cambridge University Press, 1980.

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

COSTA, Iraneide. O uso do artigo definido diante de nome próprio de pessoa e possessivo do século XIII ao século XVI. In: \_ MATTOS e SILVA, R. V.; MACHADO FILHO, A.V. L. (orgs). **O português quinhentista**: estudos lingüísticos. Salvador: EDUFBA/UEFS, 2002. p.285-306.

COSTA, Iraneide. **Uso do artigo definido diante de possessivo e de nome próprio de pessoa**: período arcaico. Dissertação (Mestrado). Instituto de Letras, Universidade Federal da Bahia, Salvador, 1998.

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

DU BOIS, J. *et al.* **Dicionário de Lingüística**. 10. ed. São paulo: Cultrix, 1998.

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

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: 18 de setembro de 2010.

GIVÓN, Talmy. Definiteness and referentiality. In J. Greenberg, C. Ferguson; E. Moravcsik (Eds.) *Universals of human language*, vol. 4. Stanford: Stanford University Press, 1978, p. 291-330.

GNERRE, Maurizio. **Linguagem, escrita e poder**. 4. ed. São Paulo: Martins Fontes, 1998.

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

HOUAISS, Antônio. **O português do Brasil**. Rio de Janeiro: Unibrade Centro de Cultura, 1985.

HUBER, J. **Gramática do português antigo**. Lisboa: Gulbenkian, 1986.  
INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Rio de Janeiro, 2006. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 08 de setembro 2010.

JANSON, T. **Articles and plural formation in creoles: change and universals**. *Lingua*, 1984, p. 291-329.

KATO, Mary Aizawa. **A semântica gerativa e o artigo definido**. São Paulo: Ática, 1974.

LACA, Brenda. Presencia y ausencia de determinante. In: I. Bosque & V. Demonte. (orgs.). **Gramática descriptiva de la lengua española**, v. 1, Sintaxis básica de las clases de palabras. Madrid: Editorial Espasa, 1999, p. 891-928.

LABOV, William. **Philadelphia**: University of Pennsylvania Press, 1972.

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

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.

LABOV, William. **Principles of Linguistic Change**: Internal Factors. Cambridge: Blackwell Publishers, 1994.

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.

LOBO, Tânia Conceição Freire; MACHADO FILHO, Américo Venâncio Lopes; MATTOS E SILVA, Rosa Virgínia (2006). Índícios de língua geral no sul da Bahia na segunda metade do século XVIII. In: LOBO, Tânia Conceição Freire; RIBEIRO, Ilza; CARNEIRO, Zenaide & ALMEIDA, Norma. (Orgs.). **Para a história do português brasileiro**. Volume VI: novos dados, novas análises, t. II. Salvador: EDUFBA.

LUCCHESI, Dante. **A Variação na Concordância de Gênero em uma Comunidade de Fala Afro-brasileira**: Novos Elementos sobre a Formação do Português Popular do Brasil. Rio de Janeiro: UFRJ. Tese de doutorado.ms, 2000.

LUCCHESI, Dante. **O português se teria crioualizado no Brasil?** – refletindo sobre uma velha questão. *APB*. v.1. Frankfurt, 2000. p.25-41

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.

BAXTER, A; LOPES, Norma. O artigo definido. In.: BAXTER, Alan; LUCCHESI, Dante; RIBEIRO, Ilza (Org.). **O português afro-brasileiro**. Salvador: EDUFBA, 2009. p. 319-330.

LUCCHESI, Dante. The article systems of Cape Verde and São Tomé Creole Portuguese: general principles and specific factors. **Journal of Pidgin and Creole Language**, 1993, p. 81-108.

LYONS, John. **Semântica**. Vol. 1. Lisboa: Presença/ Martins Fontes, 1977.

LYONS, Christopher. **Definiteness**. Cambridge: Cambridge University Press, 1999.

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

MATTOS E SILVA, Rosa Virgínia. **Caminhos da linguística histórica: ouvir o inaudível**. São Paulo: Parábola, 2008.

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.

MATTOS E SILVA, Rosa Virgínia. **O português são dois: novas perspectivas, velhos problemas**. São Paulo: Parábola Editorial, 2004b.

MATTOS E SILVA, Rosa Virgínia. Notas sobre avaliações lingüísticas nos gramáticos Fernão de Oliveira e João de Barros. In: SILVA, Rosa Virgínia Mattos e; MACHADO FILHO, Américo Venâncio. (Org.). **O português Quinhentista: estudos lingüísticos**. Salvador: Edufba; Feira de Santana: UEFS, 2002. p. 43-60.

MATTOS E SILVA, Rosa Virgínia. Uma compreensão histórica do Português Brasileiro: velhos problemas repensados. In: CARDOSO, Alice Marcelino; MOTA, Jacyra Andrade Mota; SILVA, Rosa Virgínia Mattos e (Org.). **Quinhentos anos de história lingüística do Brasil**. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. p. 219-254

MEILLET, Antoine. **Linguistique historique et linguistique générale**. Paris Champion, 1921.

MEINTEL, D. **The creole of the island of Brava**. In: VALKHOFF (Org.), 1975, p. 205-256.

MENDES, Soélis Teixeira do Prado. **A ausência/presença do artigo definido diante de nomes próprios no português mineiro da comunidade de Barra Longa: um caso de retenção?** Dissertação (Mestrado) - Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2000.

MIRA MATEUS, Maria Helena et all. **Gramática da Língua portuguesa**. Elementos para a descrição da estrutura, funcionamento e uso do português actual. Coimbra: Almedina, 1983

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

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

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 sociolingüística: 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 sociolingüística: o tratamento da variação**. [S.I.:s.n], 2004. p. 43-50.

NEGRI, Lígia. **A natureza dêitica do artigo definido**. 1986, 135 f. Dissertação (Mestrado) – Universidade Estadual de Campinas, São Paulo, 1986.

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.

NOLL, V. **O português brasileiro: formação e contrastes**. São Paulo: Globo, 2008.

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.

PAIVA, Maria da Conceição. Transcrição de dados lingüísticos. In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). **Introdução à Sociolingüística: o tratamento da variação**. 3. ed. São Paulo: Contexto, 2003, p. 135-146.

PAIVA, Maria da Conceição. A variável gênero/sexo. In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). **Introdução à Sociolinguística: o tratamento da variação**. 3. ed. São Paulo: Contexto, 2003, p. 33-42.

PAIVA, Maria da Conceição; DUARTE, Maria Eugênia Lamoglia. Mudança linguística: observações no tempo real. In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). **Introdução à Sociolinguística: o tratamento da variação**. 3. ed. São Paulo: Contexto, 2003, p.179-190.

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

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.

PROGRAMA Brasil Quilombola. Disponível em : <<http://bvms.saude.gov.br/bvs/publicacoes/brasilquilombola2004.pdf.htm>>. Acesso em: 10 out. 2010.

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

REIS, João José; GOMES, Flávio dos Santos (Orgs.). **Liberdade por um fio**. História dos quilombos no Brasil. 4 ed. São Paulo: Companhia da Letras, 2008.

RIBEIRO, Darcy. **O povo brasileiro: a formação e o sentido do Brasil**. São Paulo: Companhia das Letras, 2006.

RIBEIRO, Ilza. O sistema de definitude e de referencialidade de uma falante afro-brasileira idosa. In: ABECS – Associação Brasileira de Estudos Crioulos e Similares. Salvador, 2010 **Anais...** Salvador: UFBA, 2010.

ROBERTS, Ian; KATO, Mary A. (Orgs.). **Português Brasileiro: uma viagem diacrônica**. Campinas: Unicamp, 1993. p. 163-184.

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.



RODRIGUES, Aryon D. As outras línguas da colonização do 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.143-162.

RODRIGUES, Bruno de Andrade. **Estudo descritivo dos usos do clítico lhe na variedade formal do português**. 2007. Dissertação (Mestrado em Letras) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007.

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

SALLES, Ricardo e SOARES, Mariza. **Episódios de história afro-brasileira**. Rio de Janeiro: DPA/FASE, 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, Marta; NARO, Anthony. Duas dimensões do paralelismo formal na concordância verbal no português popular do Brasil. **DELTA**, São Paulo, v.9, n.1, p.1-14, 1993.

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

SILVA-CORVALÁN, C. **Sociolingüística: teoria y análisis**. Madrid: Alhambra, 1988.

SILVA, Giselle Machline de Oliveira. Estertores do possessivo seu na terceira pessoa. In: SILVA, Giselle Machline de Oliveira; SCHERRE, Maria Marta. P. (Orgs.). **Padrões sociolingüísticos: análises de fenômenos variáveis do português falado no Rio de Janeiro: Tempo Presente**, 1996, p. 171-181.

SILVA, Giselle Machline de Oliveira; SCHERRE, Maria Marta Pereira. **Padrões sociolingüísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro**. 2 ed. Rio de Janeiro: Tempo Brasileiro/Departamento de Lingüística e Filologia, UFRJ, 1996.

SILVA, Giselle Machline de Oliveira; PAIVA, Maria da Conceição de. Visão de conjunto das variáveis sociais. In: SILVA, Giselle Machline de Oliveira; SCHERRE, Maria Marta Pereira (Orgs.). **Padrões sociolingüísticos: análises de fenômenos variáveis do português falado no Rio de Janeiro**. Rio de Janeiro: Tempo Brasileiro, 1996, p. 336-78.

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.

SILVA NETO, S. da. **Introdução Estudo da Língua Portuguesa no Brasil**. 2ª ed. Rio de Janeiro: INL, 1963.

SILVA NETO, S. da. **História da Língua Portuguesa no Brasil**. 5ª ed. Rio de Janeiro: Presença, 1988.

TARALLO, Fernando. **A pesquisa sociolingüística**. 7. ed. São Paulo: Ática, 2005.

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.

TARALLO, F.; ALKIMIM, T. **Falares crioulos**: línguas em contato. São Paulo: Ática, 1987.

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

VOTRE, Sebastião José. Relevância da variável escolaridade In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza (Org.). **Introdução à Sociolingüística**: 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 lingüística**. Tradução: Marcos Bagno. São Paulo: Parábola, 2006 [1968].