



UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA E CULTURA
Rua Barão de Geremoabo, nº 147 – CEP 40170-290 – Campus Universitário Ondina – Salvador-BA
Tel./Fax: (71) 3283-6256 – Site: <http://www.ppglinc.letas.ufba.br/> – E-mail: pgletba@ufba.br

DÉBORA CARVALHO TRINDADE GOMES

**O USO VARIÁVEL DO *MAIS* NO PORTUGUÊS AFRO-BRASILEIRO:
COORDENAÇÃO E SUBORDINAÇÃO**

Salvador
2014

DÉBORA CARVALHO TRINDADE GOMES

**O USO VARIÁVEL DO *MAIS* NO PORTUGUÊS AFRO-BRASILEIRO:
COORDENAÇÃO E SUBORDINAÇÃO**

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

Orientadora: Profa. Dra. Maria Cristina Vieira de Figueiredo Silva

Co-Orientador: Pr. Dr. Alan Norman Baxter

Salvador
2014

Sistema de Bibliotecas da UFBA

Gomes, Débora Carvalho Trindade.

O uso variável do mais no português afro-brasileiro: coordenação e subordinação / Débora Carvalho Trindade Gomes. - 2014.
150 f.: il.

Inclui anexos.

Orientadora: Prof^a. Dr^a. Maria Cristina Vieira de Figueiredo Silva.

Co-Orientador: Prof. Dr. Alan Norman Baxter.

Dissertação (mestrado) - Universidade Federal da Bahia, Instituto de Letras, Salvador, 2014.

1. Língua portuguesa - Sintaxe. 2. Contato linguístico. 3. Sociolinguística. 4. Língua portuguesa - Português falado - Bahia. 5. Quilombos - Bahia. I. Silva, Maria Cristina Vieira de Figueiredo. II. Baxter, Alan Norman. III. Universidade Federal da Bahia. Instituto de Letras. IV. Título.

CDD - 469.5

CDU - 811.134.3'367

DÉBORA CARVALHO TRINDADE GOMES

**O USO VARIÁVEL DO MAIS NO PORTUGUÊS AFRO-BRASILEIRO:
COORDENAÇÃO E SUBORDINAÇÃO**

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

Aprovada em: 30 de julho de 2014.

BANCA EXAMINADORA

Professora Doutora Maria Cristina Vieira de Figueiredo Silva – UFBA (Orientadora)

Professor Doutor Alan Norman Baxter – UFBA (Co-Orientador)

Professor Doutor Rerisson Cavalcante de Araújo – UFBA

A Jesus Cristo, o Mestre dos mestres,
dedico esta dissertação.

AGRADECIMENTOS

A Deus, meu Pai amado, agradeço por sempre me guiar e me cercar de amor e de cuidados em todos os momentos, por estar sempre presente, por ser minha fonte de vida, de alegria e de força. Agradeço pelo amor incondicional e inexplicável, sem o qual não sei viver.

Ao meu esposo Francisco, por toda a dedicação e cumplicidade, por sempre acreditar em mim e reacender a esperança em meu coração. Agradeço por fazer meus dias mais leves com seu bom-humor de sempre, independente de qualquer circunstância. Como sempre digo, “o amor com você é mais bonito, você me faz feliz, meu bem!”.

Aos meus pais, Trindade e Carminha, e às minhas irmãs, Quésia e Priscila, a quem amo “desde sempre e para sempre”, pela confiança, preocupação e cuidado. Agradeço por torcerem pelas minhas conquistas e me apoiarem nas minhas decisões.

Às minhas amigas-irmãs Lene e Vanessa, por entenderem as vezes que eu não podia sair para dar uma volta, conversar, rir. Agradeço por entenderem meu afastamento necessário e por torcerem por mim, principalmente para eu terminar logo esta dissertação.

À minha orientadora, Profa. Dra. Cristina Figueiredo, agradeço imensamente por toda a generosidade, dedicação, paciência, encorajamento, oportunidades, e pelo exemplo de pessoa e profissional, a quem aprendi a admirar há muitos anos, desde o início de minha graduação. Agradeço por sempre abrir as portas de sua casa para me receber, orientar, e emprestar (muitos) livros que serviram como fonte para minha pesquisa. Agradeço pela disponibilidade de tempo na UFBA, em casa, por telefone, pelo Skype, pelo e-mail. Enfim, o tamanho da gratidão que sinto não se explica por meio de palavras. Existem pessoas que mudam a nossa vida quando as conhecemos, e minha orientadora é uma dessas pessoas para mim.

Ao meu co-orientador, Prof. Dr. Alan Baxter, que desde 1987 já chamara a atenção para o uso variável do *mais* em dialetos rurais do português brasileiro, agradeço por ter aceitado esta co-orientação e pelos muitos textos enviados, pelas ricas sugestões de leitura e de pesquisa e pelas excelentes discussões teóricas. Agradeço por sempre disponibilizar seu tempo e experiência para acrescentar à minha pesquisa maior riqueza de informações.

Ao Prof. Dr. Dante Lucchesi, coordenador do Projeto Vertentes, por sugerir a investigação do fenômeno que possibilitou a concretização deste trabalho e por acolher meu projeto de pesquisa junto ao Projeto Vertentes. Agradeço ainda pelas preciosas sugestões durante o andamento da pesquisa.

Ao Prof. Dr. Rerisson Cavalcante, pelas sugestões de importantes leituras e por fazer parte da avaliação final deste trabalho.

À Prof. Ms. Isis Barros, por me ajudar, gentil e carinhosamente, com a revisão final deste trabalho.

Agradeço ainda a todos os que contribuíram de alguma forma para que eu chegasse até aqui, perto ou longe, com orações, palavras de ânimo, e outras muitas demonstrações de carinho e amizade.

Que Deus abençoe a todos vocês!

RESUMO

No português popular do estado da Bahia, é muito comum o uso do item de vocabulário (IV) *mais* estabelecendo relação de subordinação, variando com o IV *com*, ou de coordenação de partes da sentença, variando com o IV *e*, além de seu uso culto como intensificador. Essas relações ampliam as características e a funcionalidade do IV *mais*, que pode exercer função de coordenador de DPs e preposição funcional/lexical introduzindo DPs comitativos. O presente estudo, realizado de acordo com os princípios teóricos e metodológicos da Sociolinguística Quantitativa, apresenta uma descrição do uso do IV *mais* como coordenador e subordinador em quatro comunidades rurais afrodescendentes localizadas no interior do estado da Bahia: *Helvécia*, *Cinzento*, *Sapé*, e *Rio de Contas*, cujas amostras de fala fazem parte do *Projeto Vertentes do Português Popular do Estado da Bahia* (www.vertentes.ufba.br). Tais comunidades, que guardam vestígios de contato linguístico, foram selecionadas considerando-se a hipótese de que os fenômenos variáveis *mais/com* e *mais/e* resultam do contato entre a língua portuguesa e línguas do substrato africano no Brasil. O tratamento sociolinguístico dado à investigação dos fatores condicionantes da variante *mais*, no português rural afro-brasileiro, demonstrou que este IV é favorecido na fala dos indivíduos mais velhos, demonstrando que essa é a variante mais antiga na comunidade, e dos que permaneceram na comunidade. O *mais* subordinador é favorecido ainda na fala das mulheres, considerada mais conservadora no âmbito dessas comunidades. Esse resultado parece apontar para a confirmação da hipótese proposta de que a ampliação funcional do IV *mais* se deve ao multilinguismo ocorrido no Brasil no período colonial. Os resultados revelam um perfil de mudança em progresso, em que a variante *mais* está sendo substituída, nessas comunidades, pelas formas urbanas cultas *e/com*, por conta de um processo de nivelamento linguístico (cf. Lucchesi, 2001, 2006). Os resultados referentes às duas variáveis dependentes *mais/com* and *mais/e* evidenciam ainda que a *referencialidade* e o *estatuto sintático do DP* são condicionantes linguísticos relevantes na escolha da variante *mais*. Assume-se, neste trabalho, que o IV *mais* foi a escolha dos falantes de português L2, do substrato africano, para relexificar uma entrada copiada de sua(s) língua(s) nativa(s), um IV com as funções de coordenador entre DPs e subordinador comitativo.

Palavras-chave: Português rural afro-brasileiro. Contato entre línguas. Relexificação. Coordenação. Subordinação.

ABSTRACT

In the popular Portuguese of Bahia state, the vocabulary item (VI) ‘more’ frequently occurs in functions beyond its standard spoken language use as an intensifier. Thus, ‘more’ occurs in variation with the VI ‘with’, marking a subordination relationship, and it also occurs in variation with ‘and’ signaling coordination of parts of the sentence. These relationships extend the features and functionality of VI ‘more’ whereby it may act as coordinator of DPs and also as a functional/lexical preposition introducing comitative DPs. This study, conducted in accordance with the theoretical and methodological principles of Quantitative Sociolinguistics, presents a description of the use of VI ‘more’ as coordinator and subordinator in four Afro-Brazilian rural communities located in the state of Bahia: *Helvecia*, *Cinzento*, *Sape* and *Rio de Contas*, whose speech samples are part of the *Projeto Vertentes do Português Popular do Estado da Bahia* (www.vertentes.ufba.br). These communities, which retain vestiges of language contact, were selected considering the hypothesis that the variables ‘more’/‘with’ and ‘more’/‘and’ result from contact between Portuguese and African substrate languages in Brazil. The sociolinguistic evaluation of the conditioning factors of the variant ‘more’, in the data considered, demonstrated that this VI is favored in the speech of older individuals, and is thus the oldest variant of the particular variable in the community, and those who have remained in the community. The subordinator ‘more’ is favored even by women, whose speech is considered to be the more conservative variety within these communities. This result suggests confirmation of the hypothesis proposed, namely that the functional expansion of the VI ‘more’ is due to multilingualism present in Brazil during the colonial period. The results reveal a profile of change in progress, in which the variant ‘more’ is being replaced in these communities by the standard urban forms ‘and’/‘with’, due to a process of linguistic leveling (cf. Lucchesi, 2001, 2006). The results for the two dependent variables ‘more’/‘with’ and ‘more’/‘and’ also show that the *referentiality* and *syntactic status of the DP* constitute relevant linguistic constraints on the choice of the variant ‘more’. The study assumes that the VI ‘more’ was the choice of African substrate speakers of Portuguese L2 to relexify an entry copied from their native language(s), an VI with the functions of coordinator of DPs and comitative subordinator.

Keywords: Rural Portuguese Afro-Brazilian. Contact between languages. Relexification. Coordination. Subordination.

LISTA DE FIGURAS

Mapa 1: Línguas com valor de identidade e diferenciação de formas no APiCS	39
Mapa 2: Línguas com valor de identidade e diferenciação de formas no WALs	42
Mapa 3: Línguas com valor de identidade de formas na África – dados do APiCS	52
Mapa 4: Línguas com valor de identidade de formas na África – dados do WALs	53
Gráfico 1: Distribuição das variantes <i>mais/e</i> segundo a <i>comunidade</i>	103
Gráfico 2: Distribuição das variantes <i>mais/e</i> segundo a <i>faixa etária</i>	105
Gráfico 3: Hierarquia de referencialidade proposta por Cyrino, Kato, Duarte (2000).	108
Gráfico 4: Distribuição das variantes <i>mais/com</i> segundo a <i>faixa etária</i>	125

LISTA DE QUADROS

Quadro 1: Etimologia de conjunções em línguas crioulas do Atlântico	33
Quadro 2: Comitativo e coordenadores em línguas europeias e nigero-congolesas	35
Quadro 3: Conector de sintagmas nominais, frases e orações em línguas crioulas.....	36
Quadro 4: <i>Corpus</i> base do português rural afro-brasileiro de Helvécia.....	85
Quadro 5: <i>Corpus</i> base do português rural afro-brasileiro de Rio de Contas.....	85
Quadro 6: <i>Corpus</i> base do português rural afro-brasileiro de Sapé	86
Quadro 7: <i>Corpus</i> base do português rural afro-brasileiro de Cinzento.....	86
Quadro 8: Comitativo e coordenadores em línguas europeias e nigero-congolesas	96
Quadro 9: Testes para DPs comitativos e recíprocos.	121

LISTA DE TABELAS

Tabela 1: Total de ocorrências do fenômeno variável <i>mais/com</i>	92
Tabela 2: Total de ocorrências do fenômeno variável <i>mais/e</i>	92
Tabela 3: Distribuição das variantes <i>mais/e</i> segundo o <i>tipo de categoria coordenada</i>	96
Tabela 4: Distribuição das variantes <i>mais/e</i> segundo o <i>traço semântico dos nomes</i>	98
Tabela 5: Distribuição das variantes <i>mais/e</i> segundo a <i>comunidade</i>	102
Tabela 6: Distribuição das variantes <i>mais/e</i> segundo a <i>estada fora da comunidade</i>	103
Tabela 7: Distribuição das variantes <i>mais/e</i> segundo a <i>faixa etária</i>	104
Tabela 8: Distribuição das variantes <i>mais/e</i> segundo a <i>estrutura do primeiro DP</i>	107
Tabela 9: Distribuição das variantes <i>mais/e</i> segundo a <i>referencialidade dos DPs</i>	110
Tabela 10: Distribuição das variantes <i>mais/e</i> segundo o <i>estatuto sintático dos termos</i>	114
Tabela 11: Distribuição das variantes <i>mais/com</i> segundo a <i>função semântica do PP</i>	122
Tabela 12: Distribuição das variantes <i>mais/com</i> segundo a <i>comunidade</i>	123
Tabela 13: Distribuição das variantes <i>mais/com</i> segundo a <i>faixa etária</i>	124
Tabela 14: Distribuição das variantes <i>mais/com</i> segundo a <i>estada fora da comunidade</i>	126
Tabela 15: Distribuição das variantes <i>mais/com</i> segundo o <i>gênero</i>	126
Tabela 16: Distribuição das variantes <i>mais/com</i> segundo o <i>estatuto sintático do PP</i>	129
Tabela 17: Distribuição das variantes <i>mais/com</i> segundo a <i>referencialidade do DP</i>	131

SUMÁRIO

INTRODUÇÃO	13
1 O TEMA	16
1.1 MECANISMOS SINTÁTICOS DE FORMAÇÃO DE ESTRUTURAS COMPLEXAS.....	17
1.1.1 Coordenação	17
1.1.2 Subordinação	19
1.2 DESCRIÇÃO DOS IVs ENVOLVIDOS NOS FENÔMENOS VARIÁVEIS.....	22
1.2.1 Catacterização do IV <i>e</i>	22
1.2.2 Caracterização do IV <i>com</i>	23
1.2.3 Caracterização do IV <i>mais</i>	25
1.2.3.1 O <i>mais</i> segundo a tradição gramatical.....	25
1.2.3.2 O português popular e o português rural afrodescendente.....	27
1.3 COORDENAÇÃO DE DPs E COMITATIVO NAS LÍNGUAS.....	30
1.3.1 Coordenação de DPs e comitativo em pidgins e crioulos: Parkvall (2012)	33
1.3.2 Coordenação de DPs e comitativo em pidgins e crioulos: APiCS (2013)	36
1.3.3 Coordenação de DPs e comitativo em diferentes línguas: Stassen (2000)	40
1.4 CONSIDERAÇÕES PARCIAIS.....	43
2 A TEORIA E O MÉTODO	45
2.1 A LÍNGUA PORTUGUESA NO BRASIL.....	45
2.1.1 A história sociolinguística do Brasil	47
2.1.2 Considerações acerca das línguas africanas no Brasil	51
2.1.3 Hipóteses para a formação do português brasileiro	56
2.1.3.1 A deriva secular e a confluência de motivos no PB.....	57
2.1.3.2 O português teria se crioualizado no Brasil?.....	59
2.1.3.3 A transmissão linguística irregular.....	61
2.2 TEORIAS DE FORMAÇÃO DE LÍNGUAS EM SITUAÇÃO DE CONTATO.....	65
2.2.1 Bioprograma da Linguagem (BICKERTON, 1988)	72
2.2.2 Relexificação (LEFEBVRE, 1998, 2001; LEFEBVRE E LUMSDEN, 1994)	73
2.3 A SOCIOLINGUÍSTICA.....	77
2.3.1 O estudo da variação	78
2.3.2 O estudo da mudança	80
2.4 O MÉTODO.....	81
2.4.1 As comunidades analisadas	81

2.4.2 O corpus	84
2.4.3 O tratamento dos dados	86
2.4.3.1 As variáveis dependentes.....	87
2.4.3.2 As variáveis independentes	88
2.4.4 O Varbrul	91
2.5 CONSIDERAÇÕES PARCIAIS	93
3. A ANÁLISE DOS DADOS	95
3.1 ESTRATÉGIAS DE COORDENAÇÃO ENTRE DPs: MAIS/E	95
3.1.1 Condicionamentos sociais	101
3.1.1.1 Comunidade.....	101
3.1.1.2 Estada fora da comunidade.....	103
3.1.1.3 Faixa etária	104
3.1.2 Condicionamentos linguísticos	105
3.1.2.1 Estrutura do primeiro DP coordenado	105
3.1.2.2 Referencialidade dos DPs coordenados.....	109
3.1.2.3 Estatuto sintático dos termos coordenados.....	111
3.1.3 Síntese dos fatores condicionantes da variável <i>mais/e</i>	115
3.2 ESTRATÉGIAS DE SUBORDINAÇÃO EM ESTRUTURAS COMITATIVAS: MAIS/COM.....	115
3.2.1 Condicionamentos sociais	122
3.2.1.1 Comunidade.....	123
3.2.1.2 Faixa etária	124
3.2.1.3 Estada fora da comunidade.....	125
3.2.1.4 Gênero	126
3.2.2 Condicionamentos linguísticos	127
3.2.2.1 Estatuto sintático do PP encabeçado pelo conector <i>mais/com</i>	127
3.2.2.2 Referencialidade do DP encabeçado por <i>mais/com</i>	130
3.2.3 Síntese dos fatores condicionantes da variável <i>mais/com</i>	132
CONSIDERAÇÕES FINAIS	133
REFERÊNCIAS	136
ANEXOS	147

INTRODUÇÃO

O português brasileiro (doravante, PB) é marcado pela diversidade linguística e a sócio-história do Brasil aponta para um passado de contato étnico e linguístico entre diferentes povos. Essa diversidade, associada a fatos sócio-históricos, tem levado pesquisadores a investigar o contato da língua portuguesa (doravante, LP) com outras línguas, durante os três séculos de colonização, principalmente, com as línguas africanas, uma vez que, como demonstra a história, por bastante tempo os africanos e afrodescendentes compunham grande parte da população brasileira, diferentemente das populações indígenas que não estiverem presentes massivamente durante esse período. Como aponta Mattos e Silva (2004), o segmento afrodescendente é o principal agente na propagação da língua portuguesa, em sua forma popular, pelo território brasileiro.

Dentre os inúmeros fenômenos variáveis que se observam no português popular do estado da Bahia, está o uso do item de vocabulário (doravante, IV) *mais* como coordenador entre sintagmas nominais (*determiner phrases*, DPs) alternando com o IV *e*, como em (1), e como subordinador em estruturas comitativas¹ variando com o IV *com*, como em (2), relações que ampliam a funcionalidade do IV *mais*.

(1) João *e/mais* Maria saíram.

(2) Pedro saiu *com/mais* Ana

A investigação desse fenômeno variável é norteada pela hipótese de que o uso do IV *mais* como coordenador e como subordinador resulta do contato entre línguas que marcou a história sociolinguística do Brasil e a existência de fenômenos semelhantes em pidgins e crioulos pode corroborar para a confirmação dessa hipótese. Neste trabalho, assume-se que as peculiaridades do PB devem-se, não só a motivações internas à estrutura da língua, mas, principalmente, a motivações externas, sócio-históricas, considerando o contato massivo entre falantes de línguas indígenas, africanas e o português trazido pelo colonizador (LUCCHESI, 2000, 2001, 2009). Esse fato justifica a escolha do *corpus* composto por amostras de fala de comunidades rurais afrodescendentes que vivem em relativo isolamento e, portanto, devem

¹ Nesta investigação, verificou-se ainda o uso do IV *mais* subordinador em estruturas chamadas, pela literatura, recíprocas, como em João casou *mais/com* Ana. Sempre que o termo “comitativo” aparecer de maneira geral opondo-se à estrutura de coordenação, refere-se tanto a construções com valor comitativo quanto com valor de reciprocidade.

guardar não só costumes culturais bastante antigos, mas também padrões linguísticos que permitem encontrar evidências de processos de variação e mudança resultantes do contato entre línguas.

Foram analisadas amostras de fala de quatro comunidades rurais afrodescendentes do interior da Bahia: *Helvécia*, localizada no município de Nova Viçosa, no extremo sul do estado da Bahia; *Cinzento*, localizada no município de Planalto, no semiárido; *Sapé*, localizada no município de Valença, na região do Recôncavo Baiano; e *Rio de Contas*, na Chapada Diamantina. As amostras de fala vernácula analisadas fazem parte do acervo do *Projeto Vertentes do Português Popular do Estado da Bahia*² (doravante, Projeto Vertentes), coordenado pelo Prof. Dr. Dante Lucchesi, na Universidade Federal da Bahia.

Realizado numa perspectiva Sociolinguística Quantitativa (LABOV, 1966, 1994, 2008 [1972]), neste trabalho, que objetiva fazer uma descrição sócio-histórica do uso do IV *mais* como coordenador e como subordinador, foram analisados 48 inquéritos distribuídos em três faixas etárias: Faixa I (20 a 40 anos), Faixa II (40 a 60 anos) e Faixa III (mais de 60 anos). Além da *faixa etária*, foram controladas as variáveis sociais *gênero* e *estada fora da comunidade*.

Os resultados encontrados, a partir da quantificação dos dados, parecem apontar para a confirmação da hipótese proposta, de que a ampliação funcional do IV *mais* se deve ao multilinguismo ocorrido no Brasil, durante os três séculos de colonização. Os resultados das variáveis sociais fixadas revelam que o IV *mais*, como coordenador e como subordinador, é favorecido pelos falantes mais velhos e pelos indivíduos que permaneceram na comunidade. Fixaram-se ainda variáveis linguísticas a fim de verificar os condicionamentos para as variantes. A *referencialidade* e o *estatuto sintático do DP* foram selecionados como relevantes pelo programa de análise estatística, o VARBRUL, e os resultados para o IV *mais* coordenador e subordinador parecem apontar para um traço [+referencial] desse IV. Verifica-se a existência de fenômeno semelhante em pidgins e crioulos, línguas formadas em condições sócio-históricas específicas, bem como em línguas africanas que chegaram ao Brasil. Assume-se que o IV *mais* da LP pode ter passado por um processo de relexificação na formação do PB, por conta do contato linguístico com o substrato africano.

² “O Projeto tem como objetivo central traçar um panorama sociolinguístico do português popular do Estado da Bahia, considerando os seguintes parâmetros: a relevância do contato entre línguas na sua formação histórica, por um lado, e os processos atuais de difusão linguística a partir dos grandes centros urbanos, por outro. Esses dois parâmetros definem os dois grandes vetores que atuam sobre a fala popular atualmente, condicionando os processos de variação e mudança que caracterizam e individualizam as diversas variedades do português popular no país”. Na página <www.vertentes.ufba.br> é possível encontrar mais informações a respeito do projeto.

Esta dissertação encontra-se organizada em três capítulos. No capítulo 1, *O Tema*, são feitas considerações a respeito dos mecanismos sintáticos de formação de estruturas complexas, coordenação e subordinação, bem como a caracterização dos IVs envolvidos nos fenômenos variáveis: *e/com/mais*. Apresenta-se ainda uma descrição desse fenômeno em diferentes línguas, a fim de relacionar as regiões onde o fenômeno ocorre na África à região de onde vieram as línguas africanas que chegaram ao Brasil, no período colonial. No capítulo 2, *A teoria e o método*, apresentam-se algumas considerações sobre a língua portuguesa no Brasil e sobre o contato entre línguas, as teorias sobre formação de línguas em situação de contato, e os pressupostos teóricos e metodológicos da Sociolinguística Quantitativa que orientam esta pesquisa variacionista, bem como a descrição dos métodos que a fundamentam empiricamente. No capítulo 3, apresenta-se a análise dos dados referentes aos dois fenômenos variáveis investigados, desse modo, esse capítulo encontra-se dividido em duas grandes seções, referentes a cada variável dependente. Por fim, nas considerações finais, apresenta-se a síntese dos resultados e dos aspectos mais importantes abordados neste trabalho em defesa da hipótese norteadora desta pesquisa.

1 O TEMA

No português popular do estado da Bahia, verifica-se o emprego do IV *mais* como conector tanto em construção coordenativa como subordinativa³: i) como coordenador, o *mais* alterna com a conjunção aditiva *e* (João *e/mais* Maria saíram); ii) como subordinador, o *mais* varia com a partícula comitativa *com* (João saiu *com/mais* Maria). Baxter (1987) já havia apontado o uso variável do IV *mais* como um vestígio crioulizante em dialetos rurais do PB. A fim de investigar tal uso, realizou-se uma análise variacionista, de acordo com os princípios teóricos e metodológicos da Sociolinguística Quantitativa, assumindo-se que os fenômenos variáveis *e/mais* e *com/mais* no português popular do estado da Bahia resultam do contato da LP com outras línguas faladas nos três primeiros séculos de colonização do Brasil, sobretudo com as línguas africanas. Nessa perspectiva, buscou-se analisar o uso do IV *mais* como conector no português rural afro-brasileiro, mais especificamente em amostras de fala vernácula de quatro comunidades afrodescendentes que fazem parte do acervo do Projeto Vertentes.

Este capítulo inicia-se com algumas considerações a respeito dos mecanismos sintáticos de coordenação e subordinação, cujo resultado é a formação de estruturas complexas. Além disso, é feita uma caracterização dos IVs envolvidos na realização do fenômeno variável: *e/com/mais*. Considerando que o fenômeno investigado neste trabalho, a possibilidade de um mesmo IV exercer a função de coordenador de sintagmas nominais (DPs⁴) e introdutor de DP comitativo, pode ser observado em línguas que sofreram processos de pidginização e crioulização (cf. Capítulo 2), reforçando a hipótese assumida, apresenta-se a descrição desse fenômeno em diferentes pidgins e crioulos, sobretudo os do Atlântico.

Além disso, são apresentadas algumas considerações a respeito do trabalho tipológico de Stassen (2000) sobre as *NP-conjunctions* (estratégias de coordenação entre DPs) em diferentes línguas, buscando verificar se, no continente africano, a localização geográfica das línguas que realizam coordenação entre DPs por meio da estratégia comitativa encontra um paralelo com a localização geográfica de pidgins e crioulos em que se verifica tal fenômeno. Objetiva-se, então, relacionar as línguas africanas que chegaram ao Brasil, no período colonial, às regiões onde o fenômeno ocorre na África.

³ A definição para esses termos, assumida neste trabalho, apresenta-se na seção 1.1 a seguir.

⁴ Utilizou-se a nomenclatura da teoria gerativa para se referir aos sintagmas de acordo com a seguinte legenda: DP = *Determiner Phrase*; PP = *Prepositional Phrase*; AdjP = *Adjective Phrase*; VP = *Verbal Phrase*; IP = *Inflectional Phrase*; CP = *Complementizer Phrase*.

Na seção seguinte, são feitas algumas considerações a respeito dos mecanismos sintáticos de combinação de unidades gramaticais: a coordenação e a subordinação, que envolvem, entre outros, o uso dos conectores *mais* (em seu uso popular), *e* e *com*.

1.1 MECANISMOS SINTÁTICOS DE FORMAÇÃO DE ESTRUTURAS COMPLEXAS

A *coordenação* e a *subordinação* são processos de construção que formam unidades complexas a partir da combinação de unidades gramaticais, porém alguns aspectos diferenciam sintaticamente tais processos. De modo geral, há um consenso na definição desses processos por parte da tradição gramatical (ALMEIDA, 2005; BECHARA, 2009; AZEREDO, 2008; dentre outros), embora em alguns poucos aspectos apresentem uma ou outra distinção. A formalização das regras gramaticais assenta na visão de que a coordenação envolve termos sintaticamente independentes/autônomos, ao passo que a subordinação pressupõe relação de dependência e complementação entre os constituintes envolvidos, como se vê nas subseções a seguir.

1.1.1 Coordenação

Uma forma de as unidades gramaticais se combinarem é por meio de coordenação. Segundo Matos (2003, p. 551), esse processo constroi unidades complexas a partir da combinação de constituintes do mesmo nível categorial, ou seja, de núcleos expandidos, que desempenham as mesmas funções sintáticas e semânticas.

Na proposta de Matos (2003), uma estrutura coordenada caracteriza-se pela presença de uma conjunção, ou seja, trata-se de um constituinte que tem por núcleo uma conjunção que explicita o nexo entre os termos coordenados. A coordenação assindética, como em (1), é utilizada nos casos de enumerações com mais de dois termos coordenados, mas também é possível ocorrer apenas entre dois termos. Mesmo não apresentando conjunção lexicalmente realizada, o nexo pode ser inferido nessas construções.

- (1) Maria comprou arroz, feijão, carne.

Azeredo (2008, p. 294) afirma que, no processo de coordenação, a combinação ocorre entre unidades gramaticais “que se situam no mesmo nível de modo que a presença de uma independe da presença da outra”. Segundo Matos (2003, p. 574), tal processo pode ocorrer

entre constituintes de diferentes categorias: orações, como em (2), e sintagmas dos tipos nominal, verbal, adjetival, adverbial, preposicional, conforme exemplos em (3); a coordenação pode ainda combinar núcleos lexicais em unidades complexas, como em (4), a seguir.

- (2) [As crianças foram ao cinema] *e* [nós ficamos em casa].
- (3) a. DP: [O presidente] *e* [o secretário] cumprimentaram-se.
 b. IP: O rapaz [saiu de férias] *e* [viajou para o exterior].
 c. AdjP: Uma casa [antiga] *e* [confortável] pode custar caro.
 d. AdvP: Ele só executa algumas tarefas [depressa] *e* [bem].
 e. PP: Todos gostam [do João] *e* [da Maria].
- (4) O [pai *e* patrão] do Antônio é o Sr. Gonçalves.⁵

(MATOS, 2003, p. 574-575)

Para Bechara (2009), o tipo de conector conjuntivo estabelece a relação semântica entre os termos coordenados e classifica os tipos de coordenação. Haspelmath (2004) afirma que a coordenação pode ser monossindética, com a ocorrência de apenas um coordenador como em (5), ou bissindética, com dois coordenadores conforme exemplo em (6). A estrutura coordenada também pode ser assindética, quando há apenas a justaposição dos termos coordenados, como em (1), apresentado anteriormente. Os exemplos a seguir são do autor.

- (5) Iraqw⁶
kwa/angw nee du'uma
 lebre e leopardo
 “a lebre e o leopardo”
- (6) Upper Kuskokwim Athabaskan⁷
dineje i « midzish i «
 alce com rena com
 “alce e rena”

(HASPELMATH, 2004, p. 2)

⁵ Com exceção do exemplo em (3b), todos os demais são de Matos (2003). Alguns são adaptados. Os exemplos que não apresentarem referência a outros autores são meus.

⁶ Língua da família Afro-asiática falada na Tanzânia.

⁷ Língua indígena do interior do Alasca.

A estrutura assindética, conforme exemplo em (1), é atestada no português, assim como a coordenação monossindética; contudo, não se verifica o tipo de coordenação bissindética. Haspelmath (2004) afirma que em algumas línguas europeias, como o inglês, o uso do marcador conjuntivo não leva em conta o significado dos termos coordenados, porém, muitas línguas fazem uso de coordenadores diferentes a depender de fatores semânticos. Um dos fatores apresentados pelo autor é a *animacidade* dos termos coordenados, que é relevante para a distinção entre *me* e *ma* em Nêlêmwa⁸. Em Takia, língua oceânica falada em Papua Nova Guiné, nomes com traço semântico [+ humano] são coordenados pela conjunção comitativa *da*, como em (7a), enquanto nomes com traço [- humano] o são por justaposição, como exemplo em (7b):

- (7) Takia
 a. *Meit Kabun da*
 Meit Kabun COM
 “Você e eu”
 b. *mau dabel fud*
 “taro, inhame e banana”

(HASPELMATH, 2004, p. 9)

Dos casos de coordenação apresentados nesta subseção, interessa à pesquisa a coordenação entre sintagmas nominais, como em (3a), nos moldes estabelecidos neste capítulo, uma vez que neste trabalho investiga-se o uso variável dos IVs *mais/e* nesse tipo de estrutura.

1.2.2 Subordinação

A diferença entre a subordinação e o processo apresentado na subseção anterior está, primeiramente, na “noção de hierarquia”. Na subordinação, as unidades gramaticais combinadas situam-se em níveis distintos de forma que uma unidade subordinada está contida em uma unidade maior, estabelecendo uma hierarquia gramatical (AZEREDO, 2008, p. 294). Essa ideia é compartilhada por diversos autores⁹.

Nesse processo, o termo subordinado desempenha uma função sintática e uma função semântica em relação ao termo subordinante, diferentemente do que acontece na coordenação,

⁸ Dialeto da língua Kanak falada na Nova Caledônia, um arquipélago localizado no Oceano Pacífico.

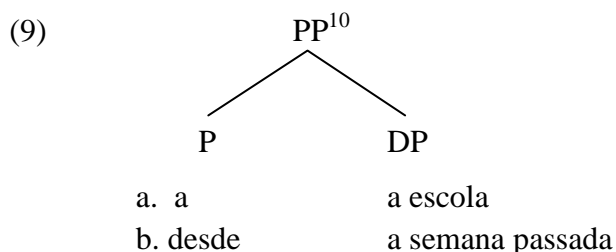
⁹ Almeida (2005), Bechara (2009), Matos (2003).

cujos termos coordenados desempenham a mesma função sintática e semântica de maneira independente. Além disso, o tipo de conector subordinativo, conjunções subordinativas e preposições, na terminologia da tradição gramatical, estabelece a relação semântica entre os termos subordinados. Quanto a esse aspecto, Azeredo (2008, p. 295) acrescenta que a unidade subordinada adquire uma função sintática, os “subordinantes são indicadores de funções novas”.

Para este trabalho, interessa a operação de subordinação estabelecida pela preposição. Bechara (2009) afirma que a preposição “se junta a substantivos, adjetivos, verbos e advérbios para marcar as relações gramaticais que elas desempenham no discurso, quer nos grupos unitários nominais, quer nas orações”. Essas categorias podem constituir o termo subordinado, enquanto o termo subordinante pode ser de diferentes categorias, como em (8). Os exemplos são de Bechara (2009):

- (8)
- a. casa de Pedro_n
 - b. pulou de contente_{adj}
 - c. gosta de estudar_v
 - d. ficou por aqui_{adv}

Na proposta de Matos (2003, p. 392-393), as preposições ou locuções prepositivas formam um sintagma preposicional juntamente com a categoria que introduz (um DP ou CP). O PP pode ser *complemento* de um DP, AdjP ou VP; pode ser *predicativo do sujeito* depois dos verbos *ser*, *estar* e semelhantes; bem como pode ser *adjunto* de várias categorias, sobretudo de VP. O PP é representado de acordo com o seguinte esquema estrutural:



No exemplo em (9), o núcleo P é seguido de um DP, a formação padrão do PP, conforme Azeredo (2008, p. 149). O autor afirma que semanticamente esta categoria

¹⁰ A autora utiliza outra nomenclatura, mas decidi modificar a fim de padronizar o trabalho. Matos refere-se ao PP como SP (sintagma preposicional), e ao DP como SN (sintagma nominal).

sintagmática é bastante versátil por conta dos diversos valores semânticos e mesmo da polissemia das preposições. Segundo Azeredo (2008, p. 196-197), em uma construção subordinativa, a preposição pode ser selecionada pelo seu significado (preposição lexical), como em (10), ou pode ser imposta pelo contexto sintático, sendo selecionada pelo seu predador (preposição funcional), como ocorre nos exemplos em (11):

- (10) a. Maria viajou **com** a família.
 b. Maria viajou **pelo** Nordeste.
 c. Maria viajou **entre** os meses de abril e junho.
- (11) a. Maria casou_v **com** João.
 b. Maria está confiante_n **em** mais uma vitória.¹¹

Em (10), a preposição é selecionada pelo significado que atribui ao termo que introduz e constitui com ele um sintagma com a função de adjunto de companhia em (10a), adjunto de lugar em (10b) e adjunto de tempo em (10c). Nesse caso, a preposição lexical atribui caso e papel temático ao DP que introduz. Já em (11), as preposições *com* e *em* fazem parte do verbo em (11a) e do nome em (11b) ao qual se ligam e os DPs introduzidos por elas exercem a função de complemento dos núcleos verbal e nominal, respectivamente. A preposição funcional em (11) atribui apenas caso ao DP que introduz; o papel temático desse DP é atribuído pelo predador, verbal ou nominal, que o seleciona.

Neste trabalho investiga-se o uso variável dos subordinantes *mais/com* nas funções de adjunto, como em (10a), e de complemento de verbos, como em (11a), nos moldes estabelecidos neste capítulo. Vale ressaltar que, neste trabalho, utiliza-se o termo subordinado apenas com o objetivo de diferenciar as posições sintáticas em que ocorre o DP introduzido pela preposição. Não se assume, considerando o objetivo de realizar uma pesquisa sócio-histórica de cunho variacionista, se esse fenômeno é realmente uma subordinação em que há dois papéis temáticos, um para cada DP envolvido na construção, ou se há coordenação entre os termos que possuem o mesmo papel temático por serem estruturas simétricas/paralelas. Mira Mateus *et. al.* (2003, p. 397) afirmam que uma frase com estrutura subordinada em que o PP possui valor comitativo, (12a) e (12b), pode ser parafraseada por coordenação (12c) e por construções que exprimem reciprocidade (12d):

¹¹ Os exemplos em (10) e (11b) são adaptados de Azeredo (2008, p. 196-197).

- (12) a. Eu dancei *com* a Maria.
 b. A Maria dançou *comigo*.
 c. Eu *e* a Maria dançamos juntos.
 d. Eu e a Maria dançamos um *com* o outro.

Os trabalhos sintáticos de Lakoff e Peters (1969) e Kayne (1994) consideram que as formas em (12a) e (12c) são sinônimas, o que é questionado nos trabalhos de Dimitriadis (2005) e Ilari (1987). Como já foi dito, neste trabalho não se assume que haja ou não sinonímia nessas estruturas, tendo em vista o objetivo inicial desta pesquisa.

1.2 DESCRIÇÃO DOS IVs ENVOLVIDOS NOS FENÔMENOS VARIÁVEIS

Esta seção apresenta uma caracterização dos IVs envolvidos nos fenômenos variáveis investigados nesta pesquisa: *e*, *com* e *mais*. Apresenta-se ainda o uso do IV *mais* como conector em estrutura coordenativa e subordinativa no português popular do estado da Bahia e no português rural afrodescendente.

1.2.1 Caracterização do IV *e*

O valor semântico atribuído ao conector coordenativo *e* é sempre o de adição (MOURA NEVES, 2000; MATOS, 2003; AZEREDO, 2008; BECHARA, 2009). Moura Neves (2000) considera que, por marcar uma relação de adição, o *e* possui um caráter mais neutro em relação a outros coordenadores, a não ser que os termos coordenados mantenham entre si uma relação semântica marcada, como de contraste, em (13a), e causa-consequência, em (13b).

- (13) a. Depenava frangos *e* não ganhava nada.
 b. Superministro arma crise *e* entra na fritura.

(MOURA NEVES, 2000, p. 739)

Segundo Moura Neves (2000, p. 741), os elementos coordenados pelo *e* podem ser de diferentes tipos: elementos de composição de uma palavra, (14a); palavras, (14b); sintagmas, (3); orações, como em (13), ou até mesmo enunciados. Também pode ocorrer coordenação

com *e* entre um sintagma oracional e um sintagma não oracional, como no exemplo da autora em (14c), desde que ambos tenham o mesmo estatuto sintático.

- (14) a. Os *leva e traz* ocupam-se de falar da vida alheia.
 b. Só poderei chegar por volta das três, *três e meia*.
 c. *Atirando o guardanapo sobre a mesa e com um temor nos lábios*, exclamou:

O conectivo *e* liga sintagmas que exercem a mesma função sintática: sujeitos (15a), objetos indiretos (15b), adjuntos (15c), dentre outros, como se vê nos exemplos de Azeredo (2008, p. 301):

- (15) a. *O porteiro e o zelador* conhecem o dono do apartamento.
 b. Ofereça café *ao porteiro e ao zelador*.
 c. O condomínio não pode funcionar bem *sem porteiro e sem zelador*.

Em suma, a leitura dos exemplos em (13), (14) e (15) mostra que o conector coordenativo *e* liga elementos ou termos. O uso de *e* implica sempre a presença de “pelo menos mais um”, e, por conseguinte, pode-se dizer que esse IV possui o traço de *adição de um segundo elemento*. O fenômeno variável *mais/e*, observado no português popular do estado da Bahia, ocorre em contextos sintáticos como o de (15a).

1.2.2 Caracterização do IV *com*

Nos termos da tradição gramatical, o conector subordinativo *com* é considerado uma preposição essencial. Uma busca pelos seus significados em diferentes trabalhos (BECHARA, 2009; MOURA NEVES, 2000; MIRA MATEUS *et. al.*, 2003) mostra que os primeiros valores semânticos atribuídos são os de *companhia* (comitativo), *ajuntamento* e *simultaneidade*, como nos exemplos em (16):

- (16) a. Estar ***com*** *alguém*. (companhia)
 b. Conversar ***com*** *alguém*. (ajuntamento)
 c. Comprar um livro ***com*** *o outro*. (simultaneidade)

Bechara (2009) considera que cada preposição, de modo geral, tem um significado primário que se desdobra em outros significados contextuais, particulares, que surgem da experiência de mundo de cada indivíduo. Nessa perspectiva, sobre a preposição *com* o autor afirma que:

A língua portuguesa só atribui a *com* o significado de “copresença”; o que, na língua, mediante o seu sistema semântico, se procura expressar com esta preposição é que, na fórmula *com + x*, *x* está sempre presente no “estado de coisas” designado. Os significados ou sentidos contextuais, analisados pela nossa experiência de mundo e saber sobre as coisas (inclusive as coisas da língua, que constitui a nossa competência linguística) nos permitem dar um passo a mais na interpretação e depreender uma acepção secundária (BECHARA, 2009).

O autor exemplifica tal posição com as frases *cortar o pão com faca*, *dancei com Marlit* e *estudei com prazer* e explica que em todos esses casos subtende-se a “presença”, seja da faca, de alguém ou do sentimento de prazer no evento. Além disso, no primeiro caso, entende-se ainda que a faca foi o instrumento utilizado para realizar o ato de cortar o pão; no segundo exemplo, emerge o sentido de companhia; e no terceiro caso, entende-se que o prazer também foi o modo como a ação verbal se realizou.

Considerando o que Bechara diz a respeito de um significado base que pode ser ampliado, é possível perceber que a preposição *com* permite uma leitura de adição nas construções, principalmente as de companhia.

Diversos valores semânticos são encontrados para a preposição *com* em diferentes estudos: companhia, ajuntamento, simultaneidade, instrumento, adição, modo, condição, causa, tempo, conformidade, lugar, concessão, meio, oposição, maneira, posse e conteúdo. Entretanto, apenas os usos que apresentam uma noção mais evidente de copresença/adicação são relevantes para a investigação realizada.

Moura Neves (2000) afirma que a preposição *com* funciona no sistema de transitividade, introduzindo complemento de verbo e de nomes, como os exemplos em (17), e também fora do sistema de transitividade introduzindo adjuntos de verbos e de nomes, como se vê em (18), expressando diferentes relações semânticas.

- (17) a. Maria namora_v **com** João.
 b. Estou desgostoso_{adj} **com** vocês.
- (18) a. Morei_v dez anos **com** meus avós.
 b. Pedro é casado_{adj} **com** Ana.

De acordo com Azeredo (2008, p. 296), a preposição *com*, sintaticamente, é um transpositor quando forma PPs que podem alternar com sintagmas adjetivais e adverbiais, como ocorre respectivamente nos exemplos em (19a) e (19b) do autor:

- (19) a. Viajar **com** *os amigos* (= viajar *acompanhado*).
 b. Misturou a massa **com** *as mãos* (= misturou a massa *manualmente*).

O autor explica que a transposição é “um mecanismo que permite expandir infinitamente os enunciados, mediante a utilização de um número limitado de meios – os transpositores – e de um número limitado de relações semânticas fundamentais” (AZEREDO, 2008, p. 296).

Os exemplos em (16), (17), (18) e (19a) mostram que, assim como ocorre com o IV *e*, o uso do conector subordinativo *com* também implica sempre a presença de “pelo menos mais um”, ou nas palavras de Bechara “copresença”. Desse modo, é possível dizer que o IV *com* também possui o traço de *adição de um segundo elemento*. O fenômeno variável *mais/com*, no português popular do estado da Bahia, investigado neste trabalho, ocorre em contextos sintáticos como o de (17a) e (18a).

1.2.3 Caracterização do IV *mais*

Nesta seção, apresenta-se inicialmente o IV *mais* de acordo com a tradição gramatical, e, em seguida, suas formas de uso no português popular do estado da Bahia e no português rural afrodescendente, que se constitui o objeto de estudo deste trabalho.

1.2.3.1 O *mais* segundo a tradição gramatical

Nesta subseção, apresentam-se possibilidades de análise do IV *mais*, com a finalidade de buscar os traços sintáticos e/ou semânticos que licenciam sua variação com os IVs *e* e *com*.

De acordo com a tradição gramatical¹², o IV *mais* expressa o valor semântico de intensificador: é um pronome adjetivo indefinido quando intensifica um nome, como em (20), ou um advérbio quando intensifica adjetivos, verbos e advérbios como em (21).

¹² Almeida (2005), Bechara (2009), Cunha (1986), e Rocha Lima (2006).

(20) Pronome adjetivo indefinido

Chegaram *mais* flores_n para Maria.

(21) Advérbio

- a. Pedro está *mais* douto_{adj.}.
- b. Meu irmão correu_v *mais*.
- c. João mora *mais* longe_{adv.}.

Como se pode observar em (21), quando o *mais* é um advérbio, a leitura obtida é sempre comparativa. Desses exemplos, infere-se uma relação de comparação a um estado ou situação anterior ou a alguém. Nesse contexto, *mais*, de alguma maneira, é um “predicador”, mesmo quando o termo requerido para que se efetive a comparação não seja realizado. Além do grau comparativo, o IV *mais* pode expressar o grau superlativo, quando o adjetivo é elevado ao grau máximo. Esse IV pode participar tanto das construções do superlativo absoluto analítico como do superlativo relativo. Em (22), são apresentados alguns exemplos dessas funções retirados de Almeida (2005).

- (22) a. O *mais* estudioso é meu filho. (superlativo absoluto analítico)
- b. João é o *mais* estudioso dos colegas. (superlativo relativo)

Embora, segundo Almeida (2005, p. 150-154), o superlativo absoluto modifique “a coisa expressa pelo substantivo, sem fazer referência à outra coisa congênere (...)”, como no exemplo em (22a), é possível inferir o universo em que se enuncia tal comparação. Diferentemente, o superlativo relativo, “além de atribuir a certa coisa uma qualidade no grau sumo, põe em relação essa coisa com outras congêneres (...)”, como em (22b). Nesse caso, a relação entre duas porções da sentença é bastante clara.

Além disso, o IV *mais* pode funcionar como pronome substantivo, como em (23a), quando precedido de artigo, pronome adjetivo ou numeral, e ainda como pronome substantivo indefinido, como em (23b), quando tiver como referente um nome que não consta na frase. Os exemplos são de Almeida (2005, p. 195-196).

- (23) a. Quem é fiel no menos é fiel *no mais*.
- b. Ele perdeu *mais* do que eu.

Em (24), apresentam-se as características do IV *mais*, segundo a tradição gramatical.

- (24) a. É um modificador: pronome indefinido ou advérbio
 b. Exprime intensidade
 c. É um “predicador”: um termo que reclama a presença de outro

Sintaticamente, o IV *mais* é considerado, pela tradição gramatical, um adjunto adnominal ou adjunto adverbial. Alguns dicionários já registram a possibilidade de o IV *mais* comportar-se como uma preposição ou como uma conjunção, porém é um uso no português popular, conforme apresentado na seção seguinte. A característica sintática em (24c) permite uma leitura de *acréscimo*, e aproxima o IV *mais* dos IVs *e/com*. É possível afirmar que *e/com/mais* permitem a leitura de *adição de um segundo elemento*. Esses traços podem licenciar a variação *mais/e* e *mais/com* no português popular e no português rural afrodescendente, como se vê na seção seguinte.

1.2.3.2 O português popular e o português rural afrodescendente

No português popular do estado da Bahia, é muito comum o uso do IV *mais* estabelecendo relação de subordinação, variando com IV *com*, ou de coordenação entre partes da sentença, variando com IV *e*, como em (25).

- (25) a. SUBORD: Ana namora *mais* Paulo. = Ana namora *com* Paulo¹³.
 b. COORD: João *mais* Maria saíram. = João *e* Maria saíram.

Observa-se que o *mais* introduz um complemento oblíquo, como em (25a), estabelecendo uma relação hierárquica de subordinação entre o verbo e seu argumento. Nesse contexto, o IV, além da função de introdutor de argumento, exerce ainda a função de atribuidor de caso ao DP. Ao introduzir um DP adjunto, o *mais* também atribui papel temático ao DP, além do caso, como em (26).

- (26) Viajei *com/mais* Antônio.

¹³ Embora não seja essa a estrutura assumida pela tradição gramatical para esse verbo, optou-se por utilizá-la, pois é um uso muito comum no português popular do estado da Bahia.

O IV *mais*, com a função de preposição, também pode ser utilizado com verbos que expressam a ideia de reciprocidade como em (27); de combinação, como em (28); ou estabelecer uma relação semântica de companhia introduzindo DP comitativo, como em (26).

(27) João casou *com/mais* Alice.

(28) Pedro juntou suas economias *com/mais* as do irmão para comprar o carro.

Nos exemplos em (27) e (28), como introdutora de complemento verbal, a preposição é considerada funcional, pois não atribui papel semântico ao DP que a complementa. Já em (26), como introdutora de um adjunto adverbial, é uma preposição lexical e atribui caso e papel temático ao DP adjacente.

Considerando as características propostas para o *mais* na subseção anterior, pode se verificar que o uso desse IV nos dialetos populares amplia sua funcionalidade. Tais usos podem ser observados no português rural afrodescendente, além do uso prescrito pela tradição gramatical, como intensificador, como se pode ver nos exemplos em (29), do *corpus*:

(29) a. **conjunção**: “eu *mais* ele ia, pegava o ônibus” (HV12)

b. **preposição funcional**: “Denestina (...) é casado *mais* Nezim...” (HV22)

c. **preposição lexical**: “Cantava *mais* os ôto (...) eu sozinha num cantava não” (CZ10)

d. **advérbio**: “era *mais* alto do que aquele pé de coquêro” (HV03)

e. **pronome indefinido**: “isso já tá cum *mais* de quatro mês” (HV07)

Constituem-se objeto de estudo deste trabalho: i) o uso do IV *mais* em construções como (29a), como coordenador entre DPs; ii) e o uso do IV *mais* conforme exemplos em (29b) e (29c), em que exerce a função de subordinador, introduzindo complemento ou adjunto de verbo com valor comitativo, no português rural afrodescendente do estado da Bahia.

Da leitura dos exemplos em (29), verifica-se que o *mais*, ao longo de sua trajetória no português, tem perdido traços, o que amplia suas possibilidades de uso, como se observa em línguas que passaram por situação de contato. É possível afirmar que, quando varia com os IVs *e/com*, o IV *mais* perde traços que o identificam como modificador (do nome, do verbo, do adjetivo e do advérbio) e passa a ter o papel de conector; sintaticamente, o IV *mais* se gramaticaliza resguardando o traço semântico de *acréscimo*.

Segundo Lucchesi (2000, p. 111), durante o contato massivo entre línguas, elementos bastante gramaticais são perdidos e outros elementos são introduzidos num processo de reconstituição do sistema linguístico. Essa reconstituição pode se dar “por duas vias que envolvem processos de gramaticalização”¹⁴: i) ampliação (ou maximização) das funções dos (poucos) itens gramaticais que se conservaram no uso linguístico da comunidade, e ii) utilização de itens lexicais para expressar relações gramaticais, ponto de partida do processo de gramaticalização.

Desses pressupostos, pode-se inferir que o uso da variante *mais*, como em (29a), (29b) e (29c), seja explicado tendo em vista o primeiro desses itens. Neste trabalho, assume-se a possibilidade de um processo de gramaticalização sintática em que o determinante ou modificador *mais* passa a um elemento de conexão, expressando a ideia de adição/acréscimo, ora como um coordenador, ora como uma preposição funcional ou lexical.

Neste trabalho, assume-se que muitas mudanças que se observam no PB, principalmente em sua variedade popular, foram motivadas pelo contato com falantes de diversas línguas no período de colonização do Brasil, principalmente com falantes de línguas africanas, conforme Lucchesi (1994, 2001, 2009). Tal variedade pode possuir aspectos semelhantes a línguas que sofreram processos de pidginização e criouliização (cf. Capítulo 2), o que justifica buscar em línguas pidgins e crioulas ocorrências do fenômeno semelhante ao investigado neste trabalho: a possibilidade de um mesmo IV funcionar como coordenador entre DPs e como partícula comitativa, conforme o IV *mais* no português popular do estado da Bahia e no português rural afrodescendente. Esse será o tema da seção 1.3.

Feitas as considerações iniciais sobre a caracterização dos IVs *mais/com/e* envolvidos na realização do fenômeno variável, sobre o uso do *mais* como conector no português popular do estado da Bahia e no português rural afrodescendente, bem como a respeito dos processos sintáticos de coordenação e subordinação que envolvem o uso desses conectores, na seção seguinte, apresenta-se uma descrição do fenômeno investigado – um mesmo IV na função de coordenador de DPs e partícula comitativa – em diferentes línguas.

¹⁴ O conceito de gramaticalização utilizado por Lucchesi refere-se “a mudança de uma categoria léxica para uma funcional, associada à perda de conteúdo lexical”, conforme propõe Roberts (1997).

1.3 COORDENAÇÃO DE DPs E COMITATIVO NAS LÍNGUAS

A natureza polissêmica das partículas de comitativo¹⁵ tem sido estudada em diversos trabalhos. Heine e Kuteva (2002, *apud* ARAÚJO, 2013) listam dez possíveis processos de gramaticalização que têm o comitativo como origem. Stolz, Stroh e Urdze (2006, *apud* ARAÚJO, 2013), em um trabalho tipológico, elaboram um mapa semântico dos casos em que uma mesma partícula é usada como comitativo e com outras funções gramaticais. Haspelmath (2004) trata da natureza polissêmica dos coordenadores e afirma que, ao se olhar para diferentes línguas, observam-se padrões recorrentes de polifuncionalidade. Em (30), apresentam-se exemplos do autor:

- (30) a. Iraqw *NEE*: conjuntivo, comitativo, instrumental, agente
 b. Hausa *DÀ*: conjuntivo, comitativo, instrumental, existencial
 c. Fongbe *KPÓÓ ... KPÓ*: comitativo, instrumental, modo
 d. Upper Kuskokwim Athabaskan *I «*: conjuntivo, comitativo, ‘também’
 e. Dargi *-RA*: conjuntivo, ‘também’
 f. Riau Indonesian *SAMA*: conjuntivo, comitativo, instrumental, recipiente
 g. English *WITH*: comitativo, instrumental

¹⁵ Segundo Ignácio (2009), a tradição gramatical tem considerado o comitativo como Adjunto Adverbial de Companhia (ROCHA LIMA, 1973; CUNHA, 1970; FARACO e MOURA, 1999, dentre outros). Contudo, no *Dicionário gramatical de verbos do português contemporâneo do Brasil* (BORBA, 1991) e no *Dicionário de usos de português do Brasil* (BORBA, 2002) o comitativo é sempre classificado como complemento, embora não se mencione o termo “comitativo”. Para Ignácio (2009, p. 17), “o comitativo se caracteriza por participar da estrutura oracional na posição de complemento, porém com função idêntica a do sujeito da oração (...). Do ponto de vista da estrutura sintática (ou sintagmática), vem normalmente introduzido pela preposição *com*, mas pode também, comutativamente, ser introduzido pela locução *junto com*”.

Para o autor, o sujeito e o comitativo podem associar-se mediante um processo de coordenação que forma um sujeito composto, fato que comprova, nas palavras do autor, “a identidade dos papéis semânticos desses dois constituintes”.

- a) O patrão discutia *com o empregado*.
 b) O patrão e o empregado discutiam.
 c) Eles discutiam.

Ignácio (2009, p. 17) salienta que o *comitativo* deve ser distinguido do *destinatário* que ocorre em estruturas semelhantes introduzidas por *com*. no caso deste último, não se verifica “identidade de funções semânticas entre sujeito e complemento, a preposição pode ser substituída por *a* ou *para* e o complemento pode comutar-se com o pronome *lhe*”.

- a) Vou falar com o diretor sobre esse assunto.
 b) Vou falar ao diretor.
 c) Vou falar-lhe.

h. Russian *-OM/ -OJ/ -JU* (= caso instrumental): instrumental, agente¹⁶
(HASPELMATH, 2004, p. 16)

É possível acrescentar à lista de Haspelmath o *com* da língua portuguesa, que pode ser comitativo e instrumental, como mostra o exemplo de Araújo (2013) em (31):

- (31) a. Estou *com* uma faca na mão.
b. Eu corto o pão *com* uma faca.

Entre as línguas africanas, como o baulê, o ewe, o fon, o quicongo, dentre outras, é bastante comum a *identidade*¹⁷ entre marcas de comitativo e instrumental (ARAÚJO, 2013). Outro fenômeno comum é a *sobreposição* de marcadores, que pode ser observada, dentre outras línguas, no Crioulo de Reunião, conforme exemplo em (32) retirado do *Atlas of Pidgin and Creole Language Structures* (APiCS)¹⁸, em seu Capítulo 70. O marcador *ek* pode ser usado para expressar comitativo e instrumental, enquanto *ansanm* é usado exclusivamente para o comitativo:

- (32) a. *Li travay ek son tonton.*
3SG trabalhar **com** POSS 3SG tio
“Ele trabalha com o tio.” (comitativo)
- b. *Lontan té i fé èk sak trésé na vakoi.*
no passado PST FIN do **with** bag woven of vacoa
“No passado, isso foi feito com um saco feito de Vacoa.” (instrumental)

¹⁶ Tradução minha.

¹⁷ A *identidade* aqui é entendida conforme Michaelis *et. al.* (2013) para os casos em que, em determinada língua, há apenas um marcador para expressar mais de uma função gramatical. A *diferenciação* indica que há dois marcadores diferentes para duas funções, e a *sobreposição* se refere à existência de dois marcadores, em que um deles se refere a apenas uma função e o outro pode ser usado para expressar mais de uma função gramatical. Exemplos de línguas com os valores de *identidade*, *diferenciação* e *sobreposição* apresentam-se ao longo desta seção.

¹⁸ O APiCS é uma publicação da *Oxford University Press* que apresenta dados de estruturas gramaticais e lexicais de 76 pidgins e crioulos. A obra contém tanto as línguas do Atlântico e Oceano Índico, as mais estudadas, quanto línguas pidgins menos conhecidas, crioulos da África, do Sul da Ásia, do Sudeste da Ásia, da Austrália e da Melanésia, variedades extintas, além de várias línguas mistas. O APiCS *On-line* contém informações sobre 76 línguas e 130 características estruturais, reunidas por uma equipe de 88 colaboradores. É uma publicação independente editada por Susanne Michaelis, Philippe Maurer, Martin Haspelmath e Magnus Huber, com o apoio do *Max Planck Institute for Evolutionary Anthropology* e do *Deutsche Forschungsgemeinschaft*. Encontra-se disponível no seguinte endereço eletrônico: <<http://apics-online.info/>>.

c. *Alon partir la butik ansanm li.*
 IMP. IPL ir DEF loja **com** 3SG
 “Vamos para a loja com ele” (comitativo)

(MICHAELIS *et. al.*, 2013, p. 277)

Outra possibilidade encontrada nas línguas é a identidade entre marcadores de comitativo e possessivo. De acordo com Araújo (2013), uma das principais estratégias para expressar a posse nas línguas do mundo é o comitativo em construções como “X está com Y” equivalendo a “X tem Y”, como atestam os trabalhos tipológicos de Heine (1997) e Stassen (2009). As grandes representantes desse uso são as línguas bantas, a exemplo da língua Tsuana, porém outras línguas também possuem essa construção, como afirma Stassen (2000).

(33) Tsuana (Nigero-congolesa, banto do Sudoeste)
Ke na le madi.
 1sg cop **com/poss** dinheiro
 “Eu tenho dinheiro.” (Eu estou com dinheiro)

(ARAÚJO, 2013, p. 13).

A leitura do exemplo em (30) mostra ainda que um padrão recorrente nessas línguas é o uso de um mesmo IV para expressar comitativo e para fazer coordenação, como ocorre com o IV *mais* no português popular do estado da Bahia e no português rural afrodescendente. No quimbundo, língua africana falada em Angola, o IV *ne* significa *e* (= *and*, do inglês) e *com* (= *with*) no sentido comitativo, conforme dados do documento histórico de Dias (1697), apresentados em Bonvini (2008, p. 109 e 116). Esse documento¹⁹ revela que, em Salvador, no século XVII, os africanos oriundos de Angola utilizavam a língua quimbunda para a comunicação, uma das línguas do substrato africano no Brasil.

As subseções seguintes fazem referência ao uso de um mesmo IV para expressar comitativo e para fazer coordenação entre DPs em diferentes línguas. Parkvall (2012) e Michaelis *et. al.* (2013) descrevem o fenômeno em pidgins e crioulos; o trabalho tipológico de Stassen (2000) organiza as línguas investigadas em *WITH-languages*, que coordenam DPs por meio da estratégia comitativa, e em *AND-languages*, que possuem dois IVs diferentes, um para cada função. Além de apresentar o fenômeno variável em diferentes línguas, os autores situam geograficamente tais línguas. Para esta pesquisa é relevante a localização dessas

¹⁹ *Arte da língua de Angola, oferecida a Virgem Senhora N. do Rosario, Mãe & Senhora dos mesmos Pretos, pelo P. Pedro Dias da Companhia de Jesu*, Lisboa (1697). Petter (2006, p. 127) afirma que o trabalho de Dias “testemunha o emprego corrente, naquela época, de uma língua africana, o quimbundo, (...) trata-se de uma língua plenamente africana, próxima da que se fala hoje em Angola”.

línguas no continente africano, sobretudo nas regiões de origem dos africanos que chegaram ao Brasil nos primeiros séculos de colonização.

1.3.1 Coordenação de DPs e comitativo em pidgins e crioulos: Parkvall (2012)

Parkvall (2012) afirma que o uso de um marcador nas funções de coordenador entre DPs e comitativo pode ser observado em diferentes pidgins e crioulos. Tal fenômeno pode ser observado em 19 línguas, como mostra o trabalho de Stolz (1998, *apud* PARKVALL, 2012). Contudo, é entre os crioulos Atlânticos que esse fenômeno se mostra mais frequente. De acordo com o autor, o rol de conjunções nesses crioulos foi incrementado “quer pela reintrodução de morfemas [da língua lexificadora], quer pela atribuição de um papel conjuncional a outros itens”.

O Quadro 1 apresenta diferentes crioulos Atlânticos em que se observam uma conjunção homófona, sincrônica ou diacronicamente, de uma preposição comitativa (PARKVALL, 2012, p. 139-140).

Variedade	Conjunção	Étimo	Fonte
CI <i>gullah</i>	/lʌŋ(ə)/	<i>Along</i>	Warantz (1986, p. 87)
CI da Costa Miskito	/wi/	<i>With</i>	Holm (1988, p. 206)
CI de St. Kitts	<long>	<i>Along</i>	Baker e Bruyn [orgs.] (1999)
CI <i>sranan</i>	/(n)ʌŋga/	<i>Along</i>	Holm (1988, p. 206)
CI saramacano	/ku/	<i>Com</i>	McWhorter (1997b, p. 46)
CI de Camarões	/weti ~ witi/	<i>With</i>	Todd (1982, p. 70)
CF da Luisiana	/avek ~ ave/	<i>Avec</i>	Valdman e Klingler (1997, p. 137)
CF do Haiti	/ak ~ akɛ ~ avek ~ ave/	<i>Avec</i>	Orjala (1970, p. 36)
CF de Santa Lúcia	/ɛk ~ ɛvek/	<i>Avec</i>	Carrington (1984, p. 125-126)
CF da Guiana	/kɛ/	<i>Avec</i>	St-Jacques-Fauquenoy (1972)
CP de Cabo Verde	/ku/	<i>Com</i>	Almada (1961, p. 136)
CP da Guiné-Bissau	/ku/	<i>Com</i>	Bartens (1996, p. 125)
CP de São Tomé	/ku/	<i>Com</i>	Lorenzino (1998, p. 191)
CP de Príncipe	/ki/	<i>Com</i>	Günther (1973, p. 80)
CP angolár	/ki/	<i>Com</i>	Lorenzino (1998, p. 192)
CP de Annobón	/ku/	<i>Com</i>	Stolz (1986, p. 238)
CH <i>negerhollands</i>	/mi/	<i>Met</i>	Stolz (1986, p. 237)
CH de Berbice	/mɛtɛ/	<i>Met</i>	Kouwenberg (1994c, p. 163)
CE papiamento	/ku/	<i>Com</i>	Holm (1988, p. 206)

Quadro 1: Etimologia de conjunções em línguas crioulas do Atlântico

Fonte: PARKVALL (2012, p. 139-140)

No caso do Crioulo Português de Cabo Verde, o exemplo do Quadro 1 refere-se apenas às variedades de Sotavento, já que, nos dialetos do Barlavento, utiliza-se a forma /ma/, do português *mais* (PARKVALL, 2012, p. 140), como ocorre no português rural afrodescendente do estado da Bahia. Segundo Baptista (2002, p. 134), no Crioulo cabo-verdiano o IV *ku* pode funcionar como conjunção coordenativa, adicionando DPs, como em (34a), pronome forte (PF) e nome, como em (34b), dois nomes, como em (34c), e pronome forte e genitivo, como em (35). A coordenação aditiva, nos demais contextos, é realizada pela conjunção *i*.

(34) a. *Mi ku nha pai.*

PF e meu pai
“Meu pai e eu.”

b. *Mi ku Brankinha.*

PF e Brankinha
“Brankinha e eu.”

c. *Nha mai ku nha pai.*

minha mãe e meu pai
“Minha mãe e meu pai.”

(35) *A(y)es ku dises*

PF e deles
“Eles e deles mesmos”

(BAPTISTA, 2002, p. 134)

Nos exemplos em (34) e em (35), observa-se que os termos coordenados possuem os traços [+definido, +específico]²⁰, comportamento semelhante verifica-se no português rural afro-brasileiro, em que a referencialidade é um fator condicionante para a variante *mais*, como se vê no capítulo 3 deste trabalho. No crioulo cabo-verdiano, o IV *ku* funciona também como uma preposição (*com*), introduzindo um DP com valor de companhia, como em (36a), e de instrumento, como em (36b).

²⁰ Inferência nossa.

(36) a. *E fika ku povu la.*

CL ficar **com** pessoas lá

“Ele fica com as pessoas lá.”

b. *E sapa ku faka.*

CL cortar **com** faca

“Ele corta com uma faca.”

(BAPTISTA, 2002, p. 133)

Parkvall (2012, p. 140) afirma ainda, baseado em diversos trabalhos, que a homofonia entre *com* e *e* é mais difundida no Níger-Congo e em toda a costa ocidental africana nas mais diferentes línguas, como: wolof, fula, manjako, mandinka, ewe, fon, fânti, twi, iorubá, engenni, hauçá, teque, dentre outras, e em um grande número de línguas bantas.

O autor afirma que as conjunções das línguas da África Ocidental caracterizam-se por dois fatores:

Em primeiro lugar, na grande maioria das línguas nigero-congolesas, contrariamente ao que acontece nas línguas europeias, a conjunção que coordena sintagmas nominais não pode ser usada para unir sintagmas verbais ou sentenças (Welmers, 1976, p. 129). Em segundo lugar, a coordenação de sintagmas nominais é frequentemente derivada e/ou homófona de uma preposição comitativa, isto é, uma palavra que significa ‘com’. (PARKVALL, 2012, p. 139)

Uma palavra africana que significa *com* não equivale a todas as formas europeias de *e*, mas apenas a algumas, como mostra o Quadro 2:

	Adposição comitativa	Morfema usado para ligar sintagmas nominais	Morfema usado para ligar sintagmas verbais ou frases
Línguas europeias	COM	E	
Línguas nigero-congolesas	COM		E

Quadro 2: Comitativo e coordenadores em línguas europeias e nigero-congolesas

Fonte: PARKVALL (2012, p. 140)

Segundo Parkvall (2012, p. 141), o padrão apresentado no Quadro 2 referente ao uso de uma mesma partícula para expressar o comitativo e para conectar DPs é atestado em diversas línguas, incluindo o iorubá e o hauçá, que vieram para o Brasil durante os séculos de colonização (cf. Capítulo 2). Sobretudo nos crioulos Atlânticos, é possível perceber o uso de

um conector de DPs, derivado de uma partícula comitativa, em distribuição complementar com um conector de orações/frases, como mostrado a seguir.

	Para ligar sintagmas nominais	Para ligar orações/frases
Sranan	/ɲ(n)anɣa/ (< E <i>along</i>)	/ɲ(h)ɛn/
Saramacano	/ku/ (< P <i>com</i>)	/ɲ(h)ɛn/
CI de Camarões	/witi/ (< I <i>with</i>)	/an/ (< I <i>and</i>)
Angolar	/ki/ (< P <i>com</i>)	/i/ (< P <i>e</i>)
CH <i>negerhollands</i>	/mi/ (< H <i>met</i>)	/en/ (< H <i>en</i>)
CH de Berbice	/metɛ/	/an/
CE papiamentu	/ku/ (< S <i>con</i>)	/i/ (< E <i>y</i>)

Quadro 3: Conector de sintagmas nominais, frases e orações em línguas crioulas.

Fonte: PARKVALL (2012).

No português rural afrodescendente do estado da Bahia, verifica-se o padrão apresentado no esquema acima, o mesmo das línguas nigero-congolesas, como mostrado no Quadro 2, o que corrobora para a hipótese do contato linguístico. O conector *mais* é utilizado para coordenar DPs, variando com o IV *e*, e para introduzir DP comitativo, variando com o IV *com*. O *mais* não coordena frases e orações, tal função é exercida pelo *e* ou por outros conectores²¹. De outro modo, na variedade culta do PB, verifica-se o padrão das línguas europeias apresentado no Quadro 2, pois tanto a coordenação entre DPs como a coordenação entre orações com valor semântico aditivo são feitas pelo conector *e*.

A subseção seguinte dá continuidade à descrição do fenômeno investigado nesta pesquisa em pidgins e crioulos, a partir dos dados apresentados por Michaelis *et. al.* (2013) no APiCS.

1.3.2 Coordenação de DPs e comitativo em pidgins e crioulos: APiCS (2013)

Dentre as línguas reunidas no APiCS, Michaelis *et. al.* (2013) mostram, no Capítulo 71, que a identidade de marcas entre comitativo e coordenador de DPs é encontrada em 34% das línguas estudadas, conforme exemplos em (37) e (38):

- (37) Crioulo de Martinica (base francesa)
 a. *Joj épi Jéra ay péché.*
 Joj **com** Jéra ir pescar
 “Joj e Jéra foram pescar.”

²¹ O conector *e* possui valor semântico aditivo. Frases ou orações também podem ser ligadas por conectores com outros valores semânticos, como adversativo, alternativo, conclusivo e explicativo.

b. *Joj ay péché épi Jéra.*

Joj ir pescar **com** Jéra

“Joj foi pescar com Jéra”

(MICHAELIS *et. al.*, 2013, p. 280)

Em (37a), observa-se o uso de *épi* como coordenador dos DPs [*Joj*] e [*Jéra*]; em (37b), a forma *épi* é usada para expressar o valor comitativo, introduzindo o adjunto [*Jéra*]. De acordo com Michaelis *et. al.* (2013, p. 280), esse mesmo crioulo também pode usar os marcadores *ek* ou *é* para ambas as funções.

(38) Sango (base africana)

a. *Laso mbi na mo, i ke gwe biani.*

hoje eu **com** você nós COP ir verdadeiramente

“Hoje, você e eu vamos, com certeza.”

b. *Mbi lango na ita ti mbi.*

Eu dormir **com** irmão de eu

“Eu morava com meu irmão.”

(MICHAELIS *et. al.*, 2013, p. 280)

O exemplo em (38a) apresenta o uso de *na* como coordenador dos DPs [*você*] e [*eu*], enquanto em (38b) a mesma forma é usada para expressar o comitativo, introduzindo o adjunto [*meu irmão*].

Os casos em que há *diferenciação*, ou seja, em que a partícula comitativa é diferente do conector de DPs, como em (39), representam 41% das línguas que aparecem no APiCS.

(39) Krio (base inglesa)

a. *Di uman en ã pikin don go na os.*

ART mulher e POSS filho PFV ir LOC casa

“A mulher e seu filho têm ido para casa.”

b. *Di uman don go na os wit ã pikin.*

ART mulher PFV ir LOC casa **com** POSS filho

“A mulher foi para casa com seu filho.”

(MICHAELIS *et. al.*, 2013, p. 280)

Como se vê em (39a), a coordenação entre DPs é feita por meio da partícula *en*, enquanto o valor comitativo realiza-se por meio de *wit*.

As línguas que apresentam o valor *sobreposição* possuem dois marcadores, em que um deles se refere a apenas uma função e o outro pode ser usado para expressar mais de uma

função gramatical. Os casos de sobreposição, em que uma forma pode representar tanto a partícula comitativa como a coordenação entre DPs, representam 25%, a exemplo do Crioulo cabo-verdiano de São Vicente, como se vê em (40). A partícula *má* (do português *mais*) no crioulo cabo-verdiano, de base portuguesa, é utilizada com a função de coordenador dos DPs [*Adriano*] e [*eu*], e com a função de introdutor do DP comitativo [*seu irmão*], em uma relação de adjunção. Porém, a sobreposição não ocorre com a partícula *y*, que exerce a função de coordenador de DPs, mas não a função de introdutor de comitativo.

(40) Crioulo cabo-verdiano de São Vicente (base portuguesa)

a. *Mi má Adrianu trubaiá djunt n'un bárk.*

ISG **com** Adrianu trabalhar junto em=DET navio

“Adriano e eu trabalhamos juntos em um navio.”

b. *El bai pa Praia má se irmon.*

3SG ir para Praia **com** 3SG.POSS irmão

“Ele foi para a praia com seu irmão.”

c. *Relasãu entre mi y nha mãi senpr foi mut bon.*

Relacionamento entre ISG e POSS.ISG mãe sempre COP.PST muito bom

(MICHAELIS *et. al.*, 2013, p. 281)

No Kinubi, conforme mostram os exemplos em (41), a partícula *ma* é utilizada com a função de coordenador dos DPs [*O Nubi de Kibera*] e [*os de Mombasa*], e com a função de introdutor do DP comitativo [*os Britânicos*], em uma relação de adjunção. Entretanto, não se observa sobreposição com a partícula *u*, que exerce a função de coordenador de DPs, mas não a função de introdutor de comitativo.

(41) Kinubi (base africana)

a. *Núbi ta Kíbra ma Mombása ma 'endisi tofauti.*

Nubi GEN Kibera **com** Mombasa NEG ter diferença

“O Nubi de Kibera e os de Mombasa não são diferentes.”

b. *Úmun já ma British.*

3PL vir **com** Britânicos

“Eles vêm com os Britânicos.”

c. *Kan bes anás ta Mijikénda u anás ta Rabái.*

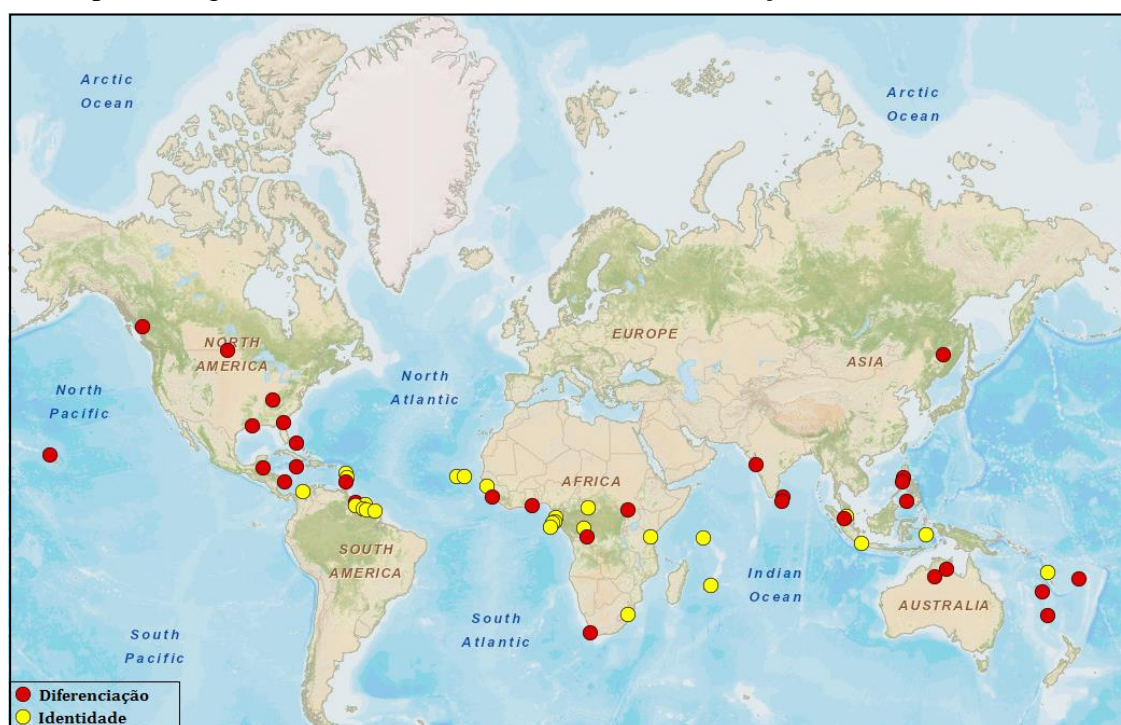
ANT só pessoas GEN Mijikenda e pessoas GEN Rabai

“Havia apenas o Mijikenda e o Rabai.”

(MICHAELIS *et. al.*, 2013, p. 281)

Diversas línguas que se formaram em condições sócio-históricas específicas, marcadas em sua origem pelo contato entre línguas, exibem o fenômeno em que uma forma pode representar tanto a partícula comitativa como a coordenação entre DPs. De acordo com o Capítulo 71 do APiCS, as línguas dessa natureza que apresentam o valor *identidade* de formas concentram-se na área atlântica. A concentração é ainda maior na África subsaariana. As línguas com o valor *diferenciação*, quase metade de base lexical inglesa, são mais comuns nas regiões Sul e Sudeste da Ásia, mas algumas línguas atlânticas também apresentam esse valor. Observa-se ainda que na América do Norte e parte superior da América Central há uma predominância de línguas crioulas em que as formas para coordenar DPs e expressar o valor comitativo são diferentes, ao passo que na parte inferior da América Central predominam línguas crioulas que apresentam identidade de formas para expressar tais funções gramaticais. O Mapa 1, a seguir, sintetiza essas informações:

Mapa 1: Línguas²² com valor de identidade e diferenciação de formas no APiCS



Fonte: APiCS On-line. Disponível em: <<http://apics-online.info/wals/71#1/14/10>>

²² Os nomes das línguas que apresentam o valor *identidade* de formas, apresentadas pelo APiCS neste mapa, constam no Anexo A (línguas da América) e no Anexo B (línguas dos demais continentes).

Na seção seguinte, apresenta-se o trabalho tipológico de Stassen (2000) sobre as *NP-conjunctions*, estratégias de coordenação de DPs em diferentes línguas, a fim de investigar a localização geográfica das línguas que realizam a coordenação de DPs por meio da estratégia comitativa no continente africano.

1.3.3 Coordenação de DPs e comitativo em diferentes línguas: Stassen (2000)

Um trabalho tipológico sobre as estratégias de coordenação de DPs (*NP-conjunction*) é realizado por Stassen (2000), a partir de uma amostra de 260 línguas selecionadas dos maiores grupos e áreas linguísticas. O autor considera que as línguas apresentam um contraste tipológico, porque podem ter duas estratégias formalmente diferentes para realizar as *NP-conjunctions*. Línguas que fazem uso de duas marcas diferentes para as estruturas “A e B” e “A com B”, ou seja, uma estratégia coordenativa e uma estratégia comitativa²³, fazem parte do grupo que o autor denomina de *AND-languages*. Por outro lado, línguas que fazem uso da mesma marca para as estruturas coordenativas e comitativas, isto é, que fazem uso apenas da estratégia “A e/com B”, são denominadas *WITH-languages*.

Segundo Stassen (2000), na estratégia de codificação coordenativa, como em (42a), os DPs coordenados possuem posição estrutural igual, mesma função sintática, mesmo papel temático e mesmo caso, além disso, formam um mesmo constituinte e estabelecem concordância. Entretanto, na estratégia comitativa, a exemplo de (42b) e (42c), os DPs referentes aos dois participantes do evento descrito pelo verbo estão em posição estrutural desigual; enquanto um pode ter qualquer caso, o outro só pode ser núcleo de um DP oblíquo; os dois DPs não são parte do mesmo constituinte e não forçam concordância. Diferentemente da estratégia coordenativa, esta é formalmente mais uniforme nas línguas e manifesta-se por meio de um marcador *com* em um dos DPs.

- (42) a. João e Maria saíram.
 b. João saiu com Maria.
 c. Maria saiu com João.²⁴

(STASSEN, 2000, p. 6)

²³ São as estruturas que se entendem como subordinadas, nos moldes estabelecidos neste capítulo.

²⁴ Tradução minha.

São *WITH-languages* o haitiano, o fongbe, o hauçá, a maioria das línguas bantas e muitas línguas oceânicas. Em (43) e (44), apresentam-se exemplos de Stassen:

- (43) Akan (Niger-Kordofanian, Kwa)

Kwesi nye Amba a-ba

Kwesi **com** Amba PST-come

“Kwesi e Amba vieram./ Kwesi veio com Amba.”

- (44) Urubu-Kaapor (Tupi)

Nasui riki ihẽ namõ i-hon

Nasui EMPH 1SG **com** 3-go

“Nasui e eu fomos./ Nasui foi comigo.”

(STASSEN, 2000, p. 22-23)

Em oposição, existem as *AND-languages*, em que há uma clara distinção entre a estratégia comitativa e a coordenativa. São exemplos praticamente todas as línguas da Europa, mas outras línguas também fazem parte desse grupo, como atestam os seguintes exemplos de Stassen:

- (45) Somali (Afroasiatic, Cushitic)

a. *waraabe iyo dawaco ayaa wada ugaadshsi tegey*

hiena e chacal FOC juntos caçar ir.3M.PST

“A hiena e um chacal foram caçar juntos.”

b. *na la ra'*

1PL **com** vir.IMP

“Vem com a gente!”

- (46) Malagasy (Austronesian, West Indonesian)

a. *misotro Raby sy Rakoto*

beber Raby e Rakoto

“Raby e Rakoto bebem.”

b. *niaraka tami-ni aho*

ir.PST com-ele eu

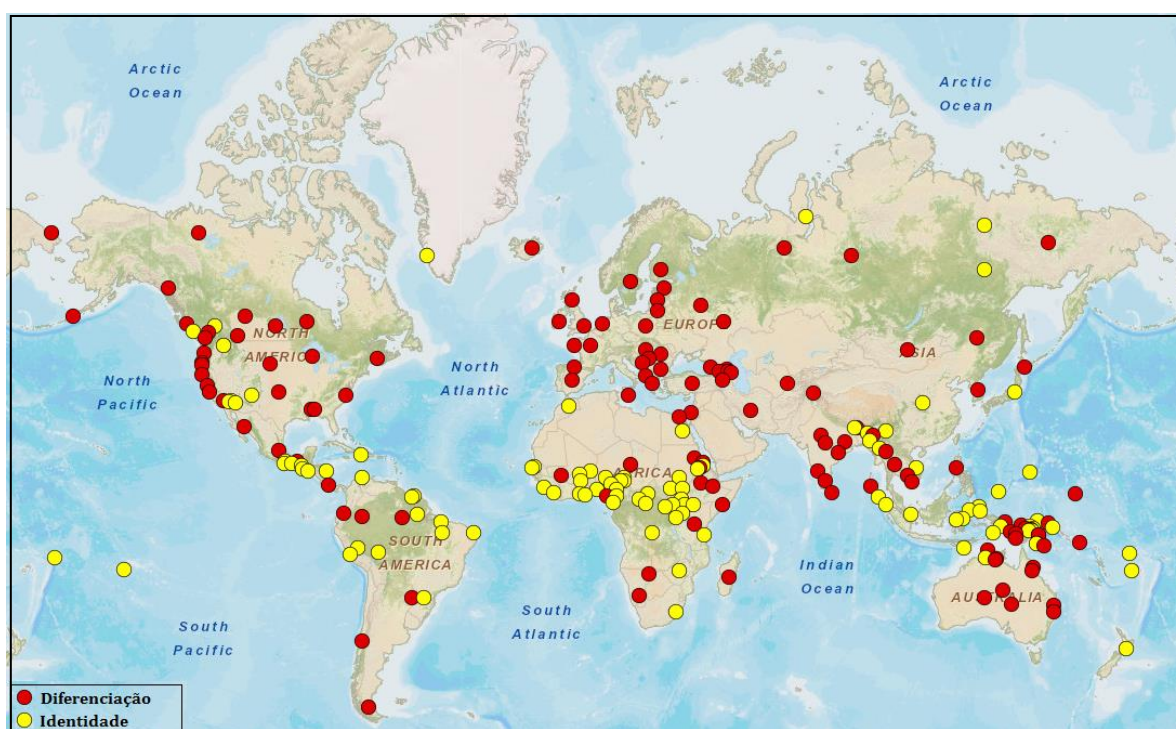
“Fui com ele.”

(STASSEN, 2000, p. 24)

De acordo com o autor, há uma diferença marcante na frequência dos dois tipos de línguas, pois há um número muito maior de *AND-languages*.

O Mapa 2, retirado do *The World Atlas of Language Structures (WALS)*²⁵ (2005a) localiza as línguas estudadas por Stassen (2000) no que tange aos valores *identidade* e *diferenciação* de marcadores para as funções de coordenador entre DPs e comitativo. Nesse mapa, observa-se que o valor *identidade* é predominante na África subsaariana e no Pacífico, e ocorre em 103 das 234 línguas investigadas, o que equivale a 44%. O valor *diferenciação* é mais comum na Europa e no Sul da Ásia; ocorre em 131 línguas representando um percentual de 56%.

Mapa 2: Línguas²⁶ com valor de identidade e diferenciação de formas no WALS



Fonte: WALS On-line. Disponível em: <<http://wals.info/feature/63A#2/16.6/153.5>>

²⁵ O WALS é um grande banco de dados de propriedades estruturais fonológicas, lexicais e gramaticais de diferentes línguas reunidos por uma equipe de 55 autores a partir de estudos descritivos. Foi publicado inicialmente em 2005 como um livro com CD-ROM, pela *Oxford University Press*, e teve sua primeira versão *On-line* em 2008. O *WALS On-line* é uma publicação do *Max Planck Institute for Evolutionary Anthropology*, editado por Matthew Dryer e Martin Haspelmath. Encontra-se disponível no seguinte endereço eletrônico: <<http://wals.info/>>.

²⁶ Os nomes das línguas deste mapa, com *identidade* de formas, no continente africano, constam no Anexo C. Os nomes das demais línguas, bem como das línguas com valor *diferenciação*, podem ser visualizadas no seguinte endereço: <<http://wals.info/feature/63A#2/16.6/153.5>>.

Como se pode observar, a língua portuguesa não foi objeto de análise do autor, nem em sua vertente europeia, nem brasileira, nem em outros países nos quais é língua oficial²⁷. Comparando-se o mapa do APiCS com o mapa do WALS, percebe-se que, em ambos, a África subsaariana concentra um grande número de línguas em que há identidade de formas para expressar coordenação de DPs e valor comitativo, como ocorre com o IV *mais* no português rural afrodescendente da Bahia. Na seção 2.1.2, busca-se verificar se há coincidência entre as regiões africanas onde esse uso é recorrente e os locais de origem dos africanos que foram trazidos para o Brasil nos três séculos de colonização, a fim de investigar se a línguas trazidas contribuíram para a formação do português rural afrodescendente no que diz respeito ao fenômeno analisado.

Vale ressaltar, por fim, que, como mostra o Mapa 2, a identidade de formas não parece ser comum entre as línguas europeias, o que corrobora a hipótese do contato entre a LP com outras línguas no Brasil para explicar o uso de um mesmo IV para as funções de coordenador de DPs e subordinador comitativo.

1.4 CONSIDERAÇÕES PARCIAIS

O capítulo 1 teve por objetivo apresentar o tema que este trabalho se propõe a investigar: a variação *mais/e* na coordenação de DPs e a variação *mais/com* em estruturas subordinadas com valor comitativo. As informações apresentadas nas seções anteriores apontaram alguns conceitos e informações relevantes para a análise desses fenômenos variáveis:

- Pode-se afirmar que os IVs *e*, *com* e *mais* compartilham o traço semântico de *adição/acréscimo*. Esse pode ser o traço em comum que licencia as variações *mais/e* e *mais/com*;
- O uso de um mesmo IV para coordenar DPs e expressar o valor comitativo é atestado em diferentes pidgins e crioulos. Nos crioulos de base lexical portuguesa, em geral, a partícula é derivada do IV *com*. No Crioulo de Cabo Verde, na variedade de Barlavento, a partícula deriva do IV *mais*;

²⁷ A Língua Portuguesa possui o *status* de língua oficial nos seguintes países: Portugal, Brasil, Cabo Verde, Angola, Guiné-Bissau, São Tomé e Príncipe, Moçambique, Timor Leste e Macau.

- O conector coordenativo de frases e orações em muitas línguas crioulas que apresentam tal fenômeno não é o mesmo que coordena DPs. Entretanto, o coordenador entre DPs possui identidade com a partícula comitativa;
- No continente africano, as línguas estudadas por Stassen (2000), em que se verifica a estratégia comitativa para coordenar DPs, localizam-se na região subsaariana, assim como os pidgins e crioulos que também apresentam tal possibilidade de construção.

Foram assumidos os seguintes posicionamentos, considerando os aspectos discutidos neste capítulo:

- i. A possibilidade de um processo de gramaticalização sintática em que o determinante ou modificador *mais* passa a um elemento de conexão, expressando a ideia de adição/acréscimo ora como um coordenador, ora como uma preposição funcional ou lexical;
- ii. Muitas mudanças ocorridas no PB, principalmente aquelas percebidas no português popular, foram motivadas pelo contato com falantes de diversas línguas no período de colonização do Brasil, principalmente com falantes de línguas africanas, conforme Lucchesi (1994, 2001, 2009);
- iii. A hipótese de que os fenômenos variáveis *e/mais* e *com/mais* no português popular do estado da Bahia resultam do contato da LP com outras línguas da colonização do Brasil, sobretudo as línguas africanas.

A hipótese em (iii), apresentada também na introdução deste capítulo, orientou a investigação a respeito da sócio-história do português brasileiro (PB), bem como do aporte teórico e metodológico da Sociolinguística, que se apresentam no capítulo seguinte. Além disso, buscou-se verificar as teorias que explicam a formação de línguas em situação de contato, o Bioprograma da Linguagem (BICKERTON, 1988) e a Relexificação (LEFEBVRE, 1998, 2001; LEFEBVRE E LUMSDEN, 1994), uma vez que se busca explicar a existência do fenômeno investigado considerando o contato entre línguas na formação do PB.

2 A TEORIA E O MÉTODO

A investigação dos fenômenos variáveis *mais/com* e *mais/e* neste trabalho, conforme apresentado no capítulo anterior, é norteada pela hipótese de que o uso do IV *mais* como conector resulta do contato entre línguas que marcou a história sociolinguística do Brasil; a existência de fenômenos semelhantes em pidgins e crioulos pode corroborar para a confirmação dessa hipótese. Neste capítulo, apresentam-se os pressupostos teóricos e metodológicos que orientam a investigação e realização desta pesquisa.

Este capítulo encontra-se estruturado da seguinte maneira: na seção 2.1, apresentam-se algumas considerações a respeito da língua portuguesa no Brasil, no que diz respeito à história sociolinguística do Brasil, considerações acerca das línguas africanas no Brasil e hipóteses para a formação do PB; na seção 2.2, abordam-se alguns conceitos relacionados à pidginização e à criouliização, assim como duas teorias de formação de línguas em situação de contato, a do Bioprograma da Linguagem e a da Relexificação; na seção 2.3, trata-se dos princípios teóricos e metodológicos da Sociolinguística Quantitativa; por fim, na seção 2.4, descreve-se o método empregado para a realização desta pesquisa, de acordo com o modelo variacionista proposto por Labov (1994, 2008 [1972]).

2.1 A LÍNGUA PORTUGUESA NO BRASIL

A língua portuguesa, além de oficial desde o século XVIII²⁸, é falada em todas as regiões do Brasil pela maior parte da população, embora o Brasil atual possa ser classificado como um país multilíngue (RODRIGUES, 2006, p. 153). Muitos estudos têm proposto que o PB é distinto do português europeu (PE). Diferenças fonológicas, morfossintáticas e lexicais que se observam no PB em relação ao PE têm sido estudadas por muitos pesquisadores e teóricos (KATO, 2012; CYRINO, 2007; DUARTE, 1993; dentre outros) que objetivam descrever e explicar a formação da variante brasileira, que surgiu num contexto de multilinguismo, conforme Mattos e Silva (2001, 2004, 2006) e Lucchesi (2000, 2001).

²⁸ Em 1758 o Marquês de Pombal promulgou a Lei do Diretório, que proibia o uso da língua geral em favor da oficialização da língua portuguesa. Segundo Trouche (2000) “a imposição da língua portuguesa foi uma questão fundamental para Portugal, no sentido da preservação da colônia, contudo, a concorrência do português com a língua geral ainda perdurou até a segunda metade do século XVIII, quando o discurso das autoridades portuguesas se centrou numa política de difusão e obrigatoriedade do ensino da língua portuguesa”.

Disponível em: <http://www.filologia.org.br/anais/anais%20iv/civ12_9.htm>.

Para explicar o distanciamento entre o PB e o PE, as evidências linguísticas têm sido analisadas sob perspectivas diferentes: à luz dos fatores sócio-históricos que corroboraram para a formação da nação e do povo brasileiro, como fez Adolpho Coelho que, em 1880, já apontava para a interferência das línguas africanas na formação do PB, ou Gregory Guy (2005) que considera a possibilidade de formação de crioulos no Brasil devido a fatores sócio-históricos; por outro lado, à luz da lógica interna do sistema, a exemplo de Tarallo (1993) e Galves (2001) que estudam os aspectos estruturais da língua, numa perspectiva gerativista; ou mesmo como têm feito Naro e Scherre (2007), que explicam a constituição histórica do PB pela deriva secular, fundamentando-se em evidências tanto da história interna como da história externa da LP.

De acordo com Mattos e Silva (2006), não se deve focalizar somente o horizonte comparativo entre o PB e o PE, mas também as variantes sócio-histórica e linguística do PB conviventes em nosso território nacional. Lucchesi (2000, 2001), na investigação sobre o PB, propõe que a realidade linguística do PB é polarizada: por um lado, existe um conjunto de características linguísticas comuns aos falantes das normas cultas e, por outro, um conjunto comum aos falantes das normas populares usadas pelas classes mais baixas, em cuja origem se verifica forte presença dos ex-escravos, os quais aprenderam o português com seus pais que tinham o português como segunda língua. Essa polarização se delineou desde o período colonial do Brasil, como indica Lucchesi (2012, p. 5).

No caso brasileiro, a polarização sociolinguística é, na atualidade, o reflexo inexorável das abissais diferenças materiais e culturais que apartam as classes sociais no país. Mas suas raízes são mais profundas e, sobretudo, antigas, pois remontam à colonização do território brasileiro, quando o colonizador português escravizou muitos povos que falavam centenas de línguas diversas entre si – e a eles impôs sua língua, como parte do violento processo de dominação espiritual e material da escravatura –, constituindo assim, por mais de trezentos anos, a força motriz da formação da sociedade brasileira.

Considerando que essa pluralidade de normas no PB, bem como o distanciamento desse em relação ao PE, se deve à própria forma como se desenvolveu a nação e o povo brasileiro, nesta seção, será apresentada uma breve caracterização da história sociolinguística do Brasil, sobretudo no que se refere ao contato da LP com as outras línguas faladas no período de colonização do Brasil. Além disso, serão apresentadas algumas considerações acerca das línguas africanas que chegaram ao Brasil e à Bahia devido à sua grande relevância

para formação da variante brasileira, assim como três hipóteses que tentam explicar a constituição do PB.

No âmbito desta pesquisa, serão levados em conta tanto os aspectos linguísticos quanto os fatores sócio-históricos que serviram de cenário para a implementação das variações *mais/com* em estruturas comitivas e *mais/e* em contexto de coordenação entre DPs.

2.1.1 A história sociolinguística do Brasil

No início do século XVI, a LP chegou ao Brasil com os primeiros colonizadores portugueses, que encontram tribos indígenas habitando a costa. De acordo com Rodrigues (1993a, p. 91), mais de mil línguas indígenas eram faladas no início da colonização portuguesa, porém, na costa, onde primeiro se estabeleceram os portugueses, eram faladas duas línguas indígenas, o *tupi* e o *tupinambá*, da mesma família Tupi-guarani, pertencentes ao mesmo subgrupo, e, por esse motivo, muito semelhantes entre si.

Inicialmente, a colonização portuguesa, promovida por D. João III, foi feita somente por homens e, do relacionamento deles com as índias, formou-se uma sociedade de mamelucos, filhos de europeus e indígenas. Por conta da necessidade de comunicação entre esses povos nessa nova comunidade, surgiu uma língua de intercurso, denominada de língua geral, de base indígena (RODRIGUES, 1996)²⁹. Rodrigues (1996, p. 6) afirma que *língua geral* é um termo específico para determinada categoria de línguas que surgiram na América do Sul nos séculos XVI e XVII em condições especiais do contato entre europeus e povos indígenas. O autor considera que o termo toma um sentido mais definido no Brasil nos séculos XVII e XVIII para designar as línguas de origem indígena faladas pelos mamelucos.

A língua tupi serviu de base para a formação da *língua geral paulista*, que se estabeleceu a partir da sociedade de mamelucos na capitania de São Vicente, província de São Paulo, e foi levada pelos bandeirantes mamelucos para Minas Gerais, Goiás, Mato Grosso e região Sul. Já a *língua geral amazônica*, de base tupinambá, se estabeleceu a partir da população de mamelucos do Maranhão e Pará e se expandiu para o interior do Amazonas, onde ainda é falada³⁰.

²⁹ No litoral, formou-se ainda uma língua moldada pelos jesuítas, no século XVI, com base na língua dos indígenas e destinada à catequização desses povos. Alguns trabalhos referem-se a essa língua como geral (FREYRE, 2002; CARDEIRA, 2006). A língua geral a que Rodrigues se refere não é esta.

³⁰ O nheengatu, a *língua geral amazônica* modificada, é falado por uma parte dos moradores de São Gabriel da Cachoeira, no rio Negro (RODRIGUES, 2006, p. 150).

As línguas indígenas se fixaram como língua popular nos primeiros séculos de colonização do Brasil devido a diferentes fatores: i) as semelhanças entre as línguas *tupi* e *tupinambá* as tornavam intercambiáveis, portanto, tais línguas eram práticas na medida em que já eram faladas em um vasto território; ii) os mamelucos aprendiam a língua das mães e davam continuidade ao uso dessas línguas; iii) os portugueses não se estabeleceram uniformemente ao longo da costa, tanto no que diz respeito tanto ao tempo de estabelecimento como ao tipo de relacionamento com os índios; iv) a LP falada pelos portugueses não era uniforme, pois, devido aos diferentes locais de procedência dos colonos, chegaram ao Brasil diferentes dialetos do PE (RODRIGUES, 2006, p. 145-146). Consequentemente, a implantação da língua portuguesa como meio geral de comunicação foi inicialmente desfavorecida.

Ainda segundo o autor, as bandeiras contribuíram para a difusão das línguas gerais para diversas regiões do interior do país e suas expedições foram fundamentais para a expansão do território brasileiro e para a definição das fronteiras do Brasil atual. Na busca por riquezas e escravos, os exploradores subjugarão as populações indígenas para utilizar sua força de trabalho nas plantações, porém, devido a não adaptação dos indígenas ao trabalho agrícola, foi preciso buscar outra fonte de mão-de-obra para manter a exploração na Colônia. Assim, em 1549 o tráfico negreiro é oficializado no país.

De acordo com Mattoso (1990, p. 22), no século XVI, ocorreu o primeiro ciclo de importação de escravos africanos, o ciclo da Guiné, que trouxe para o Brasil africanos sudaneses para trabalhar no cultivo da cana de açúcar e do tabaco. Desse período até o fim do tráfico negreiro, em 1860, foram trazidos ao Brasil milhares de africanos de diferentes regiões. Ao final do primeiro século da Colônia, os escravos africanos já representavam 42% da população e “já se estendia a todas as capitanias, ultrapassando no conjunto, qualquer um dos outros grupos – portugueses (30%) e índios (28%)” (COUTO, 1992, *apud* MATTOS e SILVA, 2004, p. 96).

De acordo com Mattoso (1990, p. 22), Portugal adotava “a política de misturar diferentes etnias africanas, para impedir a concentração de negros de uma mesma origem numa só capitania”, desse modo, os escravos precisavam assimilar o português, mesmo que precariamente, para que pudessem se comunicar. Apesar dessa política, alguns documentos linguísticos evidenciam a existência de uma língua geral de base africana em Minas Gerais e o uso do quimbundo, uma língua plena africana, por escravos angolanos em Salvador (cf. seção 2.1.2).

O panorama dos séculos XVI e XVII permite afirmar que o Brasil passou por uma situação de multilinguismo generalizado que perdurou até meados do século XVIII (MATTOS E SILVA, 2004, p. 93). Diversos fatores contribuíram para a restrição desse cenário e, conseqüentemente, para a expansão da LP: o extermínio das populações indígenas, falantes de diversas línguas; a violência a que foram submetidos os africanos no Brasil, falantes de 200/300 línguas, que eram obrigados a não usar suas línguas maternas e adotar o português para comunicação; o decreto pombalino, que alçou a língua portuguesa ao *status* de língua oficial do Brasil no século XVIII; e a migração massiva de portugueses, que, a partir do referido século, fez crescer o acesso dos escravos aos modelos da língua alvo.

No século XVIII, por conta da descoberta das minas de ouro e diamantes, estima-se que imigraram para o Brasil 300 mil portugueses, o que contribuiu para a difusão da LP (LUCCHESI, 2009, p. 48). Callou *et. al.* (2006, p. 264) afirmam que, neste século e em todo o século XIX, o crescimento do número de brancos foi decisivo para o avanço do uso da LP.

Ainda no século XVIII, em 1763, o Rio de Janeiro torna-se a capital do Brasil e, na colônia, o crescimento de núcleos urbanos foi um reflexo direto da produção de riquezas no país. O crescimento da vida urbana na colônia favoreceu o uso da LP em suas variedades mais normativizadas, já que se criaram condições para uma produção literária mais consistente e maior difusão de padrões culturais e linguísticos institucionalizados; porém, vale ressaltar que os africanos e afrodescendentes “continuavam a assimilar precariamente o português nas mais aviltantes situações de incivilidade e, em muitos casos, conservaram as suas línguas nativas para a comunicação entre si” (LUCCHESI, 2009, p. 49).

É esse o cenário que a família real portuguesa encontra ao desembarcar no Brasil em 1808. Lucchesi (2009, p. 49) ressalta o impacto demográfico da fixação da corte na capital do país “por se tratar, não de um contingente de colonos dispersos, mas de um grupo orgânico e que constituía a elite política e cultural da época”. Entre as conseqüências de tal empreendimento, estão o crescimento da urbanização e o fortalecimento da cultura institucional, além de diversas mudanças que culminaram na independência do país em 1822.

Ainda no século XIX, ocorreu a abolição da escravidão no Brasil, em 1888. Segundo Lucchesi (2009, p. 52), a situação dos ex-cativos teve importantes implicações sociolinguísticas, pois os escravos livres receberam doações de terra e passaram a viver isolados, praticando a cultura de subsistência, ou foram para as cidades onde eram mantidos nas periferias. Em ambos os casos nos agrupamentos de ex-escravos mantinham-se as mudanças decorrentes do contato linguístico.

Dessa forma, no século XIX, a polarização sociolinguística³¹ da LP encontra-se no auge: uma pequena elite nos centros urbanos cultivava “a língua e as boas maneiras” inspiradas em Portugal, o que explica, segundo Lucchesi, a formação das variedades cultas do PB, enquanto a maioria da população (índios, africanos e mestiços) vivia segregada e explorada, dando origem às variedades populares do PB.

Nas primeiras décadas do século XX, entretanto, o processo de industrialização e de urbanização da sociedade brasileira, assim como a atuação dos meios de comunicação de massa, como o rádio e a TV, começaram a redefinir esse quadro de acentuada polarização sociolinguística em um processo de *nivelamento linguístico*, no qual os padrões de comportamento linguístico dos centros urbanos passaram a penetrar no interior do país, como afirma Lucchesi (2009, p. 536). Em relação às comunidades rurais afro-brasileiras isoladas, o autor considera que esse processo tem feito desaparecer as marcas mais acentuadas do contato linguístico por conta da integração dessas comunidades à sociedade brasileira durante o século XX.

Estaria em curso um nivelamento linguístico do interior do país a partir da influência dos modelos de prestígio das grandes cidades brasileiras. Tal processo de difusão das variantes do modelo urbano culto brasileiro seria implementado principalmente pela ação dos meios de comunicação de massa e do deslocamento populacional, com a contribuição da ampliação do sistema público de educação formal. Tal cenário sociolinguístico prevê que estejam em curso, nas comunidades de fala analisadas, processos de mudança linguística de cima para baixo nos quais a variante valorizada socialmente penetra na comunidade movida por influências externas (LUCCHESI, 2009, p. 536).

Atualmente, o Brasil caracteriza-se por um multilinguismo, pois, além do português, língua oficial e hegemônica, são faladas também, minoritariamente, línguas indígenas nacionais, europeias e asiáticas. De acordo com Rodrigues (2006, p. 153), “o Brasil atual pode ser classificado como um país multilíngue, no qual se falam cerca de 200 línguas, das quais umas 180 são indígenas e as demais são idiomas de minorias de origem europeia ou asiática”. A LP é falada em todas as regiões brasileiras, diferentemente de outras ex-colônias de Portugal, nas quais predominam línguas crioulas, a exemplo de Cabo Verde, São Tomé e Príncipe, Guiné Bissau, dentre outras. Esse *status* da LP no Brasil se deve aos motivos que já foram apresentados para sua fixação no país, sobretudo a partir do século XVIII. O multilinguismo generalizado cedeu lugar ao que Mattos e Silva (2004, p. 93) chama de multilinguismo localizado. Para a autora, a língua portuguesa é “heterogênea, plural e

³¹ Lucchesi (1994, 1998, 2001a, 2002b, 2006a).

polarizada” quando se considera o PB como um todo, não apenas a norma padrão (MATTOS E SILVA, 2006).

Na seção seguinte, são feitas algumas considerações acerca das línguas africanas e sua grande relevância para a formação do PB.

2.1.2 Considerações acerca das línguas africanas no Brasil

As línguas africanas chegaram ao Brasil com a importação maciça de africanos, de meados do século XVI até o século XIX, para trabalharem como escravos no Brasil, como apresentado na seção anterior. Petter (2006) estima que vieram para o Brasil de 200 a 300 línguas africanas procedentes de duas áreas linguísticas, a oeste-africana e a banta. Mattoso (1990) distingue quatro ciclos de importação de africanos para o Brasil:

- No século XVI: o ciclo da Guiné trouxe escravos sudaneses da região africana ao norte do Equador;
- No século XVII: o ciclo do Congo e de Angola, África equatorial e central, trouxe africanos bantos;
- No século XVIII: o ciclo da costa de Mina trouxe novamente africanos sudaneses durante todo o século XVIII. A partir de meados desse mesmo século, inicia-se o ciclo da baía de Benin;
- No século XIX: chegam ao Brasil africanos de diferentes regiões, contudo predominam os de Angola e Moçambique.

A autora afirma que “foi, pois, a cana-de-açúcar das regiões de Pernambuco, Bahia e Rio de Janeiro que obrigou à importação de escravos nos séculos XVI e XVII, enquanto o ouro gerou o aumento da demanda no século XVIII”. No século XIX, é o café a principal atividade econômica que movimenta o tráfico de escravos. (MATTOSO, 1990, p. 23).

De acordo com Castro (2002, p. 45), chegaram a Salvador as línguas *mina-jeje* (também chamada de mahi, do grupo ewe-fon/gbe), *nagô-iorubá* e *hauçá*, provenientes da região oeste-africana. Mina-jeje começou a chegar ao Brasil a partir do século XVII, mas a língua hauçá só é registrada no século XIX. As línguas bantas foram introduzidas no Brasil ainda no século XVI e continuaram chegando durante todo o período do tráfico negreiro. A autora afirma que:

Nesse emaranhado de línguas [banto], foram relativamente importantes sobre outras, no Brasil, três línguas litorâneas – *umbundo*, *quimbundo* (Angola), e *quicongo* (Angola e Congo-Brazzaville) –, consequência da procedência temporal dos seus falantes, densidade demográfica e amplitude geográfica da sua distribuição humana em território brasileiro durante quatro séculos consecutivos (CASTRO, 2002, p. 43).

Retomando as informações apresentadas no Mapa 1 (seção 1.3.2) e no Mapa 2 (seção 1.3.3), é possível perceber que, no continente africano, verifica-se a existência de línguas com o valor identidade de formas nas regiões de origem dos africanos que vieram para o Brasil e para a Bahia durante o período do tráfico negreiro. O Mapa 3, retirado do APiCS, apresenta apenas os dados referentes ao valor identidade em pidgins e crioulos no continente africano:

Mapa 3: Línguas com valor de identidade de formas na África – dados do APiCS



Fonte: APiCS On-line. Disponível em: <<http://apics-online.info/wals/71#1/14/10>>

O Mapa 4, retirado do WALS, com dados do trabalho tipológico de Stassen (2000), apresenta os dados referentes ao valor identidade em diferentes *WITH-languages* no continente africano:

Mapa 4: Línguas com valor de identidade de formas na África – dados do WALS



Fonte: WALS On-line. Disponível em: <<http://wals.info/feature/63A#2/16.6/153.5>>

Como já foi dito, as rotas do tráfico de escravos africanos para o Brasil incluem a região oeste-africana e a banta. A primeira localiza-se ao norte do Equador e compreende a região do Senegal à Nigéria (Golfo de Benin), e dos países: Gâmbia, Guiné-Bissau, Guiné-Conakry, Serra Leoa, Libéria, Burquina-Fasso, Costa do Marfim, Gana, Togo e Benin. Esta região é dominada pelos sudaneses. O domínio do povo banto recobre toda a região abaixo da linha do Equador, incluindo os seguintes países: Camarões, Gabão, Angola, Namíbia, Zâmbia, Tanzânia, Quênia, Moçambique, África do Sul (CASTRO, 2002, p. 39).

É possível verificar, nos Mapas³² 3 e 4, que o quimbundo, língua falada em Angola, não foi objeto de estudo dos autores dos trabalhos que possibilitaram tal representação. Apesar disso, um dado é importante frisar: conforme apresentado na seção 1.3, a obra de Dias (1697) revela que, no quimbundo, o *IV ne* significa *e* (= *and*, do inglês) e *com* (= *with*) no sentido comitativo (BONVINI, 2008, p. 109 e 116), como ocorre nas línguas do Mapa 4.

Como se viu na seção 1.3, os fenômenos variáveis *mais/com* e *mais/e* investigados neste trabalho encontram um paralelo com diferentes línguas do continente africano, em que um mesmo *IV* exerce as funções de coordenador de DPs e partícula comitativa. Relacionando

³² No Anexo B e no Anexo C constam os nomes das línguas representadas nos mapas.

os fatos sócio-históricos até então apresentados sobre a origem dos africanos que vieram para o Brasil, sobre as línguas que chegaram com eles e sobre as regiões na África onde se verifica fenômeno semelhante ao investigado nesta pesquisa, é possível imaginar a presença de um substrato africano no Brasil com estruturas semelhantes ao fenômeno variável *mais/com* e *mais/e* observado no português rural afrodescendente, que pudesse ter interferido na formação do PB³³.

Por muito tempo, os africanos e afrodescendentes compunham a grande parte da população brasileira. Segundo dados apresentados por Mussa (1991, p. 163), do século XVII ao XIX a presença de etnias não brancas fica no patamar de 60%. Sobre a importância dessas etnias para a difusão da LP, Mattos e Silva (2004, p. 106) considera que a massa sucessiva de escravos chegados ao Brasil, vindas de duas grandes áreas linguísticas do continente africano, teve que aprender a língua portuguesa a partir de exposições precárias. Para a autora, o segmento afrodescendente é o principal agente na propagação da língua portuguesa popular pelo território brasileiro, e fundamenta seu pensamento em fatores sócio-históricos como:

- a. A presença maciça de afrodescendentes, conforme demonstrado pela demografia histórica;
- b. A atuação de escravos na economia brasileira colonial;
- c. A mobilidade geográfica de escravos no Brasil;
- d. Os papéis desempenhados pelos escravos nas famílias do colonizador, tanto nos núcleos urbanos como na colônia rural;
- e. O significado social e linguístico dos espaços ilegítimos da escravidão.

De acordo com Petter (2006), documentos linguísticos revelam alguns dados sobre o estabelecimento de línguas africanas no Brasil do século XVII ao século XIX:

³³ Dentre as línguas estudadas por Stassen (2000) que apresentam identidade de formas para o coordenador de DPs e a partícula comitativa, encontram-se algumas línguas indígenas faladas no Brasil: *kaingang*, no Paraná, e *canela-krahô*, no Maranhão (da família macro-gê), *tupi*, no Rio Grande do Norte, e *urubu-kaapor*, no Pará (da família tupi), e *apalaí*, no Pará (da família karíbe), conforme Anexo D. Vale ressaltar que os índios logo foram dizimados ou recuaram para o interior do país. Talvez o contato tenha sido maior com os africanos. Além disso, como já foi dito, as línguas africanas que chegaram à Bahia, como quimbundo, hauçá e ioruba, por exemplo, possuem fenômeno semelhante ao investigado nesta pesquisa.

- No século XVII, os africanos angolanos utilizavam o quimbundo (língua plena africana) para a comunicação em Salvador³⁴;
- No século XVIII, há evidências de uma língua geral africana formada por um conjunto de línguas do grupo gbe³⁵;
- No século XIX, há um plurilinguismo africano, sobretudo em Salvador, onde predominava a língua nagô-ioruba³⁶. É nesse período que a imprensa e a literatura começam a registrar o “falar diferente” dos negros.

Entretanto, não há registro de línguas africanas plenas no século XX no Brasil. A autora afirma que tais línguas, desde o final do século anterior, passaram a manifestar-se como línguas especiais e seu uso está associado a grupos específicos e a duas funções principais: uso *ritual* nos cultos religiosos africanos, nas diferentes regiões do país, e *demarcação social* em algumas comunidades rurais afrodescendentes. O uso ritual das línguas africanas mantém-se como “veículo de expressão de cânticos, saudações e nomes de iniciados”. Também podem servir como meio de comunicação entre os adeptos de uma mesma comunidade de culto. Já a demarcação social está associada ao uso de “línguas secretas” que demonstram uma resistência cultural, bem como à origem africana dos seus usuários (PETTER, 2006, p. 119).

Essas línguas secretas representam um código de acesso exclusivo a um subgrupo demarcado socialmente e foram encontradas nas comunidades de Cafundó (SP) e Tabatinga (MG), que utilizam como língua de comunicação uma variedade do português regional (dialeto rural), caracterizada por um léxico reduzido de origem africana. Em linhas gerais, tem-se um léxico de origem banta (quimbundo) utilizado apenas na fala dos adultos que cumpre uma função lúdica de enganar ou ocultar algo daqueles que não conhecem a língua secreta (PETTER, 2006, p. 121 e 122).

A aquisição da LP como L2 por uma massa de escravos africanos certamente contribuiu para a alta variabilidade presente no português popular brasileiro (PPB), pois, como afirma Mattos e Silva (2004, p. 93), a voz dos afrodescendentes promoveu uma

³⁴ Conforme Petter (2006, p. 126), em *Arte da língua de Angola, oferecida a Virgem Senhora N. do Rosario, Mãe & Senhora dos mesmos Pretos, pelo P. Pedro Dias da Companhia de Jesu*, Lisboa (1697). Segundo a autora, o trabalho de Dias “é a gramática do quimbundo, falada em Salvador, pelos escravos oriundos de Angola”.

³⁵ Conforme Petter (2006, p. 127), *Obra nova da Língua Geral de Minna*, de Antonio da Costa Peixoto, Ouro Preto (1741).

³⁶ Dados revelados pelos estudos antropológicos de Nina Rodrigues, iniciado em Salvador, em 1890 (PETTER, 2006, p. 129).

reestruturação no PE que começou a chegar ao Brasil em 1500 e ao longo de todo o período colonial.

Na seção seguinte, apresentam-se três hipóteses que têm sido formuladas para explicar a constituição histórica do PB.

2.1.3 Hipóteses para a formação do português brasileiro

No debate atual sobre a constituição histórica do PB, três hipóteses tentam explicar a formação do PPB. De maneira geral, têm-se os seguintes posicionamentos teóricos:

- i. O PB é uma espécie de continuação do português arcaico (PA), com algumas alterações;
- ii. O PPB é marcado por tendências presentes nas línguas crioulizadas;
- iii. A aquisição de um modelo defectivo de português e sua nativização desencadearam um processo de transmissão linguística irregular que teve importantes consequências para a formação da atual situação linguística no Brasil, não envolvendo necessariamente a transferência de estruturas das línguas do substrato para o português brasileiro.

Na hipótese (i), defendida por Naro e Scherre (2007), o PE que chegou ao Brasil passou por mudanças que já estavam prefiguradas no sistema, desse modo, o que se observa no PB é uma ampliação das estruturas da LP ao longo de todo o seu percurso histórico. Segundo esses autores, o contato entre a LP e outras línguas no Brasil serviu apenas para acelerar o processo de deriva.

O contato linguístico é considerado a partir de outra perspectiva nas hipóteses (ii) e (iii). Na hipótese (ii), postulada por Guy (2005), o PB passou por um processo de crioulição impossível de não ter ocorrido devido aos fatores sócio-históricos imbricados em sua formação. Já na hipótese (iii), defendida por Baxter e Lucchesi (1993, 1997) e Lucchesi (2000), a variedade da língua alvo que se forma em uma situação de contato multilinguístico apresenta uma forte redução/simplificação em sua estrutura gramatical. Além disso, consideram que o contato da LP com as línguas africanas pode ter gerado crioulições leves em diferentes períodos de tempo.

2.1.3.1 A deriva secular e a confluência de motivos no PB

A deriva secular como responsável por mudanças linguísticas no PB é defendida por Naro e Scherre (2007). Os autores afirmam que a variabilidade na concordância nominal e verbal do PB, considerada como sendo de origem pós-crioula ou resultante de um aprendizado imperfeito de L2, como se verá nas seções a seguir, já estava prefigurado no novo sistema mais analítico da LP, antes mesmo de sair da Europa. Para os autores, tal estado de coisas se torna bastante plausível dada a deriva secular das línguas românicas, e indo-europeias de maneira geral, em direção à uniformização morfológica, com a sobrevivência apenas das formas irregulares mais salientes. (NARO; SCHERRE, 2007, p. 32). Desse modo, o PB é caracterizado como uma continuação do PA com algumas alterações, já que, para os autores da hipótese, ainda não foram encontrados traços do PB que já não tenham ocorrido no PA/PE.

Naro e Scherre retomam o conceito de deriva do linguista norte-americano Edward Sapir (1921), para quem as línguas são geneticamente aparentadas por serem de uma mesma origem – o proto indo-europeu – assim, não se movem no tempo e no espaço de forma aleatória, mas são resultantes de uma deriva (LINS, 2009, p. 285). Desse modo, a fundamentação da hipótese da deriva secular está principalmente em documentos do PA. Naro e Scherre buscam nesses documentos evidências dos mesmos fenômenos encontrados no PB, e analisam, dentre outros, a concordância verbal e nominal do PB.

No que se refere à concordância verbal, Naro e Scherre argumentam que a redução morfológica se deve a um fenômeno fonológico – a perda da nasalização da vogal não acentuada final –, o que também ocorre em Portugal na fala popular. Os autores afirmam que, sem influências ameríndia ou africana, o mesmo fenômeno observado no Brasil ocorre em Portugal, e vão além, afirmando que o mesmo fenômeno aparece também em textos medievais portugueses. Quanto à concordância nominal, os autores apresentam dados da variação no latim clássico e nas línguas românicas modernas e consideram que “o processo da queda do -s final no português do Brasil [teve] seu início no português dialetal da Europa, que, por sua vez, estava apenas dando continuidade a uma deriva pré-românica”.

Para Naro e Scherre (2007, p. 25), tanto a história interna como a história externa apontam para a hipótese da mudança linguística natural. Com base em dados de estudos dialetológicos, afirmam que a variação na concordância nominal no PB é um “fenômeno geral, independente de localização geográfica” e concluem que as variáveis sociais mais relevantes para o entendimento desse fenômeno são a escolarização e o contraste rural/urbano.

Para os autores, a língua portuguesa vinda da Europa já possuía uma força que a impulsionava para uma dada direção/desenvolvimento, e, no Brasil, encontrou-se com forças da América e outras vindas da África e “juntas se reforçaram para produzir o português popular do Brasil”. Daí o termo “confluência de motivos”, cunhado pelos autores para dar conta dessas forças linguísticas e sociais.

Naro e Scherre (2007, p. 67) negam que tenha havido crioulização no Brasil e afirmam que não existem evidências que apontem para um *pidgin* prévio de base portuguesa que tenha contribuído com particularidades estruturais: “rejeitamos explicitamente a posição de que o português popular do Brasil é um semicrioulo, tem uma história crioula ou tem a ele subjacente uma leve crioulização”. A influência cultural das línguas indígenas e das línguas africanas não é negada por Naro e Scherre, porém, a influência gramatical de línguas africanas ou de quaisquer outras que não a portuguesa é rejeitada (LINS, 2009, p. 287).

A proposta da deriva secular tem recebido algumas críticas. De acordo com Lucchesi (2012):

Naro e Scherre concentram todos os seus esforços, não para reunir evidências empíricas que apoiassem a hipótese da deriva, mas em coletar fatos que servissem como contra-exemplo da participação do contato linguístico na formação do português popular do Brasil. (LUCCHESI, 2012, p. 268 e 269).

O autor critica ainda o fato de que admitir que a variação do PB não está associada ao contato entre línguas, como argumentam Naro e Scherre, não significa dizer que esteja associada à deriva secular. Lucchesi (2012, p. 263) também contesta os argumentos que Naro e Scherre utilizam para explicar a variação na concordância verbal no PPB, afirmando que tal processo não poderia ter sido desencadeado por conta de uma desnasalização já existente no PE, como afirmam Naro e Scherre.

Para Lins (2009, p. 288), “a documentação escrita, notadamente do PA, de que se vale Naro e Scherre para traçar suas análises, reflete apenas as tendências (...) e não sinaliza para uma quantidade expressiva de ocorrências”. Lins também considera incoerente a generalização de que os traços encontrados no PB têm sempre um ancestral no PE, e acrescenta que

O que realmente precisa ficar mais delineado no trabalho desses estudiosos é o modo que esse “confluir de motivações” lida com a influência africana e indígena, e ainda com o multilinguismo, a mobilidade populacional, a “sociodemografia histórica” e a presença e ausência da escolarização no Brasil, haja vista serem esses aspectos significativamente consideráveis na conformação do PB (LINS, 2009, p. 290).

Por sua vez, Mattos e Silva (2006, p. 233) chama a atenção para a necessidade de estudos fundamentados na sintaxe do PA para o português moderno “não apenas no que se refere à concordância, como também a outras características sintáticas que distinguem PE e PB, como aspectos referentes ao sistema pronominal e à ordem sintática”.

Na subseção seguinte, apresenta-se outra hipótese que tem sido formulada para dar conta das particularidades do PB. Nesse caso, considera-se o contato linguístico fundamental para a formação da variante brasileira.

2.1.3.2 O português teria se criouloizado no Brasil?

Adolpho Coelho, em 1880, foi o primeiro a levantar a questão da semelhança entre o PB e as línguas crioulas afro-portuguesas ao salientar características de dialetos crioulos no PB, a exemplo da falta de concordância de número no interior do DP, em construções como *artigo + substantivo* ou *adjetivo + substantivo*, nas quais apenas uma das formas leva a marca de plural. Segundo Coelho (1880, *apud* Tarallo 1993a), os escritores brasileiros “não formam idéia clara e completa das diferenças que separam o português da antiga metrópole da linguagem do novo império”.

Porém, é com Gregory Guy (1981a, 1989) que a proposta crioulistica para a formação do PB se desenvolve de fato. Para Guy, o PB passou por um processo de crioulização inevitável dada a história sociolinguística do Brasil, conforme discutida na seção 2.1.1. De acordo com Guy, um crioulo de base portuguesa teria se formado no período de colonização do Brasil e passado por uma rápida descrioulização que deu origem às variedades do PPB. Essa descrioulização constitui um processo de aquisição de regras do PE em contextos mais salientes, por causa do contato do português popular com o português culto, a língua alvo (LUCCHESI, 2001, p. 99). John Holm (1987) se refere ao PPB como sendo um semi-crioulo, ou seja, uma variedade que compartilha traços estruturais com línguas crioulas, mas que não resultou de um processo radical de crioulização.

Guy apoia-se tanto em evidências linguísticas como em fatos históricos, demográficos, geográficos e sociais. As evidências linguísticas, segundo o autor, são:

a) Na morfologia: redução de vários tipos, como eliminação da segunda pessoa verbal nas formas -s e -ste, como em (1a), e perda de pronomes átonos, como em (1b):

- (1) a. Tu já falou com sua mãe sobre aquele assunto?
b. Encontrei ele na festa sozinho.

b) Na sintaxe: falta de concordância, como em (2a), e desuso das formas mesoclíticas, como em (2b):

- (2) a. Os meninoø comeuø o bolo todo.
b. Te convidarei para meu casamento.

c) Na fonologia: redução de codas, como em (3a), e vocalização do /l/ em sílaba final, como em (3b):

- (3) a. A gente vai jogáø bola no campinho.
b. mal [ˈmaw], mil [ˈmiw]

De acordo com Guy (2005, p. 33), a história sociolinguística do Brasil “indica a entrada e a saída de falantes de crioulos e as condições suficientes para a criouliização” e “as evidências internas do PPB indicam vários traços mais de acordo com uma história de criouliização do que com qualquer outra explicação”.

Na proposta de Guy, o PB estaria indo em direção ao PE devido ao processo de descriouliização, diferentemente da proposta de Coelho segundo a qual “o PB está sofrendo mudanças que o distanciam do alvo original” (TARALLO, 1993a, p. 36). Tarallo apresenta análises sintáticas do PB e discorda de Guy ao concluir que a possibilidade de haver descriouliização é muito remota. Contudo, Tarallo concorda que o contato do PB com outras línguas é um fato a ser considerado:

O PB é pois o resultado da alteração do PE em contato (e essa palavra é crucial aqui) com diferentes grupos étnicos e diferentes línguas, uma situação que não leva necessariamente à criouliização, mas simplesmente à diferenciação dialetal (TARALLO, 1993a, p. 59).

Além da crítica feita por Tarallo, a hipótese em questão tem sido alvo de diversas considerações. Naro e Scherre (2003) consideram que o termo *criouliização* acaba muitas vezes se confundindo com *pidginização* e questionam sobre quais documentos de fato comprovam uma língua crioula no Brasil de léxico português e gramática africana. Mattos e Silva (2006) afirma que se trata de uma hipótese muito generalizante para explicar a formação do PB. Por sua vez, Lins (2009) considera que esta hipótese não analisa consistentemente

todos os fatos e processos que dizem respeito à constituição histórica do PB, mas ressalta que seu valor, enquanto arcaçou teórico, não pode ser descartado.

Na próxima subseção, apresenta-se a hipótese da transmissão linguística irregular para explicar a formação do PB. Diferentemente da proposta crioulizante de Guy, essa hipótese pressupõe um estágio inicial de crioulização, embora não assuma que o PB tenha se tornado um crioulo.

2.1.3.3 A transmissão linguística irregular

Baxter (1992, 1995), Lucchesi (2000, 2003b) e Baxter e Lucchesi (1997, 2006) têm desenvolvido o conceito de transmissão linguística irregular (TLI) para dar conta de situações em que uma variedade histórica de uma língua se forma por força de contato multilinguístico, passando a apresentar estrutura próxima às estruturas de línguas crioulas, sem que se configure um caso típico de pidginização ou crioulização. Nas palavras de Lucchesi:

A transmissão linguística irregular é um conceito mais amplo do que o de pidginização/crioulização, pois engloba tanto os processos de mudança provenientes do contato entre línguas através dos quais uma determinada língua sofre alterações muito profundas na sua estrutura, do que resulta o surgimento de uma outra entidade linguística, quanto os processos nos quais uma língua sofre alterações decorrentes do contato com outras línguas, sem que essas alterações cheguem a configurar a emergência de uma nova entidade linguística qualitativamente distinta (LUCCHESI, 2000, p. 7).

Conforme postulada, a TLI contribui para o entendimento da constituição histórica do PB uma vez que esse se constitui uma variedade histórica da língua portuguesa que apresenta mudanças induzidas pelo contato do português com outras línguas presentes no Brasil colonial. Os pidgins e crioulos configuram um caso extremo de TLI, enquanto uma variedade histórica de determinada língua alvo está em outro nível do processo. De acordo com Lucchesi “a TLI constitui um contínuo de níveis diferenciados de socialização/nativização de uma segunda língua, adquirida massivamente, de forma mais ou menos imperfeita, em contextos sócio-históricos específicos” (LUCCHESI, 2003, p. 274).

O acesso aos modelos da língua alvo é fundamental para o processo de TLI, visto que, com o prolongamento de uma situação de contato entre línguas pode haver uma condição sociolinguística de acesso restrito ou de maior acesso aos modelos da língua alvo, e cada uma dessas condições desencadeará um processo mais ou menos intenso. Havendo uma

continuidade da situação de contato e pouco acesso aos modelos da língua alvo “a expansão gramatical do código de comunicação emergencial ocorre principalmente através dos processos de reestruturação original da gramática e através da transferência de estruturas provenientes das outras línguas” (LUCCHESI, 2003, p. 273), o que é chamado de *influências do substrato*. Tal situação dá ensejo à formação de pidgins e crioulos.

Porém, se houver uma continuidade da situação de contato e mais acesso aos modelos da língua alvo, geralmente prestigiados na comunidade de fala, tais modelos “tendem a suplantar os processos de transferências de estruturas do substrato e/ou reestruturação original da gramática” (LUCCHESI, 2003, p. 274). Nesse caso, há a formação de uma variedade da língua alvo que, contudo, apresenta certas características resultantes de sua nativização entre os falantes de outras línguas e seus descendentes.

De acordo com Lucchesi (2003, p. 275), o processo de reestruturação gramatical caracteriza a TLI em todos os seus níveis. Contudo, antes desse processo, há uma simplificação dos dispositivos morfossintáticos nas fases iniciais da TLI. Tal simplificação, de acordo com o autor, justifica-se pelos seguintes fatores:

- a. O difícil acesso dos falantes das outras línguas aos modelos da língua-alvo;
- b. Os falantes dessas outras línguas, por serem em sua maioria adultos, não mais dispõem de acesso aos dispositivos mentais da faculdade da linguagem;
- c. A ausência de ações normativizadoras.

A reestruturação gramatical ocorre nas fases seguintes e é feita com maior ou menor intensidade a depender do nível de acesso aos modelos da língua alvo, conforme foi dito acima. Lucchesi (2003, p. 274) salienta que entre as línguas crioulas também ocorrem graus variados de reestruturação gramatical. Em alguns casos, a reestruturação é mais profunda, como no Crioulo francês do Haiti e no Angolar (crioulo da Ilha de São Tomé); em outros, esse processo é menos profundo, como se deu no Crioulo francês da Ilha de Reunião e no Crioulo cabo-verdiano de base portuguesa.

Lucchesi (2003, p. 276) considera como características fundamentais do processo de TLI: (i) perda, ou variação no uso de morfologia flexional e palavras gramaticais; (ii) alteração dos valores dos parâmetros sintáticos em função de valores não marcados, que não implicam, entre outras coisas, movimentos aparentes na estruturação da sentença; e (iii) gramaticalização de itens lexicais para preencher as lacunas na estrutura linguística.

Diversas estruturas presentes atualmente no PB são atribuídas ao contato linguístico que marcou a história do Brasil colonial. Algumas delas são mais generalizantes que outras por conta dos fatores sócio-históricos atuantes em cada caso (LUCCHESI, 2003, p. 278-281). Dentre elas, estão:

a) Eliminação de certos dispositivos gramaticais mais abstratos e de uso restrito da língua alvo, como em (4). Em (4a) observa-se a eliminação da marca morfológica da 2ª pessoa da flexão verbal, que ocorre na maioria das regiões brasileiras, e em (4b) a eliminação da marca do subjuntivo.

- (4) a. Você foi ao médico ontem?
b. Talvez Maria vai na feira.

b) Manutenção da variação no esquema presença/ausência do dispositivo gramatical da língua alvo, como a ausência da flexão de número, tanto no sintagma nominal, em que apenas o primeiro elemento é marcado, como no sintagma verbal, no exemplo em (5).

- (5) Meus tio chega de viagem hoje.

c) Alteração na frequência de uso relativamente à marcação de determinados parâmetros sintáticos, a exemplo de (6), em que há uma mudança em relação ao parâmetro do sujeito nulo com a realização do sujeito pronominal.

- (6) Eu cheguei e não o encontrei em casa.

d) Recomposição da estrutura gramatical da língua alvo, eliminando a variação ou reduzindo-a a uma pequena escala, como a variação na concordância de gênero em (7):

- (7) Ana é casado com meu filho.

e) Manutenção da variação no uso do dispositivo gramatical dentro de um esquema de variável ternária, com a variante da língua alvo, (8a), uma variante oriunda de um processo original de reestruturação da gramática, (8b), e a variante zero, (8c):

- (8) a. João gostou do livro, mas terá que devolvê-lo amanhã.

- b. João gostou do livro, mas terá que devolver ele amanhã.
- c. João gostou do livro, mas terá que devolver \emptyset amanhã.

Pode-se acrescentar a esse rol de características, o fenômeno investigado neste trabalho: o uso do IV *mais* como coordenador de DPs, alternando com a conjunção aditiva *e* (João *e/mais* Maria saíram), e como subordinador em construções comitativas, alternando com a preposição *com* (João saiu *com/mais* Maria).

A hipótese da TLI, para explicar a constituição do PB, assim como a hipótese proposta por Guy, tem sido alvo de críticas, sobretudo pelos teóricos e pesquisadores Naro e Scherre (2007), que afirmam que a variação na concordância verbal e nominal, bem como outras estruturas que têm sido atribuídas ao contato da língua portuguesa com outras línguas, não são exclusivas do PB. Como já foi dito na seção 2.1.3.1, os autores fundamentam seu posicionamento apresentando exemplos de variação verbal e nominal no PA e no PE.

Não obstante, a TLI certamente amplia a discussão em torno das alterações observadas em línguas que passam pelo contexto sociolinguístico do contato. Postulada como um contínuo de níveis, a TLI abarca tanto casos em que há a formação de uma nova língua (pidgin/crioula) como os casos em que uma língua-alvo apresenta diversas alterações devido ao seu contato constante com outros sistemas. Além disso, tal hipótese se vale tanto de evidências linguísticas quanto de dados históricos, demográficos e sociais de modo geral estreitamente relacionados à formação de uma língua.

De modo geral, cada proposta teórica tem apresentado dados e análises relevantes que expõem as particularidades do PB em relação ao PE e abrem caminho para que novas interpretações entrem nesse debate. Nesse ponto, vale lembrar o posicionamento de Mattos e Silva, que não entra no debate crioulação *vs.* deriva, mas afirma que características inovadoras do PB, sobretudo das variedades populares, se devem à forma como a língua portuguesa foi aprendida nos primeiros séculos pela massa populacional: “como segunda língua; com modelos defectivos da língua-alvo, a do colonizador, mas não tão defectivos que propiciassem a formação de um crioulo estável e generalizado; na oralidade; sem o controle normativo da escolarização” (MATTOS E SILVA, 2004, p. 107). A autora está de acordo, principalmente, com a TLI, pois defende a ideia do multilinguismo generalizado e da ação normatizadora no final do século XIX, aspectos sócio-históricos também abordados por Lucchesi.

Como apresentado no início deste capítulo, a hipótese norteadora desta pesquisa situa o fenômeno investigado, o uso do IV *mais* como conector coordenativo e subordinativo, como

um reflexo do contato entre línguas ocorrido no Brasil (cf. seção 2.1.1), por esse motivo, selecionou-se um *corpus* com amostras de fala vernácula do português rural afrodescendente.

Na seção seguinte, apresentam-se teorias de formação e aquisição de línguas em situação de contato linguístico, uma vez que se verifica a existência de fenômenos semelhantes aos investigados neste trabalho em pidgins e crioulos (cf. seção 1.3).

2.2 TEORIAS DE FORMAÇÃO DE LÍNGUAS EM SITUAÇÃO DE CONTATO

As teorias acerca da mudança linguística motivada pelo contato entre línguas têm estudado sistematicamente os pidgins e crioulos, variedades linguísticas totalmente novas e distintas das línguas envolvidas em sua formação que surgem em contextos sócio-históricos específicos. Nesta seção, são feitas, inicialmente, algumas considerações acerca dos contextos que proporcionam a pidginização e a crioulição, das características estruturais dos pidgins e crioulos, bem como das teorias acerca da gênese dessas línguas. Em seguida, são apresentadas duas teorias de formação e aquisição em situação de contato: Bioprograma da Linguagem (BICKERTON, 1988) e Relexificação (LEFEBVRE, 1998, 2001; LEFEBVRE E LUMSDEN, 1994).

Diferentemente das línguas que passam por mudança linguística regular, em linhas gerais, as línguas pidgins e crioulas podem ser caracterizadas da seguinte forma:

- (i) Desenvolvem-se em comunidades multilíngues;
- (ii) Surgem da necessidade de uma língua franca;
- (iii) São formadas em comunidades onde o acesso aos modelos da língua alvo é restrito;
- (iv) Surgem em um curto espaço de tempo;
- (v) Apresentam características das línguas do substrato e do superstrato.

(LEFEBVRE, 1998)

Quando grupos de falantes adultos de diferentes línguas são postos em contato e precisam interagir, a necessidade de comunicação leva esses falantes a desenvolverem uma forma de comunicação emergencial a partir de uma acomodação e negociação linguística. De acordo com Baxter (1996), nessas situações, “restrições sociais impedem o processo de aprendizagem de uma segunda língua e nenhum grupo aprende a língua do outro grupo”, por esse motivo, um código emergencial, jargão ou pré-pidgin, começa a se esboçar quando os

falantes desenvolvem uma forma de comunicação utilizando as regras gramaticais de suas línguas nativas combinadas com um pequeno vocabulário manejado por cada falante. Esse contexto sociolinguístico é caracterizado pelas condições em (i), já que se trata de uma pequena elite que fala a língua do superstrato e uma maioria que fala as diversas línguas do substrato, e em (ii) por causa da necessidade de interação entre esses grupos.

Esse jargão inicial com funções restritas e caracterizado por uma drástica redução gramatical pode ser aprimorado e em torno dele pode surgir alguma estruturação gramatical quando a situação de contato se mantém. Nesse caso, a língua que começa a se formar passa a denominar-se *pidgin*, que, por sua vez, pode difundir-se e estender-se lexicalmente e gramaticalmente, passando ao *status* de pidgin expandido (COUTO, 1996). Segundo Baxter (1996), na pidginização, há uma diminuição das complicações gramaticais e uma redução drástica do léxico. O que caracteriza a pidginização é o fato de a nova variedade de língua ser utilizada pelos falantes do substrato para a interação entre si e não somente como forma de comunicação com o dominador.

Mühlhäusler (1986) considera a pidginização como um processo de gramaticalização dentro de um *continuum*: jargão > pidgin instável > pidgin estabilizado > pidgin expandido. O pidgin passa por estágios a depender do grau de estruturação gramatical que se organiza em torno do jargão inicial, e vai desde o pidgin instável (com pouca estruturação e pouco vocabulário que diferenciam esse pidgin das línguas envolvidas na situação de contato) ao pidgin expandido.

Por seu turno, as línguas crioulas surgem quando há a nativização de um pidgin ou mesmo jargão socializado, ou seja, quando uma nova comunidade é formada a partir da situação de contato e uma nova geração adquire como língua materna a variedade da língua formada nessa comunidade, mesmo tendo acesso às línguas nativas dos pais como aconteceu nos contextos de escravidão (MÜHLHÄUSLER, 1986)³⁷. As línguas crioulas são L1 para a maioria dos seus falantes, diferentemente dos pidgins que não têm falantes nativos (COUTO, 1996).

Ao longo da história das línguas naturais, há diversos casos de pidginização, que ocorreram devido aos contatos promovidos por conta de deslocamentos com interesse comercial, de viagens de reconhecimento e da escravidão. Segundo Couto (1996) acredita-se que o primeiro pidgin de certa importância da história das línguas tenha sido uma variedade

³⁷ A afirmação de que a língua crioula surge com a nativação é proposta por grande parte dos crioulistas, mas não é consensual, a exemplo dos posicionamentos de Singler (1992) e Mufwene (1996).

do português utilizada para comunicação ao longo de toda a costa africana e em toda a rota marítima dos portugueses, desde os primeiros anos das Grandes Navegações.

Já as línguas crioulas surgiram principalmente no contexto da expansão europeia, durante os séculos XV a XIX, quando os europeus buscavam ampliar seu território através da colonização de novas terras e fortalecer sua economia por meio de novas relações comerciais. No século XV, a expansão começa com os portugueses na África, que vão se estabelecendo em diferentes lugares até a Ásia e Oceania, o que também é feito a partir do século XVII pelos ingleses, franceses e holandeses. Por esse motivo, as línguas europeias (português, inglês, francês, holandês e espanhol) são a base da formação da maioria das línguas crioulas. A atuação pioneira dos portugueses deu ensejo à formação de diversas línguas crioulas, chamadas de crioulos de base portuguesa, conforme Couto (1996).

A colonização europeia favoreceu o surgimento de línguas crioulas por conta de sua forma de atuação: a) a necessidade de trabalhadores escravos levou diversos contingentes de diferentes origens a serem transportados para um local designado para o trabalho, sendo forçados a conviver e, conseqüentemente, a comunicarem-se; b) o grupo dominador estabelecia entrepostos comerciais em locais onde já eram faladas diferentes línguas e os dominados ainda eram obrigados a aprender a língua do superstrato. Outro contexto sociolinguístico para a formação de línguas crioulas são os quilombos, agrupamentos de escravos foragidos, dada a pluralidade étnica de seus habitantes.

Esses contextos sociolinguísticos são definidos por Bickerton (1988) e tomados como parâmetro para distinguir os tipos de línguas crioulas. O primeiro caso, segundo o autor, dá origem aos *crioulos de plantação*, que surgiram principalmente nas situações de trabalho das indústrias açucareiras, nas quais se desenvolveram sociedades pequenas e rígidas constituídas de uma pequena elite europeia e uma maioria de trabalhadores não europeus. O segundo contexto, um grupo de fora que penetra em uma área onde muitas línguas já são faladas, dá origem ao *crioulo de forte*. As fortificações eram lugares de trabalho estabelecidos por europeus em território não europeu que davam origem a populações híbridas de pais europeus, mães não europeias e crianças etnicamente misturadas (BICKERTON, 1988). Os crioulos de plantação teriam ocorrido no Caribe, na costa ocidental da África e no Havaí; os *crioulos de quilombo*, formados nos agrupamentos de escravos foragidos, conhecidos atualmente encontram-se no Suriname, na ilha de São Tomé e na Colômbia.

Bickerton (1988) salienta que essas possibilidades não são as únicas. Conforme caracterizado em (iii), o contexto de surgimento de pidgins e crioulos se verifica quando um grupo multilinguístico tem a necessidade de uma língua em comum e o acesso a uma língua

preexistente é restringido. A característica em (iv) é apresentada, entre outros, por Goodman (1964, *apud* LEFEBVRE, 1998), que salienta a diferença entre as mudanças linguísticas ocorridas em ritmo lento e percebidas apenas em períodos consideráveis de tempo, e as mudanças linguísticas ocorridas em um curto espaço de tempo, duas ou três gerações, tal como a criouliização.

Finalmente, o que se percebe sobre (v) é que se trata de uma mistura de diferentes línguas, mas não de uma mistura aleatória. Com base nas observações de diversos estudiosos³⁸ sobre as línguas crioulas, Lefebvre (1998, p. 4) afirma que estas observações sugerem que as línguas não são formadas por uma mistura arbitrária das propriedades das línguas presentes no momento da formação dos crioulos. A autora conclui que:

O padrão geral que parece emergir das observações é o seguinte: enquanto as formas das entradas lexicais de um crioulo radical são derivadas da língua do superstrato, as propriedades sintáticas e semânticas dessas entradas lexicais seguem o padrão das línguas do substrato³⁹ (LEFEBVRE, 1998, p. 4) (tradução minha).

Para a definição de uma língua como crioula, devem ser levados em consideração tanto seus aspectos linguísticos como seu processo de formação, pois os contextos sociolinguísticos específicos teriam como resultado certos tipos de estruturas que seriam comuns a essas línguas. O processo histórico que dá origem a uma língua crioula envolve, por um lado, a perda de elementos que não são essenciais à comunicação imediata, do que decorrem as restrições funcionais das línguas emergenciais. Por outro lado, a ampliação das funções do código comunicativo emergencial gera uma necessidade de expansão/reestruturação para atender a essas novas funções (LUCCHESI, 2009).

Nessa perspectiva, Bickerton (1988) afirma que alguns elementos gramaticais perdidos por conta do contato não são reconstruídos; outros se perderam mas precisam ser reconstruídos. No primeiro grupo, encontram-se os elementos gramaticais sem valor referencial, que exprimem conceitos mais abstratos e formais, bem como os elementos de conteúdo semântico mais transparente:

³⁸ Goodman, (1964); Voorhoeve, (1973); Huttar, (1971), dentre outros que a autora cita.

³⁹ “*These observations suggest that creole languages are not formed by an arbitrary mixture of the properties of the languages present at the time they are being created. The general pattern that seems to emerge from the observations reported above is the following: while the forms of the lexical entries of a radical creole are derived from the superstratum language, the syntactic and semantic properties of these lexical entries follow the pattern of the substratum languages*” (trecho original).

- Concordância de gênero;
- Concordância de número;
- Morfologia verbal presa: tempo, particípio, etc.;
- Morfologia derivacional;
- Pronome de caso e formas de gênero;
- A maioria das preposições.

O segundo grupo constitui-se de elementos que são reconstruídos por meio de *gramaticalização*:

- Artigos;
- Formas de tempo, aspecto e modo;
- Pronomes interrogativos;
- Pluralizador;
- Pronomes para todas as pessoas e números;
- Formas para marcar casos oblíquos;
- Preposição locativa geral;
- Complementizador “irrealis”;
- Partícula de relativização;
- Reflexivos e recíprocos.

Os estudos sobre os pidgins e crioulos têm se preocupado em explicar a gênese dessas línguas e as propriedades em comum que compartilham entre si. Algumas correntes teóricas seguem em direções bem distintas, mas todas têm em comum o mesmo objetivo.

Muito em voga nos anos 60 e 70, a *hipótese monogenética*, em sua versão mais forte, considera que as línguas crioulas de base portuguesa originaram-se de uma fonte única, o Pidgin Português do Oeste Africano, que também deu origem a outros crioulos por meio de um processo de relexificação, uma hipótese substratista. Esse proto-pidgin serviu como instrumento de comunicação no período da expansão marítima em um contexto que Couto (1996) resume da seguinte forma:

Os portugueses foram os primeiros a ir além do cabo Bojador e do cabo Não, tendo disseminado diversos núcleos de colonização ao longo de toda a costa ocidental africana já a partir da metade do século XV. Hoje sabemos que uma língua de contato de base portuguesa – sob a forma de pidgin ou crioulo – estava em curso na região desde pelo menos os começos do século XVI. (...) os portugueses foram também os maiores fornecedores de escravos para as colônias das outras potências europeias. Tudo isso contribuiu para a disseminação de uma espécie de língua franca de base portuguesa pelos sete mares. É devido a tudo isso que surgiu a hipótese monogenética que poderia ser mais bem chamada de *lusogênese* (COUTO, 1996, p. 156).

Outra versão da hipótese monogenética acrescenta a ideia de esse proto-pidgin ser derivado por relexificação de uma língua franca do mediterrâneo que existiu por volta do ano 1000. Em uma perspectiva mais restrita, a monogênese é postulada para explicar a ocorrência de crioulos de base inglesa e francesa, assumindo que são derivados respectivamente de um Pidgin Inglês do Oeste Africano e de um Pidgin Francês do Oeste Africano. Porém, Besten, Muysken e Smith (1994, p. 89) afirmam que a principal questão levantada nesse caso é se os dois grupos de crioulos baseados no inglês (grupos do Atlântico e do Pacífico) e se os dois grupos de crioulos baseados no francês (grupos do Atlântico e do Oceano Índico) têm a mesma origem em comum. Segundo Besten, Muysken e Smith (1994, p. 88), a ideia de que todos os pidgins e crioulos, ou todos os crioulos da área atlântica tenham derivado de um único caso de pidginização é irracional.

Outras hipóteses, superstratistas, voltadas para um *input* europeu são a *hipótese de dialeção de línguas europeias*, segundo a qual as línguas crioulas seriam dialetos regionais das línguas europeias (inglês, francês e holandês), e as teorias do *baby talk* e do *Foreigner talk*, que levam em consideração a simplificação de estruturas linguísticas. A *teoria do baby talk* considera que processos de pidginização teriam se originado a partir da fala voluntariamente simplificada dos falantes do superstrato, semelhante à forma de comunicação que um adulto utiliza com uma criança. Essa fala teria servido como modelo para o substrato e explicaria a simplificação dos pidgins e crioulos. Na *teoria do Foreigner talk*, a pidginização é a cristalização de um modelo imperfeito da língua-alvo que resulta da tentativa de aquisição da língua pelos grupos dominados. Porém, estruturas sem uma correspondência com as línguas europeias, nos pidgins e crioulos, contestam as hipóteses do *baby talk* e *foreigner talk*.

Outra hipótese que objetiva explicar o surgimento das línguas que se formam em contextos especiais de contato entre línguas é a *teoria do substrato*. Segundo Couto (1996):

“por substrato se entende as influências dos povos dominados deixadas na língua de que se trata”. Nessa perspectiva, os falantes adultos dos grupos dominados são os principais agentes da formação de línguas pidgins e crioulas por transferirem as estruturas gramaticais de suas línguas nativas para a nova língua em formação durante o período de socialização. A combinação dessas estruturas com itens lexicais da língua alvo constitui a base para a formação da nova língua. Muitas teorias conhecidas por outros nomes também consideram o substrato fundamental para a formação dos pidgins e crioulos, a exemplo da *teoria da língua mista*.

Arends, Kouwenberg e Smith (1994, p. 103-108) apresentam evidências da influência do substrato nos níveis fonológico, morfológico, lexical, sintático e semântico de línguas crioulas. Claire Lefebvre, da Universidade de Quebec, em Montreal, tem estudado o papel do substrato na gênese das línguas crioulas a partir de um estudo sistemático que fez, em mais de 20 anos de pesquisa, do crioulo haitiano, que tem como língua lexificadora o francês e uma das principais línguas do substrato o fongbe, do oeste africano. Lefebvre (1998, 2001) e Lefebvre e Lumsden (1994) propõem alguns processos cognitivos que estariam envolvidos na gênese das línguas crioulas, os quais serão apresentados na seção 2.2.2.

Em oposição aos substratistas, estão os *universalistas*. De acordo com Derek Bickerton, em sua teoria do Bioprograma da Linguagem (1988), as crianças formam as línguas crioulas devido a sua capacidade inata para a aquisição da linguagem, como se apresenta na seção 2.1.1, a seguir.

No centro da discussão promovida pelos universalistas e substratistas está a questão da reestruturação do jargão, ou melhor, quem são os principais agentes dessa reestruturação. Para os universalistas, a reestruturação se dá na nativização, desse modo, as crianças seriam os principais agentes do processo por conta da atuação da faculdade da linguagem que opera na aquisição de L1. Para os substratistas, os adultos transferem as estruturas das suas línguas nativas para a língua em formação e a reestruturação é feita por meio de uma negociação (social) em torno das soluções gramaticais.

A seção seguinte apresenta a hipótese universalista de Bickerton (1988): o Bioprograma da Linguagem.

2.2.1 Bioprograma da Linguagem (BICKERTON, 1988)

A hipótese do Bioprograma da Linguagem fundamenta-se nos aspectos similares entre as diferentes línguas crioulas. Bickerton (1988) considera que esses aspectos não podem ser atribuídos a propriedades das línguas do substrato nem tampouco do superstrato. As propriedades gramaticais dos crioulos independem das línguas do substrato, pois resultam de aspectos universais da faculdade da linguagem, comum a todos os seres humanos.

Para o autor, a crioulição ocorre quando crianças são expostas a uma comunidade de fala que tem um pidgin não estruturado, um modelo defectivo de L2, como língua. As crianças aprendem o pidgin, o *input* para a aquisição de língua materna, e usam sua faculdade da linguagem (um conjunto de estruturas gramaticais específicas que elas podem acessar quando expostas a dados incompletos) para criar/inventar as estruturas que faltam. As crianças reestruturam a língua crioula por meio dos dispositivos inatos – o bioprograma. Desse modo, segundo o autor, os responsáveis pela criação de uma nova língua são as crianças. A faculdade da linguagem atua durante a aquisição de língua materna, formando uma língua altamente estruturada a partir de dados linguísticos restritos.

De acordo com a hipótese de Bickerton, o léxico deriva da língua dominante enquanto a gramática é formada pelas crianças. Essas adquirem o léxico da língua alvo já empobrecido pelo pidgin e formam a gramática por meio do bioprograma. Sendo a faculdade da linguagem universal, a gramática das línguas crioulas apresenta aspectos similares.

Muysken e Veenstra (1994, p. 129) explicam que, nessa perspectiva, as línguas crioulas são invenções das crianças que crescem na nova comunidade (que se formou a partir de um dos contextos sociolinguísticos apresentados acima) ouvindo, ao seu redor, apenas falantes do pidgin que se formou ali. Porém, trata-se de uma língua sem estrutura suficiente para funcionar como língua natural, então, as crianças usam sua própria capacidade linguística inata para transformar a entrada pidgin de seus pais em uma linguagem de pleno direito. Desse modo, as línguas crioulas são semelhantes porque a capacidade linguística inata aplicada é universal, e são mais simples porque refletem as estruturas linguísticas mais básicas.

Com sua teoria, Bickerton dá conta de características que são gerais das línguas crioulas e situa os estudos crioulos no campo dos universais linguísticos. Para o autor, o estudo das línguas crioulas pode ser válido para as teorias de aquisição, de um modo geral, pois revelam como os dispositivos inatos atuam em situações em que a transmissão geracional de dados bem formados foi interrompida. Nesses casos, é possível perceber, segundo

Bickerton, de que modo a faculdade da linguagem determina a aquisição, já que mostram situações de línguas que se formaram em um curto espaço de tempo.

A proposta de Bickerton tem recebido diversos questionamentos e críticas, sobretudo da parte dos substratistas, que atribuem aos falantes adultos em uma situação de contato o papel de formação da nova entidade linguística através da transferência de mecanismos gramaticais de suas línguas nativas para a nova língua pidgin ou crioula em formação. Na seção seguinte, apresenta-se a hipótese substratista da *relexificação*.

2.2.2 Relexificação (LEFEBVRE, 1998, 2001; LEFEBVRE E LUMSDEN, 1994)

A hipótese da relexificação tem sido formulada para explicar a formação de línguas pidgins, crioulas e mistas, que emergem em contextos de contato entre línguas. De acordo com Couto (2002, p. 224), considera-se Whinnom (1956), Taylor (1956) e Thompson (1961) como os iniciadores dessa teoria. Porém, foi com Stewart, em 1962, que o termo *relexificação* foi cunhado.

Muysken (1981a) formulou a hipótese da relexificação em termos mais formais. De acordo com o autor, a relexificação pode ser definida como um processo de substituição vocabular que adota, das entradas lexicais da língua alvo, apenas a representação fonológica. Na proposta de Muysken, para que ocorra a relexificação as representações semânticas da língua alvo e da língua do substrato precisam se sobrepor parcialmente, caso contrário, as duas entradas nunca poderão ser associadas uma a outra. Desse modo, a relexificação é dirigida semanticamente (MUYSKEN, 1981a, p. 62 *apud* Lefebvre 1998, p. 16).

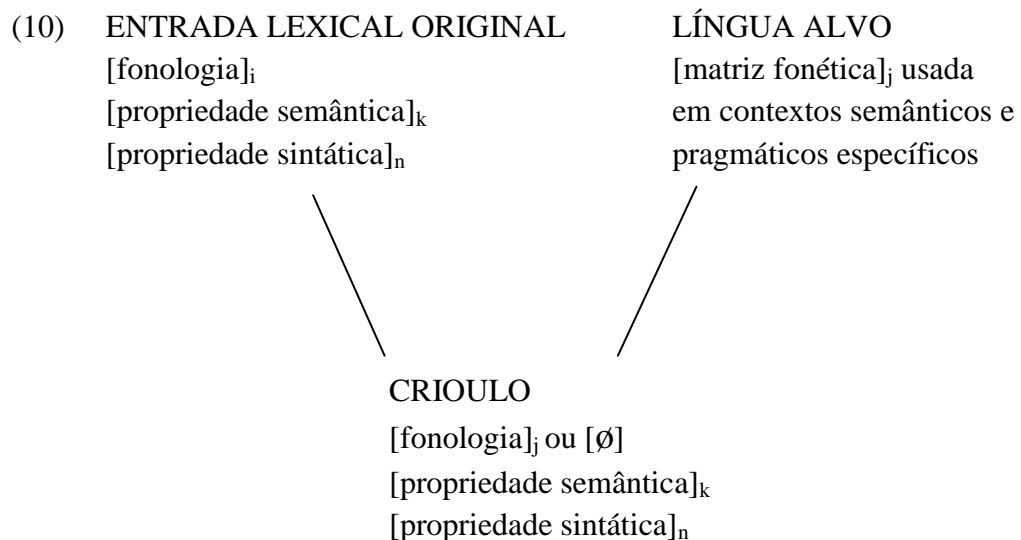
O autor fundamenta-se em dados da *media lengua*, uma língua mista falada no Equador que tem o quéchua como língua de substrato e o espanhol como língua de superstrato. No léxico da *media lengua*, as categorias lexicais têm praticamente todas as suas formas fonológicas derivadas do espanhol, enquanto os afixos e categorias funcionais são derivados do quéchua, como no exemplo apresentado em (9). Muysken (1981a *apud* Lefebvre 1998, p. 20) afirma que embora a representação fonológica seja do espanhol, o conteúdo semântico das categorias lexicais é da língua do substrato.

- (9) a. *No sé.* (Espanhol)
 não saber-1st
 “Eu não sei.”

b. *Mana yacha-ni-chu.* (Quechua)
 No *sabi-ni-chu.* (Media lengua)
 não *saber-1st-VAL*
 “Eu não sei.”

(MUYSKEN, 1981a, *apud* Lefebvre 1998, p. 19)

A fim de explicar o papel do substrato na gênese dos crioulos, Lefebvre (1998, 2001) e Lefebvre e Lumsden (1994) consideram que a relexificação é um processo mental que constrói um novo léxico da seguinte forma: as entradas lexicais das línguas do substrato são copiadas e a representação fonológica dessas entradas copiadas é substituída por uma nova forma derivada de uma matriz fonética da língua-alvo. Desse modo, uma palavra crioula possui a sintaxe e a semântica do seu equivalente substrato, mas é fonologicamente relacionada a um item lexical do superstrato. A escolha da matriz fonética da língua-alvo para relexificar uma entrada lexical copiada é baseada em seu uso em contextos semânticos e pragmáticos específicos. Lefebvre e Lumsden (1994) propõem a seguinte representação do processo de relexificação:



(LEFEBVRE, 2001, p. 12)

Na fase inicial de formação do crioulo, o processo de relexificação provê um léxico comum para que os falantes das línguas do substrato possam se comunicar. Para os falantes, criadores do léxico crioulo, a relexificação consiste em identificar, na língua-alvo, uma sequência fonética apropriada que sirva como uma representação fonológica para a entrada

lexical copiada de sua(s) língua(s) nativa(s). Lefebvre (2008. p. 199) apresenta o seguinte exemplo:

(11)	HAITIANO	FRANCÊS	FONGBE
	<i>vyann</i>	<i>viande</i>	<i>làn</i>
	‘carne’	‘carne’	‘carne’
	‘animais comestíveis’		‘animais comestíveis’
	(complemento do verbo ‘matar’)		(complemento do verbo ‘matar’)

No exemplo apresentado em (11), as entradas lexicais do francês e do fongbe, envolvidas na relexificação, compartilham o significado de *carne*. A nova entrada lexical do crioulo haitiano deriva sua fonologia do francês, mas suas propriedades sintáticas e semânticas são da língua do substrato.

De acordo com Lefebvre (1998, p. 40), as entradas lexicais crioulas são formadas, principalmente, pelo processo de relexificação, que objetiva a formação de um léxico comum que viabilize a comunicação; esse novo léxico torna-se a base para a língua franca dentro da comunidade crioula. Além disso, dois outros processos, alimentados pelo *output* da relexificação, também desempenham um papel no desenvolvimento do crioulo: a *reanálise* e o *nivelamento dialetal*.

A reanálise/gramaticalização é um processo possível nos casos regulares de mudança linguística e é um importante processo no desenvolvimento das línguas crioulas, na perspectiva de Lefebvre, já que um item relexificado pode ser gramaticalizado, tendo suas funções ampliadas.

Por conta das diferentes línguas que atuam como substrato de uma língua crioula, o léxico criado por relexificação, no início da formação do crioulo, não é uniforme. Desse modo, o nivelamento dialetal opera reduzindo a variação resultante da relexificação dos vários léxicos do substrato. De acordo com Couto:

O nivelamento dialetal, contrariamente aos processos mentais de relexificação e de reanálise, é um processo social, e consiste numa negociação entre os falantes de diversas línguas de substrato sobre qual forma adotar. Ele visa a reduzir a variação entre os léxicos produzidos pela relexificação de diferentes léxicos substratais. Os traços que são comuns a um grande número de línguas de substrato serão preferidos. (COUTO, 2002, p. 233)

Conforme postulado em (10), a relexificação não se distancia da afirmação consensual de que a gênese de pidgins e crioulos envolve mistura e redução seguida de expansão. Apesar de considerar o falante adulto como o principal agente da formação das línguas crioulas, Lefebvre (1998, p. 13) considera que, como em outras situações de mudança linguística, as crianças têm um papel importante no desenvolvimento de crioulos. Vale ressaltar ainda que, embora não discuta a questão dos fatores sociais na obra em questão, a autora assume que eles interagem com os processos envolvidos na formação dos crioulos, particularmente com o nivelamento dialetal, conforme estudos de Labov e seus associados sobre a interação de fatores linguísticos e sociais em comunidades linguísticas.

No que diz respeito ao fenômeno investigado nesta pesquisa, verifica-se, nas regiões de origem das línguas africanas que chegaram ao Brasil, a presença de línguas com fenômeno semelhante: um mesmo IV para as funções de coordenador de DPs e partícula comitativa (cf. seção 2.1.2). Esse cenário permite imaginar a presença de um substrato africano no Brasil, cujas línguas possuem estruturas semelhantes ao fenômeno variável *mais/com* e *mais/e* do português rural afrodescendente, apontando para a possibilidade de relexificação.

Considerando esses fatos sócio-históricos, bem como a proposta da relexificação de Lefebvre (1998, 2001) e Lefebvre e Lumsden (1994), é possível assumir que:

- i. Na formação do PB, o IV *mais* tenha sido a escolha fonética da LP, a língua-alvo, para relexificar uma entrada lexical copiada da(s) língua(s) do substrato, de maneira semelhante ao esquema representado em (10).
- ii. O contato da LP com línguas do substrato africano no Brasil tenha favorecido um processo de reanálise do IV *mais*, que tem suas funções ampliadas, passando a exercer a função de coordenador e subordinador no PB;
- iii. A nova entrada, o *mais* conector no PB, passa a ter propriedades semânticas e sintáticas da entrada original, ou seja, da língua do substrato, e representação fonológica da língua alvo.

Para Muysken (1981a), deve haver algo em comum entre a semântica da entrada lexical copiada e a semântica da matriz fonética escolhida para fazer a relexificação. A função de coordenar pressupõe mais de um elemento, assim como a expressão de comitativo; o IV *mais*, conforme apresentado na seção 1.2.3.1 possui o traço semântico de acréscimo. Os IVs *e*,

com e *mais* permitem a leitura de *adição de um segundo elemento*. Esse parece ser o traço semântico encontrado no IV *mais* que o licenciou para a relexificação.

O IV *mais* no português também possui mais de uma propriedade distributiva: pode ser modificador de verbo, de adjetivo, de advérbio e de nome. De certa forma é um IV multifuncional. Conforme apresentado na seção 1.2.3.1, do IV *mais* pode ser considerado um *predicador*, um termo que reclama a presença de outro, esse traço sintático encontrado no IV *mais* pode ter sido mais um traço em favor do seu licenciamento na relexificação.

Para esta pesquisa, assume-se que, no português rural afro-brasileiro falado no interior do estado da Bahia, o uso do IV *mais* como coordenador e subordinador resulta do contato da LP com as línguas do substrato africano, visto que estrutura semelhante ao uso variável *mais/com* e *mais/e* se verifica em línguas do continente africano. O substrato africano no Brasil que parece ter favorecido a relexificação do IV *mais* pode ter sido o quimbundo, o hauçá, o iorubá, dentre outras línguas. Como apresentado na seção 2.1.2, essas línguas chegaram à Bahia com os escravos africanos. Entretanto, admite-se que será necessário um trabalho mais amplo sobre a geografia linguística tanto no Brasil quanto na África, a fim de verificar com mais clareza a hipótese levantada.

Na seção seguinte, apresentam-se os pressupostos teóricos e metodológicos da Sociolinguística assumidos para este trabalho, bem como a descrição dos métodos que fundamentam empiricamente esta pesquisa.

2.3 A SOCIOLINGUÍSTICA

Rompendo com as propostas de estudo em que a língua é concebida como um sistema homogêneo e unitário, e, portanto, analisada fora de seu contexto social, a Sociolinguística tem como objeto de estudo a variação e a mudança linguísticas observáveis no contexto da comunidade de fala, concebendo a língua como um sistema heterogêneo e sistematicamente variável. A Teoria da Variação e Mudança, como também é chamada, parte dos postulados de Weinreich, Labov e Herzog (2006 [1968]) e dos estudos realizados por Labov (1994, 2008 [1972]). Como o próprio termo explica, a Sociolinguística objetiva descrever a língua a partir de seus condicionamentos sociais e linguísticos.

Labov (2008 [1972]), a partir de seu estudo pioneiro sobre a variação nos ditongos da língua inglesa falada na ilha de Martha's Vineyard, em Massachussets, propõe o estudo do uso vernacular da língua e das formas em variação dentro de uma comunidade de fala, a partir da perspectiva de que o processo de variação e mudança linguísticas só pode ser entendido a

partir do contexto social da comunidade, pois, além das pressões linguísticas, as pressões sociais também atuam sobre a linguagem. Para Labov, há uma regularidade dentro de um sistema variável, ou seja, a variação não é aleatória, mas regulada por condicionamentos linguísticos e sociais.

Weinreich, Labov e Herzog (2006 [1968]) postulam que nem toda variabilidade que se nota indica mudança linguística, porém, toda mudança pressupõe, obrigatoriamente, um estágio de variação na estrutura linguística. Nas subseções seguintes apresentam-se algumas considerações a respeito do estudo da variação e da mudança linguísticas.

2.3.1 O estudo da variação

A variação linguística dá-se no momento em que duas ou mais variantes coexistem em uma *comunidade de fala*. Esta, segundo Labov, é definida

pela participação num conjunto de normas compartilhadas; essas normas podem ser observadas em tipos de comportamento avaliativo explícito e pela uniformidade de padrões abstratos de variação que são invariantes no tocante a níveis particulares de uso. LABOV, (2008 [1972], p. 150).

A comunidade de fala não é determinada por acordos de uso linguístico, mas pelas avaliações que os falantes fazem do uso da língua na comunidade; seus membros compartilham normas estabelecidas a respeito de uma variedade linguística.

Da perspectiva de que a variação é inerente ao sistema linguístico (WEINREICH; LABOV; HERZOG, 2006 [1968]), surge a necessidade de explicar a escolha dos falantes por uma ou outra forma de uso. A partir do controle de condicionamentos sociais e linguísticos, a Sociolinguística busca descrever, explicar, sistematizar a variação no sistema linguístico. Esses condicionamentos indicam que a variação é regular e sistemática, e que a escolha de uma ou outra variante dentro do sistema não é aleatória.

Para descrever e explicar o processo de variação a partir do contexto social, uma pesquisa orientada pela Sociolinguística deve proceder à análise quantitativa de *variantes linguísticas* referentes à fala vernacular no âmbito de uma comunidade. Uma variante pode ter diferentes *status*: pode ser conservadora, inovadora, padrão ou não padrão, pode ter prestígio social ou ser estigmatizada. Um conjunto de variantes linguísticas, possibilidades de produção, é denominado de *variável*. Segundo Tarallo (2007), *variável dependente* é aquela utilizada como o objeto de estudo de determinada pesquisa sociolinguística, é o fenômeno

linguístico em variação. Delimitar a variável dependente é fazer o levantamento das possíveis variantes disponíveis para o falante. Nesta pesquisa, são variáveis dependentes a *coordenação entre DPs*, cujas variantes são *mais/e*, e a *subordinação em estruturas comitativas*, cujas variantes são *mais/com*.

Já as *variáveis independentes* são os contextos internos (linguísticas) e externos (sociais) que condicionam o uso das variantes, no sentido de “exercer pressão sobre os usos, aumentando ou diminuindo sua frequência de ocorrência” (MOLLICA, 2003, p. 11). As variáveis independentes ou explanatórias são estabelecidas mediante hipóteses, que devem ser formuladas considerando-se a estrutura social da comunidade de fala investigada e os contextos linguísticos. Por sua vez, os *fatores* são os valores de uma variável independente, por exemplo: para a variável social *faixa etária*, nesta pesquisa, os fatores são Faixa I (20 a 40 anos), Faixa II (41 a 60 anos) e Faixa III (acima de 60 anos); para a variável linguística *traço semântico do nome*, os fatores são [+ animado] e [- animado].

Para fundamentar empiricamente a pesquisa e a análise do fenômeno variável, o método sociolinguístico compreende etapas como: coletar os dados em uma determinada comunidade, constituindo um *corpus* com amostras de fala vernácula da comunidade a ser investigada; quantificar os dados coletados de acordo com as variáveis fixadas; descrever o fenômeno variável e suas possíveis variantes⁴⁰; delimitar os possíveis fatores linguísticos e sociais que condicionam o uso das variantes; bem como estabelecer a projeção do fenômeno variável no tempo a fim de identificar se o cenário variável indica ou não mudança linguística. Por se tratar de um modelo de análise quantitativa, é necessário que a investigação seja feita a partir de um *corpus* representativo, ou seja, de um volume de dados expressivo. Para estabelecer as variantes de um fenômeno variável é preciso investigar se há *covariação estrita*, ou seja, se as variantes podem permutar na mesma posição sem acarretar prejuízo semântico. O método sociolinguístico inclui ainda o suporte quantitativo que possibilita a análise estatística dos dados.

⁴⁰ No português popular do estado da Bahia verifica-se o uso de um mesmo IV com as funções de coordenador de DPs e partícula comitativa, semelhante ao que ocorre em línguas pidgins e crioulas, bem como em *WITH-languages*, na terminologia de Stassen (2000). Esse IV pode ser, no português popular do estado da Bahia, o IV *com* ou o IV *mais*. Desse modo, para a variável dependente *coordenação entre DPs* as variantes são *mais/e/com*. Nesta pesquisa, investiga-se apenas o uso do IV *mais* como coordenador e partícula comitativa, assumindo-se que esse IV tenha sido relexificado e reanalisado para exercer tais funções (cf. seção 2.2.2).

2.3.2 O estudo da mudança

A mudança linguística em uma comunidade de fala verifica-se quando uma das formas variáveis sobrepõe-se às demais e se estabelece na comunidade. É possível verificar um quadro de *mudança em progresso* a partir da observação que se faz de um fenômeno variável em *tempo real* e em *tempo aparente* (LABOV, 1994). O estudo em *tempo real* permite analisar um cenário de variação e verificar diacronicamente a possibilidade de mudança em progresso. Labov (1994) propõe duas maneiras possíveis de se fazer observações em tempo real: i) o *trend studies*, que consiste em retornar à comunidade investigada após certo período de tempo para repetir a mesma investigação com informantes distintos, mas que tenham as mesmas características sociais dos primeiros; e ii) o *panel studies*, em que o pesquisador volta à comunidade investigada para procurar os mesmos informantes que já participaram do primeiro estudo a fim de constituir um novo *corpus*. Em ambas é possível observar dois momentos da variação. A primeira abordagem permite estudar o comportamento linguístico da comunidade a partir da comparação de grupos diferentes; já a segunda favorece uma análise de comportamento linguístico individual em períodos distintos de tempo.

O *tempo aparente*, método utilizado nesta pesquisa, é uma projeção que se faz sobre o tempo real; é um procedimento metodológico que busca verificar, mediante a distribuição dos informantes por faixas etárias, um padrão ascendente ou descendente do fenômeno investigado. Um padrão que revela, por exemplo, que determinada variável realiza-se predominantemente entre os indivíduos mais velhos da comunidade, indica um quadro de mudança em progresso. Por outro lado, essa análise pode revelar um quadro de *variação estável*, quando não se verifica a predominância de uma variante linguística sobre outras.

A interpretação de dados que apontam para uma mudança em progresso ou variação estável não deve se fundamentar apenas na análise da faixa etária, mas deve ser motivada pelos resultados de outras variáveis sociais. Vale ressaltar ainda que o conceito de *mudança em progresso* proposto por Labov (1994) contraria a ideia de que a mudança linguística só pode ser observada após sua absoluta implementação, já que a variação que se percebe em um determinado tempo pode significar uma tendência à mudança futura.

Como se vê, a perspectiva Sociolinguística permite uma sistematização da diversidade linguística, a partir de uma metodologia bem delimitada capaz de descrever os condicionamentos linguísticos e sociais que operam a variação e a mudança no sistema linguístico.

Nas seções seguintes descrevem-se os métodos que fundamentam empiricamente esta pesquisa.

2.4 O MÉTODO

Nesta seção apresentam-se os métodos empregados para esta análise variacionista, de acordo com os princípios teóricos e metodológicos da Sociolinguística Quantitativa; descrevem-se as comunidades analisadas, o *corpus*, o tratamento dos dados, as variáveis dependentes, as variáveis independentes, e o suporte quantitativo – o pacote de programas VARBRUL. Os passos seguidos neste trabalho seguem o padrão utilizado nas pesquisas do Projeto Vertentes.

2.4.1 As comunidades analisadas

Assumindo que muitas mudanças ocorridas no PB, principalmente aquelas percebidas no português popular, foram motivadas pelo contato com falantes de diversas línguas no período de colonização do Brasil, principalmente com falantes de línguas africanas, conforme Lucchesi (1994, 2001, 2009), selecionaram-se quatro comunidades rurais compostas por afrodescendentes que vivem em relativo isolamento.

Segundo Lucchesi (2009, p. 75), as comunidades surgiram a partir de antigos quilombos ou de agrupamentos de ex-escravos que receberam doações de terra para viver após o fim da escravidão no Brasil. O autor afirma que essas comunidades se caracterizam pelos 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. (LUCCHESI, 2009, p. 75)

A escolha dessas comunidades se deve ao fato de que a variedade da LP falada nelas pode possuir aspectos semelhantes a línguas que sofreram processos de pidginização e criouliização, o que justifica realizar uma comparação com línguas crioulas, buscando identificar processos semelhantes aos investigados. Nesta pesquisa, são investigadas as seguintes comunidades rurais afrodescendentes do estado da Bahia: Helvécia, Cinzento, Rio de Contas e Sapé.

A comunidade de *Helvécia*, localizada no município de Nova Viçosa, no extremo sul do estado da Bahia, tem sua origem com a fundação da Colônia Leopoldina, por volta de 1818, por imigrantes europeus. Segundo Lucchesi (2009, p. 85) em 1850 a produção e a exportação do café na colônia, que utilizava mão-de-obra escrava, fizeram-na chegar ao auge; entretanto, no final do mesmo século diversos fatores corroboraram para sua decadência (entre eles, a abolição da escravatura) e consequente abandono por parte dos colonos europeus. Muitos ex-escravos não se dispersaram, mas permaneceram na região praticando a cultura de subsistência, mantendo-se em relativo isolamento até meados do século XX. De acordo com Baxter e Lucchesi (1999), até o final de 1850 o número de africanos chegava a 50% da população adulta da colônia.

Das comunidades investigadas, Helvécia é considerada a mais crioulezante no âmbito das pesquisas realizadas no Projeto Vertentes por guardar traços linguísticos que refletem um processo mais intenso de transmissão linguística irregular. Além disso, de acordo com Lucchesi (2009), essa comunidade possui características sócio-históricas que dão à comunidade esse *status*. São elas:

- a) A quantidade de africanos na colônia era muito maior do que a de brancos;
- b) Os senhores de escravos eram, em sua maioria, estrangeiros, portanto não tinham o português como língua materna;
- c) A comunidade de ex-escravos fixou-se em uma localidade isolada e permaneceu em relativo isolamento até meados do século XX;
- d) Muitas crianças escravas nasceram na Colônia Leopoldina, o que favorecia a manutenção das variedades do português utilizado para comunicação.

O fato de a dialetóloga Carlota Ferreira, numa expedição linguística ao sul do estado da Bahia, em 1961, ter registrado algumas características crioulezantes da fala de um casal de informantes, contribui também para a caracterização de Helvécia como crioulezante. O que se percebe na fala desses informantes são processos de variação “nos quais a forma padrão do português alterna com formas que teriam feito parte do repertório gramatical do antigo crioulo de Helvécia” (LUCCHESI, 2009, p. 93). São estruturas típicas de línguas crioulas encontradas em Helvécia pelos linguistas Alan Baxter e Dante Lucchesi em 1994:

- Ausência de preposição em estruturas nominais;
- Ausência do verbo copulativo;

- Orações encaixadas sem complementizador;
- Negação verbal com sujeito marcado negativamente;
- Uso de formas do presente para indicar ações e estados situados no passado;
- Estruturas de duplo objeto com inversão na ordem dos complementos verbais e supressão da posição de dativo.

A comunidade de *Cinzento*, localizada no município de Planalto, no semiárido baiano, é de origem quilombola, formada por escravos fugidos do “Arraial dos crioulos” na Chapada Diamantina. De acordo com Silva (2009), estima-se que a chegada dos primeiros negros a Cinzento tenha sido na primeira metade do século XIX. A comunidade encontra-se em um terreno de difícil acesso, acidentado, que favoreceu o isolamento dos cinzentenses de outras comunidades. Sobre a escolha de uma localidade com tantas adversidades, Silva (2009, p. 96) afirma que “certamente, fugiam da escravidão e buscavam no local a formação de uma irmandade com base primeiro na cor e depois no parentesco, já que a endogamia é uma prática comum até hoje”.

As comunidades de Barra e Bananal, conhecidas como *Arraiais de Rio de Contas*, encontram-se localizadas no município de Rio de Contas, na Chapada Diamantina. Segundo Lucchesi (2009, p. 97), a região de Rio de Contas começou a ser ocupada na última década do século XVII, por conta de uma rota de viagem que passava pela localidade em direção a Salvador, capital da Província na ocasião. A criação de um povoado chamado de Creoulos, “ponto de pouso” para os viajantes que seguiam a rota, deu início ao povoamento local, que foi crescendo assim como o desenvolvimento da mineração. As comunidades de Barra e Bananal foram habitadas inicialmente por escravos fugidos que procuravam um local seguro para habitar. Esses se fixaram às margens do rio Brumado praticando a cultura de subsistência, atividade realizada até os dias de hoje como o principal meio de sustento dos habitantes. Lucchesi (2009, p. 98) ressalta que em função da indústria do turismo instalada em Rio de Contas, muitos turistas e estudiosos têm frequentado as comunidades, cuja tradição africana foi se perdendo, a exemplo da religião. Atualmente, a religião católica predomina nas duas comunidades.

A comunidade de *Sapé*, localizada no município de Valença, na região do Recôncavo Baiano, tem sua origem após a abolição da escravatura. Figueiredo (2009) afirma que, em 1800, as terras da atual Sapé, bem como de localidades vizinhas, pertenciam a um único dono e foram repartidas como herança entre os filhos deste. Com o fim do trabalho escravo no país,

em 1888, muitos ex-escravos permanecem na região de Sapé e receberam de seus donos lotes de terra como pagamento por dívidas de trabalho. Mesmo sendo donos de terras pouco produtivas e isoladas, estes ex-escravos permaneceram ali e constituíram famílias. Atualmente, os habitantes da comunidade ainda praticam a cultura de subsistência e vivem em condições precárias de vida.

Sobre a formação sócio-histórica das comunidades investigadas, com base nos registros disponíveis, Lucchesi *et. al.*(2009, p. 100) fazem as seguintes considerações:

- a) **Helvécia**: comunidade que apresenta em sua gramática os efeitos mais notáveis do contato linguístico do PB;
- b) **Cinzento**: comunidade formada provavelmente no século XIX, período em que não havia mais condições para um contato linguístico mais profundo, semelhante a Sapé;
- c) **Rio de Contas**: as comunidades de origem mais remota e sobre as quais se poderia pensar em processos mais intensos de transmissão linguística irregular, porém, desenvolveram-se no contexto da mineração, não favorável a processos de crioulização;
- d) **Sapé**: comunidade formada provavelmente no século XIX, período em que não havia mais condições para um contato linguístico mais profundo.

2.4.2 O corpus

O *corpus* utilizado para esta pesquisa é composto de 48 entrevistas com amostras representativas da fala vernácula das comunidades afrodescendentes de Helvécia, Rio de Contas, Cinzento e Sapé. A fala dos informantes das três primeiras comunidades foi gravada em fita cassete (meio analógico) e depois digitalizadas; em Sapé a gravação já foi feita em meio digital. Todas as entrevistas foram editadas em meio digital para melhorar a qualidade do som, transcritas grafematicamente de acordo com a *Chave de transcrição*⁴¹ do Projeto Vertentes, e revisadas para correção de possíveis erros de transcrição.

Conduzidas pelo documentador, as entrevistas têm duração de 40' a 1 hora. Com o objetivo de registrar uma fala o mais natural e espontânea possível, buscou-se superar o

⁴¹ A Chave de transcrição foi elaborada pelo coordenador do Projeto Vertentes com o objetivo de orientar as transcrições das entrevistas, tanto no que diz respeito aos aspectos linguísticos a serem considerados como à uniformidade nas transcrições. A Chave encontra-se disponível para *download* em <<http://www.vertentes.ufba.br/projeto/transcricao>>.

*paradoxo do observador*⁴² minimizando a presença do documentador e o contexto de monitoramento através de estratégias como: um ambiente de descontração e conversação informal; temas de total domínio do entrevistado, como vida pessoal, conjugal, infância, doenças, morte de familiares, festas, trabalho, história da comunidade, dentre outros; liberdade para alterar o rumo da conversa; possibilidade de interação com terceiros.

A seleção dos informantes foi feita de acordo com critérios pré-estabelecidos. Buscou-se compor uma amostra estratificada com indivíduos que nasceram e viveram na comunidade, distribuídos da seguinte forma:

- Quanto ao sexo: masculino (M) e feminino (F);
- Quanto à faixa etária: Faixa I (20-40 anos), Faixa II (40-60 anos) e Faixa III (mais de 60 anos);
- Quanto ao nível de escolaridade: analfabeto (A) e semianalfabeto (S);
- Quanto à estada fora da comunidade: nunca se ausentaram (N) e saíram por período superior a seis meses (E).

Os Quadros a seguir mostram as características do *corpus* de acordo com as comunidades.

Faixa I					Faixa II					Faixa III				
Inq-02	M	28a	S	E	Inq-07	F	42a	A	E	Inq-13	F	85a	A	N
Inq-03	F	35a	S	E	Inq-08	M	44a	S	N	Inq-19	F	103a	A	E
Inq-04	M	30a	S	N	Inq-09	F	54a	A	N	Inq-20	M	70a	A	E
Inq-05	F	38a	S	E	Inq-12	M	57a	A	N	Inq-22	M	80a	A	N

Quadro 4: *Corpus* base do português rural afro-brasileiro de Helvécia

Faixa I					Faixa II					Faixa III				
Inq-01	F	26a	S	E	Inq-08	M	55a	A	E	Inq-20	M	77a	A	N
Inq-04	F	37a	S	E	Inq-09	M	43a	S	E	Inq-21	F	74a	A	N
Inq-05	M	26a	S	E	Inq-11	F	42a	S	E	Inq-24	F	75a	A	N
Inq-06	M	36a	S	E	Inq-13	F	47a	S	N	Inq-26	M	68a	S	E

Quadro 5: *Corpus* base do português rural afro-brasileiro de Rio de Contas

⁴² De acordo com Labov (2008 [1972], p. 245), trata-se de um problema metodológico em que a simples presença do documentador pode interferir na fala natural do informante: “o objetivo da pesquisa linguística na comunidade deve ser descobrir como as pessoas falam quando não estão sendo sistematicamente observadas – no entanto, só podemos obter tais dados por meio da observação sistemática”.

Faixa I					Faixa II					Faixa III				
Inq-01	F	24a	S	N	Inq-05	F	53a	S	E	Inq-09	F	76a	A	E
Inq-02	M	27a	S	N	Inq-06	M	42a	S	N	Inq-10	M	70a	A	N
Inq-03	F	36a	A	N	Inq-07	F	55a	A	N	Inq-11	F	77a	A	N
Inq-04	M	28a	A	N	Inq-08	M	48a	S	E	Inq-12	M	66a	A	N

Quadro 6: *Corpus* base do português rural afro-brasileiro de Sapé

Faixa I					Faixa II					Faixa III				
Inq-01	F	28a	S	E	Inq-05	M	41a	S	N	Inq-09	F	63a	A	N
Inq-02	F	16a	S	N	Inq-06	M	48a	S	E	Inq-10	F	107a	A	N
Inq-03	M	34a	S	E	Inq-07	F	44a	S	N	Inq-11	M	64a	S	N
Inq-04	M	41a	S	E	Inq-08	F	50 ^a	A	N	Inq-12	M	82a	A	N

Quadro 7: *Corpus* base do português rural afro-brasileiro de Cinzento

Como se pode observar nos Quadros, buscou-se estabelecer uma amostra vernacular em que o perfil dos informantes refletisse a heterogeneidade de cada comunidade. Foram selecionados 2 informantes do sexo masculino e 2 do sexo feminino por faixa etária, compondo 12 entrevistas por comunidade. Entretanto, a distribuição das variáveis apresenta algumas assimetrias em relação à escolaridade e à estada fora da comunidade, interpretadas por Lucchesi (2009, p. 157) como “reflexo do universo sócio-cultural em que foi feita a recolha” dos dados.

2.4.3 O tratamento dos dados

Os fenômenos investigados nesta pesquisa, *mais/com* em construções comitativas e *mais/e* coordenando DPs, demandaram a fixação de duas variáveis dependentes, como apresentado na seção seguinte. Após a seleção das comunidades e das entrevistas que constituem o *corpus* desta pesquisa, foi feito um levantamento exaustivo e cuidadoso das ocorrências das variáveis dependentes. Em seguida, foram estabelecidas as variáveis linguísticas que possivelmente atuariam influenciando as variáveis dependentes; o padrão de fixação das variáveis sociais seguiu os procedimentos comumente utilizados nos trabalhos realizados no âmbito do Projeto Vertentes, um padrão fixado pelo coordenador do projeto Prof. Dr. Dante Lucchesi.

As etapas seguintes foram: a elaboração de duas chaves de codificação, sendo uma para cada variável dependente; a codificação dos dados inicialmente levantados; a revisão e o

descarte das ocorrências que não atendiam a alguns critérios de seleção⁴³. Finalmente, procedeu-se à quantificação dos dados selecionados com o suporte do Pacote de programas VARBRUL (cf. seção 4.5), a fim de estabelecer as frequências de uso e, sobretudo, os grupos de fatores estatisticamente relevantes, que revelam os fatores significativos para a análise dos fenômenos.

Após a quantificação, os dados foram organizados em tabelas, para melhor visualização dos resultados estatísticos obtidos e procedeu-se à análise qualitativa desses resultados.

2.4.3.1 As variáveis dependentes

A motivação para a realização desta pesquisa foi a variação observada entre os IVs *e/com/mais*. A sistematização para o estudo do fenômeno impôs a fixação de duas variáveis dependentes, ou seja, do fenômeno linguístico em variação, objeto de estudo da pesquisa sociolinguística, visto que essa variação implica a existência de duas estruturas sintáticas distintas: *coordenação* e *subordinação*, conforme exemplos em (12) e (13), respectivamente.

- (12) a. “eu *mais* ele ia, pegava o ônibus” (HV-12)
 b. “Meu irmon *e* io nasceu” (HV-19)

A variável dependente *coordenação entre DPs* busca investigar o uso dos conectores coordenativos de DPs no português rural afrodescendente. Trata-se de uma variável binária e composta pelas variantes *mais/e*.

Conforme apresentado na seção 1.2.1, a coordenação constrói unidades complexas a partir da combinação de constituintes do mesmo nível categorial, semântica e sintaticamente independentes/autônomos.

- (13) a. “eu não queria brigá *com* ele não...” (SP-02)
 b. “Aí eu saía, viajava *com* ele, ia direto.” (CZ-07)
 c. “Denestina (...) é casado *mais* Nezim...” (HV22)
 d. “Cantava *mais* os ôto (...) eu sozinha num cantava não.” (CZ10)

⁴³ Por questões de coerência e organização textual, optou-se por apresentar esses critérios de seleção no capítulo 3, juntamente com as hipóteses norteadoras para as variáveis linguísticas. Como esta pesquisa investiga duas variáveis dependentes, as variáveis independentes linguísticas fixadas apresentam especificidades, desse modo serão apresentadas no capítulo 3.

A variável dependente *subordinação em construções comitativas* busca investigar o uso dos conectores em estruturas comitativas no português rural afrodescendente. Esta variável também é binária, composta pelas variantes *mais/com*. Para a análise desta variável, não se faz distinção entre a preposição funcional, como em (13a) e (13c), e a preposição lexical, conforme exemplos (13b) e (13d).

Retomando o que foi dito na seção 1.2.2, a subordinação pressupõe uma relação de dependência e complementação entre os constituintes envolvidos, ou seja, uma unidade subordinada está contida em uma unidade maior dentro de uma hierarquia gramatical no seguinte esquema estrutural: uma serve de base e outra serve de complemento ou adjunto.

2.4.3.2 As variáveis independentes

Na perspectiva de que a variação não é aleatória, mas condicionada por fatores linguísticos e sociais, foram elaboradas duas chaves de codificação com os grupos de fatores, ou seja, as variáveis independentes linguísticas e sociais que poderiam condicionar o uso das variantes. De acordo com Naro e Scherre (2007, p. 148):

[...] é fundamental identificar conjuntos de circunstâncias linguísticas e sociais (restrições) que tendem a favorecer ou desfavorecer o uso de uma ou outra variante [...]. Os grupos de fatores são uma forma de operacionalizar hipóteses a respeito do funcionamento dos fenômenos linguísticos variáveis, que podem ou não estar ligadas a modelos linguísticos claramente estabelecidos. (SCHERRE; NARO, 2007, p.148)

No que tange às variáveis linguísticas, para a variável dependente *coordenação entre DPs* foram estabelecidos 6 grupos de fatores, como se vê em (14), e para a variável *subordinação em construções comitativas* também foram fixados 6 grupos listados em (15). Tais variáveis linguísticas serão apresentadas no capítulo seguinte, juntamente com os resultados e análise dos dados, por apresentarem especificidades em relação às respectivas variáveis dependentes.

- (14)
- a. Estatuto sintático dos DPs coordenados
 - b. Tipo de categoria coordenada
 - c. Traço semântico dos nomes coordenados
 - d. Estrutura do primeiro DP coordenado
 - e. Estrutura do segundo DP coordenado

f. Referencialidade dos DPs coordenados

- (15) a. Estatuto sintático do PP encabeçado por *com/mais*
 b. Estrutura do DP encabeçado por *com/mais*
 c. Referencialidade do DP encabeçado por *com/mais*
 d. Traço semântico do nome
 e. Distância entre o verbo e o PP subordinado
 f. Função semântica do PP encabeçado por *com/mais*

Quanto às variáveis sociais estabelecidas, foram fixadas as mesmas estabelecidas para os trabalhos realizados no âmbito do Projeto Vertentes⁴⁴, que, dentre outros, busca verificar os efeitos históricos do contato da LP com as línguas africanas. A investigação dessas variáveis é bastante relevante para a investigação da hipótese defendida neste trabalho, a de que o uso do *mais* como coordenador de DPs e como partícula comitativa resulta do contato linguístico da LP com línguas do substrato africano (cf. seção 2.2.2). A análise de tais variáveis pode fornecer evidências empíricas que relacionam a atual realidade linguística do PB ao contato entre línguas que se deu nos três séculos de colonização do Brasil (cf. seções 2.1.1 e 2.1.2). Para a análise das duas variáveis dependentes foram fixadas as seguintes variáveis sociais:

- (16) a. Faixa etária
 b. Estada fora da comunidade
 c. Gênero
 d. Comunidade

Ao se fixar a variável *faixa etária* buscou-se verificar diferentes momentos dos fenômenos variáveis *mais/com* e *mais/e*, considerando o modelo laboviano de estudo linguístico em *tempo aparente*, que permite verificar se o fenômeno variável reflete uma variação estável ou uma mudança em progresso (LABOV, 2008 [1972]). Foram considerados para análise os seguintes intervalos etários: Faixa I (20 a 40 anos), Faixa II (41 a 60 anos) e Faixa III (acima de 60 anos). Lucchesi (2009, p. 156-157) afirma que a fala dos indivíduos mais velhos “permite visualizar os estágios mais antigos da gramática das comunidades de fala analisadas”. Nessa perspectiva e considerando a hipótese de que o conector *mais* emergiu

⁴⁴ Nesta pesquisa, entretanto, não se considerou a variável *escolaridade*, tendo em vista sua constituição por apenas dois fatores (analfabeto e semi-analfabeto).

do contato entre línguas, espera-se que a variante *mais* seja mais frequente na fala dos adultos e idosos.

Com a variável *estada fora da comunidade*, buscou-se verificar se o afastamento da comunidade, de pelo menos seis meses, condiciona o uso de uma ou de outra variante, tendo em vista que os indivíduos que nunca deixaram a sua comunidade para trabalhar fora, ou mesmo para passear ou visitar familiares, tendem a manter os padrões linguísticos próprios de seu dialeto. Diferentemente dos indivíduos que, por algum dos motivos supracitados ou mesmo por outros, se ausentaram de sua comunidade e tiveram contato com outras normas linguísticas. A expectativa é que os informantes que nunca saíram de sua comunidade ou não se ausentaram dela por mais de seis meses façam maior uso da variante conservadora⁴⁵ *mais*, tendo em vista que tais indivíduos tendem a manter os padrões linguísticos próprios de seu dialeto.

No que se refere à variável *gênero*, estudos linguísticos têm mostrado que sua relevância está atrelada às particularidades que envolvem cada comunidade analisada. De acordo com Paiva (2003, p. 33):

A análise da dimensão social da variação e da mudança linguística não pode ignorar (...) que a maior ou menor ocorrência de certas variantes, principalmente daquelas que envolvem o binômio forma padrão/forma não padrão e o processo de implementação de mudanças estejam associados ao gênero/sexo do falante e à forma de construção social dos papéis feminino e masculino (grifo nosso).

A autora cita diversos estudos de natureza sociolinguística que apontam para um conservadorismo linguístico das mulheres, como o de Fisher (1958), o de Mollica, Paiva & Pinto (1989) e o de Scherre (1996), dentre outros. Por outro lado, a autora ressalta que há estudos que apontam para a direção oposta, ou seja, que a variante inovadora é encontrada mais frequentemente entre as mulheres, como mostra um estudo feito por Labov (1966) sobre o inglês de Nova York. Assim, é possível perceber que os papéis que os homens e as mulheres exercem em suas comunidades definem seu comportamento linguístico, ora conservador, ora inovador.

No contexto cultural das comunidades rurais afro-brasileiras analisadas, os homens tendem a ter mais contato com o mundo exterior por conta de sua participação no processo produtivo e vida social, sendo expostos a diferentes padrões linguísticos. Por outro lado, as

⁴⁵ Em conformidade com a hipótese desta pesquisa, no âmbito das comunidades rurais afrodescendentes considera-se a variante *mais* conservadora em ambos os fenômenos investigados, enquanto as variantes padrão *e/com* são consideradas inovadoras.

mulheres apresentam comportamento linguístico mais conservador por conta de sua vida social mais voltada para o âmbito doméstico, resguardando formas linguísticas mais marcadas pelo contato entre línguas que fazem parte do padrão dessas comunidades (LUCCHESI, 2009, p. 313). Nessa perspectiva, espera-se que a variante *mais* seja favorecida na fala das mulheres.

A variável *comunidade* foi fixada para verificar a relação entre o grau de isolamento das comunidades selecionadas e o uso variável *mais/com* e *mais/e*. Conforme apresentado na seção 2.4.1, as quatro comunidades, compostas por afrodescendentes e formadas a partir de agrupamentos de ex-escravos ou antigos quilombos, mantiveram-se isoladas até meados do século XX, o que lhes permitiu resguardar aspectos linguísticos típicos de situações de contato entre línguas. A expectativa para o uso da variante *mais*, nos dois contextos sintáticos analisados, é que seja favorecida entre os falantes das comunidades com maior grau de isolamento, visto que são estas que tendem a conservar seus padrões linguísticos diretamente relacionados ao contato entre línguas (LUCCHESI, 2009). Com a fixação desta variável parte-se do pressuposto de que a variante *mais* seja favorecida entre os falantes das comunidades com maior grau de isolamento, que no âmbito do Projeto Vertentes são as comunidades de Helvécia e Cinzento, seguidas de Rio de Contas e Sapé.

2.4.4 O Varbrul

Para o tratamento estatístico dos dados foi utilizado o Pacote de Programas VARBRUL, um conjunto de programas computacionais estruturado para realizar análises quantitativas em dados de variação sociolinguística (GUY; ZILLES, 2007, p. 105). Por meio dos programas é possível obter os seguintes resultados estatísticos fundamentais para a análise dos fatores condicionantes das variáveis dependentes:

- Frequências percentuais;
- Variáveis/contextos relevantes;
- Pesos relativos em uma escala de 0 a 99;
- Nível de significância dos fatores relevantes.

O peso relativo acima de .50 indica um contexto favorável à realização de determinada variante, enquanto peso relativo abaixo de .50 indica desfavorecimento. Variantes com valor próximo a .50 são consideradas contextos neutros. Já o nível de significância equivale a uma

margem de erro; são considerados confiáveis os dados estatísticos com nível de significância igual ou inferior a .05.

Os programas que compõem o VARBRUL são executados um a um de modo que o *output* de um programa é o *input* para o programa seguinte. Para a análise de uma regra variável são necessários, basicamente, quatro programas: CHECKTOK, READTOK, MAKECELL e IVARB. Outros programas também fazem parte do Pacote VARBRUL. Para esta pesquisa foi utilizado ainda o TSORT.

Finalmente, sobre a relevância da análise quantitativa, Guy e Zilles (2007, p. 73) afirmam que:

A realização de análises quantitativas possibilita o estudo da variação lingüística, permitindo ao pesquisador apreender sua sistematicidade, seu encaixamento lingüístico e social e sua eventual relação com a mudança lingüística. [...] O uso de métodos estatísticos [...] tem permitido demonstrar o quão central a variação pode ser para o entendimento de questões como identidade, solidariedade ao grupo local, comunidade de fala, prestígio e estigma, entre tantas outras. (GUY; ZILLES, 2007, p.73)

Cabe ao pesquisador utilizar-se do método quantitativo, tão relevante, para realizar uma análise qualitativa dos dados, interpretando criticamente os resultados obtidos.

Após a coleta de dados e a codificação, submeteram-se os dados ao pacote de programas e chegou-se aos seguintes números:

Tabela 1: Total de ocorrências do fenômeno variável *mais/com*

	MAIS		COM		TOTAL
	Ocorr.	%	Ocorr.	%	
TOTAL	188	27	514	73	702

Tabela 2: Total de ocorrências do fenômeno variável *mais/e*

	MAIS		E		TOTAL
	Ocorr.	%	Ocorr.	%	
TOTAL	95	8	1163	92	1258

Após análise quantitativa das variáveis independentes, chegou-se a comportamentos categóricos, o que levou à redução dos dados analisados, como apresentado no capítulo seguinte.

2.5 CONSIDERAÇÕES PARCIAIS

O capítulo 2 teve por objetivo apresentar os pressupostos teóricos e metodológicos que orientam a investigação e realização desta pesquisa, bem como descrever os métodos empregados para esta análise variacionista. Para a análise dos fenômenos variáveis *mais/com* e *mais/e* investigados, é possível destacar que:

- Serão levados em conta tanto aspectos linguísticos quanto fatores sócio-históricos, como a história sociolinguística do Brasil, marcada por contato linguístico;
- Existem línguas com o valor identidade de formas no continente africano em áreas que coincidem geograficamente com as regiões de origem dos africanos que vieram para o Brasil durante o período do tráfico negreiro. As rotas do tráfico compreendem as regiões oeste-africana e banta, localizadas na África subsaariana;
- Chegaram à Bahia, com os africanos, línguas como quimbundo, hauçá, iorubá, wolof, ewe, fula, dentre outras, em que se verifica identidade de formas para o coordenador de DPs e partícula comitativa.

Neste capítulo, alguns posicionamentos foram assumidos quanto ao fenômeno que se investiga:

- O IV *mais* parece ter sido a escolha fonética da LP para o substrato africano relexificar uma entrada lexical copiada de sua(s) língua(s) nativa(s), um IV com as funções de coordenador entre DPs e partícula comitativa, semelhante ao esquema representado em (10), possibilitando a criação de um IV comum para a comunicação. Os IVs *e*, *com* e *mais* permitem a leitura de *adição de um segundo elemento*. Esse parece ser o traço semântico encontrado no IV *mais* que permitiu sua inserção na relexificação.
- A ampliação das funções do IV *mais*, que passa a atuar como coordenador de DPs e partícula comitativa, parece resultar ainda de uma reanálise/gramaticalização, em algum grau.

No capítulo seguinte, apresenta-se a análise quantitativa e qualitativa dos dados. Foram levantadas 702 ocorrências do fenômeno variável *mais/com* e 1258 ocorrências do fenômeno variável *mais/e* nas quatro comunidades rurais afrodescendentes investigadas: Helvécia, Rio de Contas, Cinzento e Sapé. A análise de amostras de fala vernácula de comunidades rurais afrodescendentes justifica-se para verificar a hipótese de que a variação dos IVs *e/com/mais* resulta de contato linguístico, assim, consideram-se bastantes relevantes os resultados referentes às variáveis sociais.

3 A ANÁLISE DOS DADOS

Neste capítulo, apresenta-se a análise quantitativa e a qualitativa dos dados tendo em vista os números obtidos após a submissão destes ao Pacote de programas VARBRUL. Este capítulo encontra-se organizado em duas seções, referentes a cada variável dependente investigada (cf. seção 2.4.3.1). Apresentam-se os critérios de seleção e refinamento dos dados, a motivação e as hipóteses que nortearam a fixação das variáveis linguísticas independentes, bem como as tabelas com a análise dos dados referentes às variáveis selecionadas pelo programa como relevantes.

Considerando a hipótese de que o uso do IV *mais* como coordenador resulta do contato entre a LP e as línguas do substrato africano no Brasil, no período da colonização do país (cf. seções 2.1.1 e 2.1.2), apresentam-se, primeiramente, os resultados das variáveis sociais, de acordo com a ordem de relevância estatística estabelecida pelo VARBRUL e, em seguida, os resultados das variáveis linguísticas selecionadas também por ordem de relevância. Consideram-se bastantes relevantes os resultados referentes às variáveis sociais, embora nem todas as variáveis fixadas tenham sido consideradas estatisticamente significativas.

3.1 ESTRATÉGIAS DE COORDENAÇÃO ENTRE DPs: MAIS/E

Esta seção trata da variação *mais/e* como estratégias de coordenação entre DPs, conforme exemplo em (1), fenômeno comum no português popular do estado da Bahia.

- (1) João *e/mais* Maria são namorados.

Para a análise desse fenômeno variável, constituiu-se um *corpus* com amostras de fala de quatro comunidades rurais afrodescendentes (cf. seções 2.4.1 e 2.4.2), de onde foram coletadas ocorrências dos coordenadores *mais/e*. Assume-se que coordenação é um processo de construção de unidades a partir da combinação de constituintes que possuem o mesmo nível categorial e são semântica e sintaticamente independentes/autônomos (cf. seção 1.1.1).

Inicialmente, fez-se um levantamento de todas as ocorrências em que havia coordenação com *mais/e*. Dessa forma, foram coletados dados de coordenação entre DPs, PPs e CPs, embora a expectativa fosse a de que essa variação ocorresse apenas entre DPs, visto que fenômeno semelhante, a possibilidade de um mesmo IV exercer a função de coordenador

de DPs e subordinador em construções comitativas pode ser observado em línguas que se formaram em situação de contato (cf. seções 1.3.1 e 1.3.2). No Quadro 2, retomado como Quadro 8, a seguir, observa-se que o coordenador de frases e orações em pidgins e crioulos, principalmente nos do Atlântico, não é o mesmo em construções de coordenação entre DPs; em contrapartida, o coordenador de DPs é o mesmo em construções comitativas. Esse padrão, o uso de uma mesma partícula para introduzir comitativo e para conectar DPs, verifica-se em diversas línguas que vieram para o Brasil com os africanos, como o iorubá e o hauçá, conforme Parkvall (2012, p. 141), e o quimbundo, de acordo com Bonvini (2008, p. 109 e 116).

	Adposição comitativa	Morfema usado para ligar sintagmas nominais	Morfema usado para ligar sintagmas verbais ou frases
Línguas europeias	COM	E	
Línguas nigero-congolesas	COM		E

Quadro 8: Comitativo e coordenadores em línguas europeias e nigero-congolesas

Fonte: PARKVALL (2012, p. 140)

Nessa perspectiva, as ocorrências foram codificadas segundo o *tipo de categoria coordenada*: DPs, PPs e CPs. Foram codificadas também ocorrências de coordenação entre PP + DP (dois PPs coordenados com a preposição não realizada no segundo) e entre DP + PP. A Tabela 3, referente à primeira quantificação realizada, apresenta a distribuição das ocorrências.

Tabela 3: Distribuição das variantes *mais/e* segundo o *tipo de categoria coordenada*

	MAIS		E		TOTAL
	Ocorr.	%	Ocorr.	%	
DP + DP	92	26	266	74	358
PP + PP	1	4	25	96	26
PP + DP	1	20	4	80	5
DP + PP	1	100	0	0	1
CP	0	0	868	100	868
TOTAL	95	8	1163	92	1258

Os primeiros resultados quantitativos, como se vê na Tabela 3, confirmam a hipótese inicial de que a coordenação entre DPs, como no exemplo em (2), é contexto propício à variação dos conectores *mais/e* com frequência igual a 26% e 74%, respectivamente.

- (2) a. “hoje, é só eu *e* o marido” (RC-11)
 b. “Minha mãe *mais* meu pai já plantô, eu nunca prantei não” (RC-04)

Os resultados apontam para o uso categórico da variante *e* na coordenação entre CPs, conforme exemplo em (3a). No caso de coordenação entre PPs, apenas uma ocorrência com a variante *mais*, (3b), foi encontrada. Trata-se de coordenação entre genitivos. Em (3c), apresenta-se um exemplo do uso da variante *e* coordenando PPs.

- (3) a. “Ali que eu pego o ônibus *e* salto, quando venho da feira...” (HV-05)
 b. “Quando chegô de manhã cedo, passô aquele fio *de... de Maria (...)* *mais de Armando*” (HV-22)
 c. “A mulhé sai de Itaperoá *e* de Valença pa dá estudo” (SP-11)

Durante a seleção dos dados, verificou-se ainda a coordenação entre PP + (\emptyset_{PREP}) DP com o IV *mais* na ocorrência em (4a) e com o IV *e* nas ocorrências de (4b) a (4e). Em (4a), embora ocorra coordenação entre categorias diferentes, verifica-se coordenação entre genitivos. Mira Mateus *et. al.* (2003, p. 385) afirmam que, em alguns casos, a preposição *de* é mais uma marca de genitivo do que uma verdadeira preposição. Como não se objetivou discutir: i) as características formais dos termos coordenados e ii) a baixa frequência de variação em contextos semelhantes a esse, resolveu-se retirar da análise as ocorrências de coordenação entre genitivos, seja entre PPs, como em (3b), ou entre PP + DP, como em (4a), o que levou à retirada das demais ocorrências em (4) e as do tipo em (3c).

- (4) a. “a terra é dele, sabe? Dele *mais* Salviano” (CZ-03)
 b. “a gente tamém...gostamos o amô de pai *e* (\emptyset_{PREP}) mãe do jeito que é” (CZ-02)
 c. “Fui criado na companhia da minha mãe, dos meus irmão, e... *e* (\emptyset_{PREP}) tia” (SP-10)
 d. “A água já tem outro pra cá perto da casa do véio *e* (\emptyset_{PREP}) a véia” (HV-05)
 e. “dado do esfoço de Deus *e* (\emptyset_{PREP}) ele” (HV-12)

Considerou-se também caso de coordenação entre genitivos, como em (4a), a sentença no exemplo em (5). Além disso, esse tipo não apresentou variação, ficando, portanto, fora da quantificação geral.

- (5) “a fazenda minha *mais* dos minino, os minino comprô...” (CZ-10)

A coordenação comitativa com DP genitivo, no *corpus* do português rural afro-brasileiro, encontra um paralelo com o crioulo de Cabo Verde, conforme apresentado no exemplo (35) da seção 1.3.1.

O resultado da quantificação dos dados quanto à variável *tipo de categoria coordenada* levou ao refinamento dos dados limitando o número de ocorrências em que se verifica a variação *mais/e*. Apenas as ocorrências do tipo em (2) foram mantidas, reduzindo o volume de dados a 358 ocorrências.

A quantificação dos dados referentes à variável *traço semântico dos nomes coordenados* apontou o uso categórico da variante *e*⁴⁶ quando os nomes coordenados possuem traço [-animado], como mostra a Tabela 4. Com a fixação dessa variável, buscou-se verificar a possibilidade de restrição semântica relacionada à animacidade dos nomes coordenados para a escolha de uma ou de outra variante, pois, segundo Haspelmath (2004), esse parece ser um traço que seleciona o conector coordenativo em algumas línguas, como Nêlêmwa, falada na Nova Caledônia, e Takia, falada em Papua Nova Guiné (cf. seção 1.2.1). A hipótese inicial de que tanto nomes [+animados] como [-animados] fossem coordenados pelo *mais* não se verificou no *corpus* investigado nesta pesquisa.

Tabela 4: Distribuição das variantes *mais/e* segundo o traço semântico dos nomes

	MAIS		E		TOTAL
	Ocorr.	%	Ocorr.	%	
[+ animado]	92	37	158	63	250
[- animado]	0	0	106	100	106
traço diferente	0	0	2	100	2
TOTAL	92	26	266	74	358

Os resultados da Tabela 4 mostram que a variação com o IV *mais*, nas comunidades analisadas, ocorre apenas entre os nomes com traço [+animado], conforme exemplo em (6). As ocorrências em que os nomes coordenados possuem o traço [-animado], como se vê em (7a), ou que apresentam esse traço em pelo menos um dos DPs, a exemplo de (7b), não apresentaram covariação estrita.

- (6) a. “Mãe... era mãe de sete fio: duas mulé *e* cinco home” (HV-05)
 b. “Eu ficô oiano Bernardino *mais* o fio dele” (HV-20)

⁴⁶ Em outro momento, verificar-se-á se o substrato teve interferência no uso categórico da variante *e*.

- (7) a. “Esse ano só plantei o milho *e* o feijão” (RC-06)
 b. “Foi casado pelo pade *e* pelo Siva⁴⁷ e largou minha fia” (SP-11)

Dessa forma, um novo corte nos dados foi necessário. As ocorrências que apontaram o uso categórico de uma das variantes, como em (7), foram desconsideradas para a análise do fenômeno variável *mais/e*⁴⁸. Verificou-se também que é categórico o uso do *e* coordenando DPs genéricos. A variável *referencialidade dos termos coordenados* contava com os fatores [+definido], [-definido] e genérico, porém, passaram para a fase seguinte da quantificação apenas as ocorrências dos dois primeiros tipos.

Após a realização dos cortes necessários devido ao uso categórico de uma das variantes, foram mantidos apenas os contextos em que se verificou covariação estrita. A distribuição entre as variantes é a seguinte: 92 da variante *mais*, 37% do total, e 158 da variante *e*, equivalendo a 63%. Essas 250 ocorrências constituem o *corpus* utilizado nesta pesquisa para verificar os condicionamentos linguísticos e sociais do fenômeno variável *mais/e*. Para a análise, inicialmente foram fixadas dez variáveis independentes, sendo quatro sociais e seis linguísticas, conforme (8) e (9), respectivamente.

- (8) a. Faixa etária
 b. Gênero
 c. Estada fora da comunidade
 d. Comunidade
- (9) a. Estatuto sintático dos DPs coordenados
 b. Tipo de categoria coordenada
 c. Traço semântico dos nomes coordenados
 d. Estrutura do primeiro DP coordenado
 e. Estrutura do segundo DP coordenado
 f. Referencialidade dos DPs coordenados

Como já foi dito, após os primeiros resultados quantitativos as variáveis linguísticas (9b) e (9c) foram descartadas.

⁴⁷ Pade = Padre. Refere-se ao casamento religioso celebrado pelo padre.

Siva = Civil. Refere-se ao casamento civil, regulamentado por lei.

⁴⁸ Embora se tenha a intuição de que seja possível a coordenação entre DPs com traço [- animado] pela variante *mais*, como em “Viajei para São Paulo e comprei uma bolsa mais um óculos de sol”.

O tratamento dos dados incluiu ainda critérios prévios de seleção. Não foram selecionadas do *corpus* ocorrências em que o *e* fosse utilizado nos seguintes casos: i) em expressões cristalizadas, como em (10); ii) na coordenação de números em geral: data, idade, valor monetário ou medidas, conforme os exemplos em (11); iii) em expressões de tempo e hora, como se vê em (12); iv) na função de marcador discursivo, como em (13), a seguir.

- (10) “quando o Amado Batista veio, eu fui no show dele. (...) já os oito que vêi que nem *Leandro e Leonardo* já eu num fui” (CZ-01)
- (11) a. “viajei dia sete de janêro... dia *vinte e cinco* de janêro” (CZ-04)
 b. “ela já tava dentro de seus *oitenta e seis* ano” (RC-11)
 c. “[Quanto tá o quilo aqui?] Cinco, *quato e cinquenta*, dependeno da... das carne” (SP-04)
 d. “mediu... dá *uma tarefa e meia*... perto de dois...” (SP-11)
- (12) a. “Eu fiquei lá *um ano e quato meses*” (CZ-01)
 b. “mais ô meno *uma hora e meia*” (RC-05)
- (13) a. “cê saí de sua casa pra i(r) numa festa ali, é tanta lambança, (...) dana a bebê, *e aí*, cum pôco, querem brigá” (HV-07)
 b. “no tempo de folga, eu ia passá na casa dos'amigo, *e tal*... encontrá colega...” (RC-05)

Todos os dados desconsiderados nesta etapa da pesquisa, por serem categóricos ou por não fazerem parte dos aspectos aqui analisados, são listados a seguir.

a) Desconsiderados por critérios prévios de seleção

- Coordenação entre números em geral: data, idade, valor monetário ou medidas;
- Expressões cristalizadas;
- Expressões de tempo e hora;
- Marcador discursivo.

b) Desconsiderados por *knockouts*

- Coordenação entre orações;

- Coordenação entre PPs;
- Coordenação entre termos em que um deles é um PP;
- Coordenação entre nomes com traço semântico [-animado];
- Coordenação entre nomes em que pelo menos um deles apresenta traço semântico [-animado];
- Coordenação entre nomes genéricos.

Após todos os recortes, foram submetidas 250 ocorrências ao pacote de programas VARBRUL, que selecionou como estatisticamente relevantes três variáveis sociais e três linguísticas, na seguinte ordem: *comunidade*, *estrutura do primeiro DP coordenado*, *estada fora da comunidade*, *referencialidade dos DPs coordenados*, *estatuto sintático dos DPs coordenados* e *faixa etária*. Na seção seguinte apresentam-se os condicionamentos sociais por ordem de relevância estatística.

3.1.1 Condicionamentos sociais

Fixaram-se quatro grupos de fatores para verificar os condicionamentos sociais das variantes *mais/e*, conforme apresentado em (8). Foram selecionados pelo VARBRUL, por ordem de relevância, *comunidade*, *estada fora da comunidade* e *faixa etária*. Como apresentado na seção 2.4.3.2, o padrão de fixação das variáveis sociais seguiu os procedimentos comumente empregados nas pesquisas realizadas no âmbito do Projeto Vertentes.

3.1.1.1 Comunidade

A análise de diversos aspectos da gramática do português rural afro-brasileiro⁴⁹ tem levado em consideração a condição de isolamento das comunidades investigadas, relacionando-a com a conservação de marcas procedentes do contato entre línguas. Com a fixação dessa variável, esperava-se que a variante *mais* fosse favorecida entre os falantes das comunidades com maior grau de isolamento, Helvécia e Cinzento, e a variante *e* fosse

⁴⁹ No âmbito do Projeto Vertentes, diversas pesquisas têm sido realizadas sobre aspectos gramaticais do português rural afro-brasileiro falado em comunidades isoladas do interior do estado da Bahia, muitas delas estão reunidas na obra *O Português Afro-brasileiro* (Lucchesi; Baxter; Ribeiro, 2009).

favorecida nas comunidades menos isoladas, ou pela presença de turistas, como ocorre em Rio de Contas, ou por estar localizada mais próxima à capital baiana, como Sapé.

Tabela 5: Distribuição das variantes *mais/e* segundo a *comunidade*

	MAIS			E			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Helvécia	44	62	.87	27	38	.13	71
Cinzento	32	42	.57	44	58	.43	76
Rio de Contas	12	28	.49	31	72	.51	43
Sapé	4	7	.07	56	93	.93	60
TOTAL	92	37	-	158	63	-	250

Nível de significância: .014

De acordo com os resultados apresentados na Tabela 5, a estratégia de coordenação com o *mais* é favorecida em Helvécia com .87 de peso relativo. Considerando que esta variante proceda do contato da LP, principalmente com línguas do substrato africano no Brasil, os números da Tabela 5 refletem o comportamento esperado, visto ser a comunidade de Helvécia considerada, no âmbito das pesquisas realizadas no Projeto Vertentes, a comunidade mais crioulezante, em cuja gramática se verificam os efeitos mais notáveis do contato linguístico e de um processo de transmissão linguística irregular (LUCCHESI, 2000; BAXTER E LUCCHESI, 1999). Além disso, de acordo com Lucchesi (2009), Helvécia possui características sócio-históricas que dão à comunidade esse *status* (cf. seção 2.4.1). A comunidade de Cinzento também tende à realização da variante *mais*, com um percentual de 42% e .57 de peso relativo.

Os resultados da Tabela 5 também confirmam a hipótese inicial em relação à comunidade de Sapé, que, por manter um grau de isolamento menor, favorece o uso da variante *e*, com um percentual de 93% e .93 de peso relativo. Quanto à comunidade de Rio de Contas, observa-se um percentual de 72% para a variante *e* e peso relativo de .51, apontando para uma neutralidade no uso das variantes *mais/e* para a coordenação entre DPs.

Em Rio de Contas e em Cinzento a diferença no uso das variantes é menos acentuada do que em Helvécia e Sapé, como mostra o Gráfico 1 a seguir.

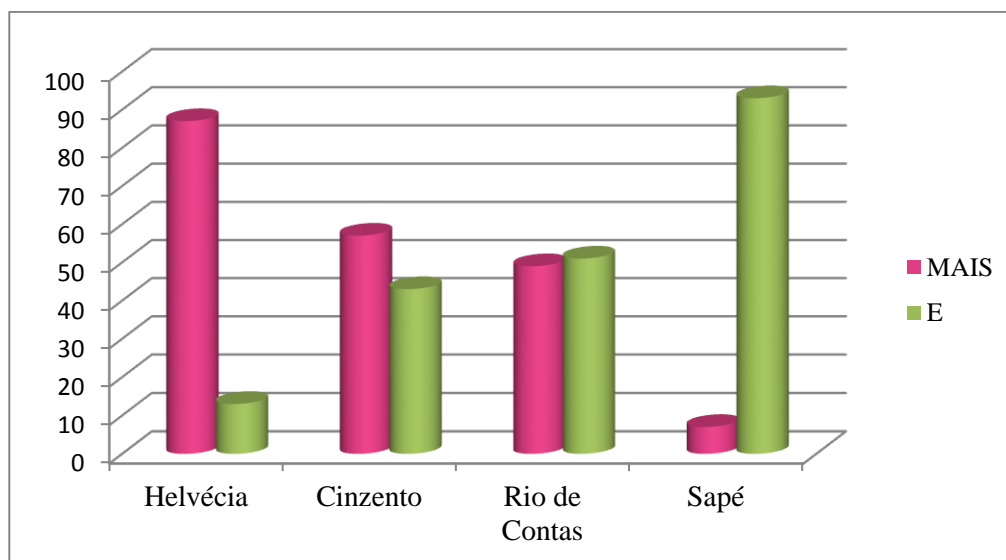


Gráfico 1: Distribuição das variantes *mais/e* segundo a *comunidade*

No Gráfico 1, em que as comunidades são apresentadas da esquerda para a direita de acordo com o grau de isolamento, observa-se um padrão decrescente de frequência de uso da variante *mais* que, segundo a hipótese deste trabalho, teria emergido do contato linguístico. Esses resultados confirmam a hipótese inicial.

3.1.1.2 Estada fora da comunidade

Com esta variável, buscou-se verificar a possibilidade de interferência linguística de outras normas nas comunidades investigadas. Nessa perspectiva, esperava-se que entre os indivíduos que nunca se ausentaram da comunidade ou que saíram por tempo inferior a seis meses a variante *mais* fosse favorecida. Os resultados são apresentados na Tabela 6, a seguir.

Tabela 6: Distribuição das variantes *mais/e* segundo a *estada fora da comunidade*

	MAIS			E			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Sim	38	31	.34	84	69	.66	122
Não	54	42	.65	74	58	.35	128
TOTAL	92	37	-	158	63	-	250

Nível de significância: .014

Como se vê na Tabela 6, a variante *mais* é favorecida com .65 de peso relativo na fala dos indivíduos que nunca deixaram sua comunidade, ou se ausentaram por período inferior a seis meses. Entre os indivíduos que deixaram a comunidade por tempo superior a seis meses

e, portanto, tiveram contato com padrões linguísticos externos, a variante *e* é favorecida com .66 de peso relativo. Tais resultados confirmam a hipótese inicial e a variante *mais* pode ser considerada o padrão linguístico para a coordenação entre DPs nas comunidades investigadas.

3.1.1.3 Faixa etária

A análise dos usos linguísticos nos intervalos etários fixados – Faixa I (20 a 40 anos), Faixa II (41 a 60 anos) e Faixa III (acima de 60 anos) – permite uma comparação de diferentes momentos da variação nas comunidades, conforme o modelo laboviano de estudo linguístico em *tempo aparente*. Desse modo, a fixação dessa variável visou verificar se a variação *mais/e* reflete uma variação estável ou uma mudança em progresso, em que a variante *mais* estaria sendo substituída pela variante *e*. Esperava-se que a variante *mais* fosse favorecida entre os falantes adultos e, sobretudo, entre os mais velhos, visto que é na fala desses indivíduos que se podem observar estágios anteriores dos aspectos linguísticos das comunidades.

Tabela 7: Distribuição das variantes *mais/e* segundo a faixa etária

	MAIS			E			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Faixa I	9	15	.31	52	85	.69	61
Faixa II	26	38	.41	43	62	.59	69
Faixa III	57	47	.65	63	52	.35	120
TOTAL	92	37	-	158	63	-	250

Nível de significância: .014

Os resultados, conforme a Tabela 7, revelam um padrão decrescente de uso da variante *mais* quando observados da Faixa III para a Faixa I. Na fala dos indivíduos mais velhos, essa variante é favorecida com .65 de peso relativo, seguido de .41 na fala dos indivíduos adultos e, .31, na fala dos mais jovens. Conseqüentemente, a variante urbana culta é mais favorecida na fala dos mais jovens, que são também os mais escolarizados das comunidades. Esse resultado geral, ilustrado no Gráfico 2, a seguir, não só ratifica a hipótese inicial levantada como também aponta para um processo de mudança em progresso, em que a variante *e* ganha espaço entre os novos indivíduos das comunidades como estratégia predominante para a coordenação entre DPs à medida que a variante *mais* tende a diminuir.

Observa-se que a forma que emergiu do contato entre línguas, como assumido neste trabalho para a variante *mais*, tem sido substituída pela forma culta. Esse cenário confirma a

hipótese de difusão da variante urbana culta para as demais regiões do país, em um processo de nivelamento linguístico (LUCCHESI, 2001 e 2006).

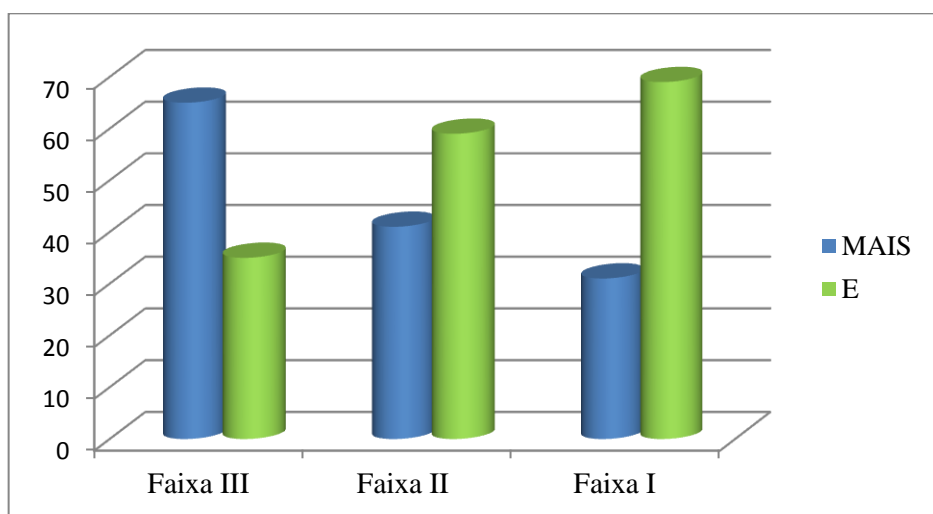


Gráfico 2: Distribuição das variantes *mais/e* segundo a *faixa etária*

A análise dos fatores sociais para esta pesquisa é bastante relevante, tendo em vista a hipótese de que o uso do IV *mais* em estruturas coordenadas resulta do contato linguístico entre a LP com o substrato africano no Brasil (cf. seção 2.1.2), por meio de relexificação e de reanálise (cf. seção 2.2.2).

3.1.2 Condicionamentos linguísticos

Nesta subseção, as variáveis linguísticas selecionadas e suas respectivas hipóteses norteadoras são apresentadas considerando a ordem de relevância apontada pelo VARBRUL. Foram selecionadas como relevantes as variáveis: *estrutura do primeiro DP coordenado*, *referencialidade dos DPs coordenados* e *estatuto sintático dos DPs coordenados*.

3.1.2.1 Estrutura do primeiro DP coordenado

Esta variável foi fixada com o intuito de investigar a relevância da composição interna do primeiro DP coordenado para a escolha do coordenador *mais* ou *e*. Inicialmente, hipotetizou-se que a escolha do tipo de determinante [+ ou - referencial], bem como a extensão do DP, pudessem condicionar a escolha do falante por uma ou outra variante. A presença de determinantes [+referenciais] e os DPs mais leves favoreceriam o uso do IV *mais*, os determinantes [-referenciais] e os DPs longos favoreceriam a variante *e*, menos saliente,

por serem mais pesados. Diversos tipos de estruturas foram codificados, porém, após os primeiros resultados quantitativos alguns foram desconsiderados. Os fatores abaixo relacionados constituem essa variável.

a) Artigo definido + Nome:

- (14) a. “adoeceu a mãe e a filha” (SP-03)
 b. “O pai *mais* a mãe que vai cuidá a terra” (HV-12)

b) Pronominal:

- (15) a. “Aí eu me casei, minha irmã casô, ficô ele e minha mãe.” (CZ-06)
 b. “Adonde eu *mais* meu marido fosse ni festa, nós levava ela.” (RC-11)

c) Nu:

- (16) a. “eu sei que tomava conta de Antôim e Antonha” (SP-11)
 b. “tem aquela venda lá de Aliomar *mais* Mané” (RC-21)

d) (Determinante) + possessivo + Nome:

- (17) a. “já num falo no meu pai e meu tio” (CZ-03)
 b. “Minha madrinha *mais* meu padrim morava lá pra cima” (HV-13)

e) (Determinante) + adjetivo + Nome:

- (18) a. “Tia Simplicia, Tia Gervarda, Tia Joana e Tia Telonha... Tudo tia, irmã de mainha...” (SP-11)
 b. “finada mãe *mais* Ilto soltava o côro mesmo” (HV-13)

f) (Determinante) + Nome + adjetivo ou locução adjetiva:

- (19) a. “tchem rapaz sortchêro e home casado caçano um pedaço de terra pa prantá e num acha” (SP-09)
 b. “as muié enfermera *mais* os dotô disse que coisa” (RC-21)

Os resultados dessa variável apresentam-se na Tabela 8, a seguir.

Tabela 8: Distribuição das variantes *mais/e* segundo a *estrutura do primeiro DP*

	MAIS			E			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Artigo definido + N	6	19	.30	26	81	.70	32
Pronominal	43	54	.72	36	46	.28	79
Nu	23	26	.33	66	74	.67	89
(D) + possessivo + N	9	39	.75	14	61	.25	23
(D) + N + adj./loc. adj.	2	22	.44	7	78	.56	9
(D) + adjetivo + N	4	36	.36	7	64	.64	11
TOTAL	87	36	-	156	64	-	243

Nível de significância: .014

Os resultados da Tabela 8 evidenciam que, quando o primeiro DP coordenado é pronominal, ou formado por (determinante) + pronome possessivo + nome, a variante *mais* é favorecida com .72 e .75 de peso relativo, respectivamente. A frequência de uso da variante *mais* nesses contextos é de 54% para DP pronominal e de 39% para DP constituído de (determinante) + pronome possessivo + nome. Essas formas são bastante referenciais quando se trata de pronomes de primeira/segunda pessoa, ou pronomes retomando referentes [+humanos], considerados mais altos na escala de referencialidade proposta por Cyrino, Duarte e Kato (2000), como se vê no Gráfico 3, a seguir. Nas ocorrências do *corpus*, os DPs pronominais são predominantemente referenciais, a exemplo de *ele* em (15a), cujo referente é o pai do informante, e *eu* em (15b), cujo referente é o próprio informante. Da mesma forma, os pronomes possessivos, como em (17). Vale ressaltar que, das 23 ocorrências de DPs nus coordenados pelo IV *mais*, 21 são referentes a nomes próprios, apenas as duas ocorrências em (20) não são. Apesar disso, trata-se de nomes mais referenciais.

- (20) “Ah! Eu vô chamá papai lá, papai *mais* mamãe. Eu vô chamá papai *mais* mamãe” (HV-22)

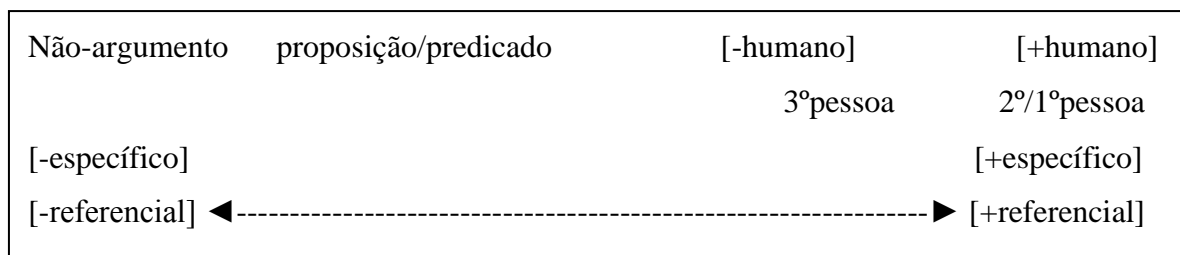


Gráfico 3: Hierarquia de referencialidade proposta por Cyrino, Kato, Duarte (2000)

Fonte: CYRINO; KATO; DUARTE (2000, p. 59).

Ocupa a posição mais alta da hierarquia os argumentos [+N, +humanos]. De acordo com as autoras, os pronomes eu (=falante) e você (=ouvinte) estão nessa posição, já que se trata de inerentemente humanos. Desse modo, é possível colocar nessa posição ainda os nomes próprios. O pronome de terceira pessoa, as proposições e os não argumentos ocupam, respectivamente, as posições mais baixas na hierarquia. As autoras concluem que à medida que o referente é [+humano] também é [+específico]. Tendo em vista a hierarquia proposta, a referencialidade dos DPs parece ser o principal fator de condicionamento da variante *mais*.

Como se vê na Tabela 8, o uso da variante *e* é favorecido quando o primeiro DP é constituído por artigo definido + nome, com .70 de peso relativo e frequência de 81%, um resultado que, inicialmente, parece contrariar a assunção de que a escolha do *mais* é condicionada pelo fator [+definido]; e nome nu, com .67 de peso relativo e 74% de frequência.

O uso do artigo definido em português exprime maior definitude, como sugere a tradição gramatical, mas esse uso não é exclusivo, o artigo definido pode também encabeçar um DP genérico. Um DP genérico pode ter a forma de um nome nu (*Macaco gosta de banana*), mas também pode ser estruturado com artigo definido, conforme exemplo de Lyons (1977, p. 160), em (20a), e de Mira Mateus *et. al* (2003, p. 221), em (20b). Esse comportamento do artigo definido no PB pode justificar o fato de ser a variante *e* condicionada em contexto de coordenação entre DPs introduzidos por esse tipo de artigo.

- (20) a. O leão é um animal pacífico.
b. O homem é um animal racional.

A variante *e* também é favorecida quando o primeiro DP é constituído por determinante + adjetivo + nome, com .64 de peso relativo e frequência de 64%, e por determinante + nome + adjetivo/locução adjetiva, com 78% de frequência e .56 de peso relativo. Esses tipos de DPs parecem ser mais pesados, com mais de uma categoria em sua

estrutura, isso pode levar à escolha do *e*, menos saliente. Contudo, como mostram os números da Tabela 8, essas estruturas não são muitas no *corpus* investigado.

As ocorrências com DPs indefinidos, constituídos por numeral ou artigo/pronome indefinido, apresentaram-se estatisticamente irrelevantes para as variantes.

Os resultados da variável *estrutura do primeiro DP coordenado* são ratificados pelos resultados da variável apresentada na subseção seguinte, também selecionada pelo VARBRUL como estatisticamente relevante.

3.1.2.2 Referencialidade dos DPs⁵⁰ coordenados

Esta variável foi fixada a fim de se investigar o condicionamento da referencialidade do DP para a escolha das variantes *mais* ou *e*. A hipótese inicial é que os DPs [+referenciais] favorecessem a variante *mais*. Os DPs coordenados foram codificados como genéricos ou como [+definido] ou [-definido], de acordo com as definições e exemplos a seguir.

a) **Genérico**⁵¹: designa o DP que tem como característica ser um elemento representativo de uma determinada classe ou espécie (LYONS, 1977, p. 160). São exemplos de DPs coordenados genéricos [*fêjão e mio*] e [*mulé e home*] nos seguintes exemplos:

- (22) a. “É a roça de café nós toca... arroz, fêjão *e* mio” (RC-13)
 b. “Ajuntô mais de cinquenta em cima de um, entre mulé *e* home”
 (SP-01)

b) [+definido]: designa os DPs coordenados identificáveis, ou seja, em que ambos os referentes são determinados, podendo ser nomeados e descritos de forma que possam ser distinguidos de outros indivíduos no universo do discurso (LYONS, 1977). Os DPs coordenados [*eu e Agripino*] e [*meu pai mais minha mãe*] nos exemplos em (23) foram caracterizados dessa forma por apresentarem essas propriedades.

⁵⁰ Refere-se à categoria DP mesmo que a categoria subordinada seja um PP, visto que aquele é que carrega traços de referencialidade.

⁵¹ Este fator foi fixado considerando que o DP genérico pode ser estruturado com determinantes ou pode ter a forma de um nome nu, como apresentado na seção 3.1.2.1.

- (23) a. “quem tomô conta foi a gente, eu *e* Agripino, meu marido...” (SP-11)
 b. “Meu pai *mais* minha mãe nunca me recramô em nada” (RC-11)

c) [-definido]: designa os DPs coordenados em que um deles ou ambos não são identificáveis pelo falante e/ou pelo ouvinte; porém são DPs que podem ter referentes específicos ou não⁵² (LYONS, 1977). Em (24a) ambos os DPs não são identificáveis para o ouvinte; em (24b), o segundo DPs não é identificável para ambos os interlocutores no universo do discurso.

- (24) a. “Eu tenho um irmão *e* um cunhado” (HV-05)
 b. “Quando começô, eu num me lembro, qu'eu num tava aí não. Minha mãe *mai* os ôto é que sabe” (HV-03)

Sobre a construção de referencialidade, Duarte e Oliveira (2003, p. 221) afirmam que a língua dispõe de diferentes operações de determinação, isto é, “processos de natureza semântico-pragmática que constroem o valor referencial de uma dada ocorrência de um nominal” expressados por meio do especificador e do número do nominal. Tais operações são aplicadas aos nomes comuns para construir seu valor referencial. A leitura [\pm definida; \pm específica] depende de uma gama de fatores.

Como já foi dito no início deste capítulo, apenas entre os DPs referenciais, com traço [+definido] e [-definido] verificou-se covariação estrita. Nos DPs genéricos, no *corpus* investigado, a coordenação foi realizada apenas pela variante *e*, como em (22). A Tabela 9 apresenta os resultados obtidos para essa variável.

Tabela 9: Distribuição das variantes *mais/e* segundo a referencialidade dos DPs

	MAIS			E			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
[+ definido]	84	40	.58	123	60	.42	207
[- definido]	5	13	.15	35	88	.85	40
TOTAL	89	36	-	158	64	-	247

Nível de significância: .014

⁵² Aos sintagmas indefinidos aplica-se a noção de especificidade. O traço [+ específico] caracteriza o DP indefinido cujo referente pode ser parafraseado pela expressão *determinado(a)*. O referente pode passar a ser tratado como identificável pelos interlocutores. O traço [-específico] caracteriza o DP indefinido cujo referente pode ser parafraseado pela expressão *algum(a)* (LYONS, 1977, p. 156). Contudo, neste trabalho foram caracterizados como DPs com o traço [-definido] tanto os [+específicos] quanto os [-específicos].

Os dados acima demonstram que há uma atuação direta do fator referencialidade no uso do *mais*, pois em contexto [+definido] essa variante é favorecida com .58 de peso relativo e 40% de frequência. Números bem acima dos apresentados para o contexto [-definido], com 13% de frequência e peso relativo de .15. A ausência de DPs genéricos coordenados pelo IV *mais* parece confirmar o condicionamento da referencialidade. Em relação à variante de prestígio *e*, os números apontam na direção oposta: o contexto [-definido] é favorecido com .85 de peso relativo e frequência de 88%, diferentemente do contexto [+definido], cujos resultados são de .42 de peso relativo e 60% de frequência.

Entre os DPs com referencialidade [+definida] estão os nomes próprios de pessoas, os pronomes de primeira e segunda pessoa, e os DPs com referencialidade [+humana], de forma geral, conforme a hierarquia de referencialidade apresentada no Gráfico 3. Da leitura da Tabela 9, é possível perceber que há uma predominância de DPs desse tipo, que parece resultar do tipo de narrativa feita pelos informantes.

A referencialidade parece ser um fator condicionante para a o uso do IV *mais* coordenador nas comunidades investigadas, confirmando a hipótese inicial, como apontam os resultados obtidos para as variáveis *estrutura do primeiro DP* e *referencialidade do DP*. A hipótese deste trabalho é que o IV *mais* tenha relexificado uma entrada copiada do substrato e teve suas funções ampliadas, passando a atuar não só como coordenador, mas também como subordinador comitativo. Tendo em vista ser a animacidade um traço presente na expressão comitativa, um traço mais alto na hierarquia de referencialidade proposta por Cyrino, Kato e Duarte (2000), é possível pensar que o uso do *mais* como coordenador seja favorecido entre DPs [+definidos] por influência do contexto comitativo.

3.1.2.3 Estatuto sintático dos termos coordenados

Esta variável foi fixada com o objetivo de investigar se a função sintática dos DPs coordenados condicionaria a escolha das variantes, considerando que as posições argumentais adjacentes ao verbo, sujeito e objeto direto, favorecessem o uso da variante *mais*. Encontrou-se covariação estrita na coordenação de DPs que exercem as funções sintáticas elencadas a seguir.

a) **Sujeito:** De acordo com Duarte (2003, p. 281), tal constituinte é o argumento mais elevado na Hierarquia Temática, a expressão com função de tópico e a expressão que desencadeia concordância verbal. É a relação gramatical típica do argumento externo de verbos transitivos,

como em (25a), e inergativos⁵³, como em (25b), do argumento externo do predicador secundário em frases copulativas e do argumento interno direto de verbos inacusativos.

- (25) a. “*ele e as menina* fazeno uma massa” (CZ-12)
 b. “se *eu mais ele* morrê é tá tudo cabado” (CZ-05)

b) **Objeto direto:** De acordo com Duarte (2003, p. 284), tem a relação gramatical de objeto direto o argumento interno direto de verbos de dois ou três lugares. Tais constituintes ocorrem tipicamente sem preposição e recebem seu papel temático do verbo predicador, que é comumente o papel temático de Tema⁵⁴. São exemplos de DPs coordenados com função de objeto direto os DPs em itálico nos exemplos em (26).

- (26) a. “ela tem muitos *irmão e irmãs*” (RC-21)
 b. “Eu vô chamá *papai mais mamãe*” (HV-22)

c) **Aposto:** De acordo com Brito (2003, p. 365-366), pertence ao grupo de modificadores apositivos, que “constituem um comentário avaliativo do locutor sobre uma entidade denotada pela expressão nominal”. Categorialmente podem ser frásicos (os teus primos, *que vivem na Califórnia*, chegam hoje), nominais ou adjetivais (Adriano, *o Imperador de Roma*, era um homem só). A codificação das ocorrências com esse fator incluiu os tipos de aposto encontrados: de sujeito, como em (27a); de sujeito clivado, como em (27b); de objeto direto, como em (27c) e de predicativo, como em (27d).

- (27) a. “Nós somo tudo primo... *eu mais Florênço, eu mais Madalena, eu mais Salviano...*” (CZ-11)
 b. “quem tomô conta foi a gente, *eu e Agripino*, meu marido” (SP-11)
 c. “Ela teve um casal... *um homezinho e uma mulhé...* gêmo” (CZ-03)
 d. “Ai, quando a parteira saia, era duas parteira, *eu e ela*” (SP-11)

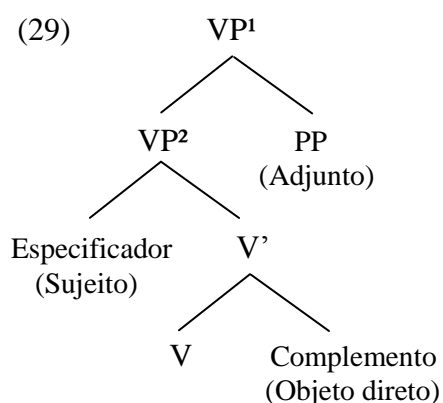
⁵³ De acordo com Mira Mateus et. al. (2003, p. 300), tanto os verbos inergativos como os inacusativos são verbos de um lugar que determinam o esquema relacional *SU V* e são denominados pela tradição gramatical luso-brasileira de *intransitivos*. Contudo, os verbos inergativos, ou verdadeiros intransitivos, selecionam *argumento externo* com a relação gramatical de sujeito, enquanto os inacusativos selecionam *argumento interno* que ocorre com relação gramatical de sujeito.

⁵⁴ Tema: a entidade que muda de lugar, de posse ou de estado, em frases que descrevem situações dinâmicas; a entidade criada pela atividade expressa pelo verbo ou afetada por tal atividade; a entidade não controladora nem experienciadora de uma situação não dinâmica (MIRA MATEUS et. al., 2003, p. 189).

d) **Adjunto de VP:** De acordo com Brito (2003, p. 416), embora participem do VP, os adjuntos não são nós irmãos de V, ou seja, não são complementos subcategorizados pelo V, desse modo, ocupam posição de adjunção à categoria VP. Nos exemplos em (28), retirados do *corpus*, observa-se a coordenação de termos com função sintática de adjunto de VP. Em (28a), o PP coordenado por *e* [*mais as fia e os fio*] modifica o verbo *vir*, e em (28b) o PP coordenado por *mais* [*com seu Quéte mais Ilto*] modifica o verbo *trabalhar*.

- (28) a. “Só essa que mora lá em valença que vem *mais as fia e os fio*” (SP-09)
 b. “trabaia aí em Helvéça, com *seu Quéte mais Ilto* na roça” (HV-13)

Em (29), apresentam-se as posições do sujeito, do objeto direto e do adjunto de VP, no modelo da Teoria X-Barra de Chomsky (1970)⁵⁵. As posições adjacentes ao núcleo V, sujeito e objeto direto podem favorecer a variante *mais*.



Acrescentou-se ainda um código para identificar os casos em que a função sintática não pôde ser depreendida do texto. Outras funções menos significativas estatisticamente não constam na análise. A distribuição das variantes conforme o estatuto sintático é apresentada na Tabela 10.

⁵⁵ Chomsky (1970) propõe um esquema estrutural que permite analisar os constituintes de uma sentença a partir de uma representação hierárquica com o seguinte modelo básico: núcleo X, lugar atribuído às categorias lexicais (nome, verbo, adjetivo, advérbio ou preposição); nível intermediário (X'), projetado pela combinação do núcleo X com um Complemento; e nível máximo de um sintagma (XP), projetado pela combinação da projeção intermediária com um Especificador. As possibilidades de estruturação de um constituinte são demonstradas pelo modelo X-Barra, e, nessa perspectiva, a forma final da estrutura será definida pelo núcleo e suas propriedades lexicais. As representações desse trabalho serão feitas mediante o modelo X-Barra por se tratar de uma teoria bastante estudada que proporciona a demonstração das possibilidades estruturais de constituintes.

Tabela 10: Distribuição das variantes *mais/e* segundo o *estatuto sintático dos termos*

	MAIS			E			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Sujeito	65	52	.67	61	48	.33	126
Objeto direto	5	11	.22	40	89	.78	45
Adjunto	7	20	.25	28	80	.75	35
Aposto	5	42	.61	7	58	.39	12
Indefinida	5	31	.46	11	69	.54	16
TOTAL	87	37	-	147	63	-	234

Nível de significância: .014

A leitura da Tabela 10 revela que a variante *mais* é favorecida na coordenação de DPs com função sintática de sujeito, com .67 de peso relativo e 52% de frequência, uma posição adjacente ao verbo, confirmando a hipótese proposta. Essa é a posição *mais* alta na hierarquia temática que abriga DPs com papel de agente, experienciador e causativo, papéis temáticos que predominantemente exigem DPs com traço [+animados].

A função de aposto também é favorável ao uso da variante *mais*. Isso se confirma pelo percentual de 42% e .61 de peso relativo. Vale ressaltar que o tipo de aposto *mais* encontrado foi o aposto de sujeito, como em (27a) e (27b), desmembrando seu núcleo, cujos tipos de papel temático possuem, geralmente, traço [+animado].

Considerando que não foi encontrada no *corpus* a coordenação de DPs com traço [-animado] pela variante *mais*, os números dessa tabela parecem apontar *mais* para a relevância do papel temático dos DPs coordenados do que para a estrutura propriamente dita, como se considerou inicialmente.

Para a função de objeto direto o VARBRUL apontou .78 de peso relativo para a variante *e*, e frequência de 89%. Os DPs com papel temático de tema podem constituir-se tanto de nomes com traço [+animando] quanto de nomes com traço [-animado]. Os resultados do fator objeto direto não confirmam a hipótese inicial de que termos adjacentes ao verbo favoreceriam o uso do IV *mais*, contudo, reforçam a relevância da animacidade na escolha dessa variante. Quanto à posição de adjunto, um termo não adjacente ao verbo, a variante *e* é favorecida com .75 de peso relativo e 80% de frequência, um resultado que confirma a hipótese inicial.

Na seção seguinte, apresentam-se os condicionamentos sociais e linguísticos para as variantes *mais/e*.

3.1.3 Síntese dos fatores condicionantes da variável *mais/e*

Os grupos de fatores analisados mostram que os contextos mais favoráveis ao uso da variante *mais* são:

- a. As comunidades de Helvécia e Cinzento;
- b. A permanência dos informantes na comunidade;
- c. A fala dos indivíduos adultos e idosos;
- d. O traço [+definido] dos DPs;
- e. As funções sintáticas de sujeito ou aposto.

Por sua vez, a variante urbana culta *e* é favorecida pelos seguintes fatores:

- a. As comunidades de Rio de Contas e Sapé;
- b. A fala dos indivíduos que já se ausentaram de sua comunidade por pelo menos seis meses;
- c. A fala dos indivíduos com menos idade, da Faixa I;
- d. O traço [-definido] dos DPs;
- e. As funções sintáticas de objeto direto e adjunto.

Na seção seguinte, apresenta-se a análise dos dados referentes à variação *mais/com*.

3.2 ESTRATÉGIAS DE SUBORDINAÇÃO EM ESTRUTURAS COMITATIVAS: MAIS/COM

Esta seção trata da variação entre as preposições *mais/com* como estratégias de subordinação, conforme exemplos em (30), assumindo-se que o uso do *mais* subordinador resulta do contato entre a LP e as línguas do substrato africano no Brasil (cf. seções 2.1.1 e 2.1.2). Nesta pesquisa, entende-se subordinação como um processo sintático em que um termo é hierarquicamente dependente de outro, numa relação de complementação ou de adjunção. Nesse processo, o termo subordinado desempenha uma função sintática e uma função semântica em relação ao termo subordinante (cf. seção 1.1.2).

- (30) a. João casou **com/mais** *Maria ontem*. (= complemento)
 b. Eu saí **com/mais** *minha amiga*. (= adjunto)

Em (30a), o PP [*com/mais Maria*] exerce função sintática de complemento do verbo oblíquo *casar*; ao introduzir argumento interno o IV *mais* é uma preposição funcional que atribui caso ao DP complemento. Em (30b) o PP [*com/mais minha amiga*] desempenha função sintática de adjunto do verbo *sair* e o IV *mais*, nesse caso, é uma preposição lexical já que introduz o adjunto e atribui a ele caso e papel temático. A covariação estrita entre as variantes *mais/com* verifica-se em estruturas subordinadas em que a preposição introduz DP comitativo, como em (30b), e recíproco, como em (30a)⁵⁶.

Na seleção dos dados a serem investigados, coletaram-se ocorrências, considerando a covariação estrita entre os IVs *mais* e *com*. Descartaram-se ocorrências do conector *com* nos contextos elencados a seguir.

a) Nas expressões *junto com/juntamente com*, como em (31), e em outras expressões cristalizadas, como se vê em (32).

- (31) a. “tinha que passá *junto com* a mãe aqui” (HV-05)
 b. “eu *juntamente com* eles...senti que era (...) a principal homenagem”
 (CZ-04)

- (32) “a gente tem que *apegá* muito **com** *Deus*” (HV-07)

b) Em expressões com o verbo *estar* (estar com) significando posse, como no exemplo em (33).

- (33) “Agora mesmo, tô ca...**com** (= tenho) uma peruzinha em casa e uma pata” (CZ-12)

⁵⁶ Neste trabalho não se discute se essas estruturas são derivadas de uma coordenação ou se são construções de complemento ou de adjunto. O que se pretende fazer neste trabalho é apresentar uma descrição sócio-histórica do uso do IV *mais*, em construções que, por motivo metodológico, são consideradas subordinativas, comitativas ou recíprocas, e coordenativas, como ocorre em diferentes línguas.

Durante o processo de seleção das ocorrências, verificou-se ainda outro tipo de fenômeno: o uso do IV *com* como coordenador entre DPs, variando com o conector *e*. Contudo, tal fenômeno foi desconsiderado nesta etapa da pesquisa⁵⁷ e as ocorrências de *com* coordenador não foram selecionadas para a análise da variável dependente *mais/com* em estrutura subordinativa.

Após seleção criteriosa, codificaram-se 702 ocorrências de acordo com dez variáveis independentes: quatro sociais, listadas em (34), e seis linguísticas, listadas em (35). A distribuição entre as variantes, conforme Tabela 1, no capítulo 2, é a seguinte: 188 da variante *mais*, correspondendo a 27% do total, e 514 da variante *com*, 73%.

- (34) a. Faixa etária
 b. Estada fora da comunidade
 c. Gênero
 d. Comunidade
- (35) a. Estatuto sintático do PP encabeçado por *com/mais*
 b. Estrutura do DP encabeçado por *com/mais*
 c. Referencialidade do DP encabeçado por *com/mais*
 d. Traço semântico do nome
 e. Distância entre o verbo e o PP subordinado
 f. Função semântica do PP encabeçado por *com/mais*

Na primeira rodada dos dados, foram apontados alguns *knockouts*. O primeiro deles diz respeito à variável *traço semântico do nome*, fixada com o objetivo de identificar se o traço semântico do DP, [+animado] ou [-animado], restringe a ocorrência da variante *mais* ou da variante *com*. Tendo em vista que, segundo Haspelmath (2004), a animacidade é um traço que restringe o uso de coordenadores em línguas (cf. seção 1.1.1), buscou-se investigar se o mesmo se verifica com os subordinadores *mais/com*. Verificou-se que a animacidade parece restringir o uso das variantes e o IV *mais* é condicionado por DPs [+animados].

Os resultados dessa rodada apontaram para a ausência da variante *mais* introduzindo DPs com traço [-animado], revelando um contexto categórico da variante *com*. Desse modo, foram excluídas da análise 71 ocorrências, reduzindo o *corpus* para 631 dados. Em (36),

⁵⁷ Objetiva-se investigar posteriormente o uso do *com* coordenativo, incluindo-se os resultados na análise que tem sido desenvolvida aqui (em estruturas coordenativas).

observa-se a ocorrência das variantes *mais/com* em contexto [+animado], e em (37) a ocorrência de *com* em contexto [-animado].

- (36) a. “a gente divestia, saia por aí **com** as amiga, os amigo, né?” (CZ-03)
 b. “eu fui criada **mais** minhas tia” (CZ-10)

- (37) “vem a enfermera **com** aquelas injeção...” (HV-04)

No exemplo em (37), a noção expressa pela preposição *com* não é de companhia e também não há uma relação de reciprocidade entre o argumento externo do verbo e o argumento interno. A noção expressa pela preposição nesse exemplo é a de *continente/conteúdo*. Da análise dessa variável, pode-se inferir que os DPs com traço [-animado] não exprimem de forma tão evidente a noção de copresença/adição, presente nas construções com comitativos e recíprocos, mas valores semânticos como os de instrumento, de modo, de causa, de tempo, dentre outras apresentadas para a preposição *com* na seção 1.2.2. O uso categórico de DPs com traço [+animado] parece confirmar tal hipótese. Entretanto, em alguns tipos de DPs com traço [+animado] não se verificou variação.

Os números referentes à variável *função semântica do PP*⁵⁸ obtidos nesta primeira rodada levaram ainda à outra redução nos dados do *corpus*. Dos 08 fatores, verificou-se em 06 o uso categórico da variante *e*, conforme Tabela 11. Os fatores semânticos fixados estão elencados a seguir.

a) **Agente**: designa “a entidade controladora, tipicamente humana, de uma dada situação” (MIRA MATEUS *et. al.*, 2003, p. 188).

- (38) a. “cheguei, fiz um exame ali **com** seu num sei quem, (...) num cunheço quem é [o] médico não...” (CZ-08)

b) **Paciente**: designa “a entidade que sofre o efeito de alguma ação” (MIOTO *et. al.*, 2007, p. 126).

- (39) “Aí eles calabora **com** a gente quando a gente vai” (CZ-03)

⁵⁸ Agradeço ao Prof. Dr. Dante Lucchesi por ter sugerido a fixação dessa variável.

c) **Alvo**: designa “a entidade para a qual algo foi transferido, num sentido locativo ou não” (MIRA MATEUS *et. al.*, 2003, p. 190).

(40) “via tudo... a gente falava **com** ela, ela entendia a gente” (SP-02)

d) **Fonte**: designa “a entidade que está na origem de uma dada situação, embora sem a controlar” (MIRA MATEUS *et. al.*, 2003, p. 189).

(41) a. “vai em Jequié e consegue **com** ela que é a secretára...” (CZ-06)

e) **Tema**: designa a entidade que muda de lugar, de posse ou de estado, em frases que descrevem situações dinâmicas; a entidade criada pela atividade expressa pelo verbo ou afetada por tal atividade; a entidade não controladora nem experienciadora de uma situação não dinâmica (MIRA MATEUS *et. al.*, 2003, p. 189).

(42) “tem dia que ela enfeza **com** nós” (CZ-02)

f) **Conteúdo**: atribuição para o DP que expressa o sentido de “estar contido” em relação ao argumento sujeito.

(43) “tava **cum** filhim no braço” (CZ-10)

g) **Comitativo**: atribuição para o DP com valor de companhia.

(44) a. “Cantá, só sei **co** os ôto” (HV-12)

b. “sempre quando eu vô dêxo ela **mais** minha mãe, né?” (HV-09)

h) **Recíproco**: atribuição para o DP que nutre uma relação de reciprocidade com o argumento sujeito.

(45) a. “Aí quano eu casei, casei **cum** fi dela” (CZ-10)

b. “Não, mora na terra **mais** o... o irmão dela.” (CZ-11)

Neste trabalho faz-se diferença entre DPs comitativos e recíprocos a fim de verificar os contextos de covariação estrita. A respeito das construções comitativas e recíprocas, Mira Mateus *et. al.* (2003, p. 397) consideram que uma frase com estrutura subordinada em que o PP possui valor comitativo pode ser parafraseada por coordenação e por construções que exprimem reciprocidade, desse modo, a hipótese inicial é que a variação *mais/com* ocorresse em estruturas como (44) e (45).

Godoy (2009) chama de recíproco o DP subordinado que nutre com o argumento sujeito uma relação lógica de reciprocidade possibilitando as formas simples, como em (46a), e descontínua, como em (46b) e (46c).

- (46) a. João e Maria conversaram.
 b. João conversou com Maria.
 c. Maria conversou com João.

Para a autora, há reciprocidade quando a informação de uma sentença B está, de fato, contida na informação de uma sentença A, como ocorre com as sentenças simples e descontínuas em (46) (GODOY, 2009). Alguns trabalhos têm proposto que as estruturas em (46) são sinônimas⁵⁹, outros têm tentado demonstrar que não há sinonímia nessas estruturas⁶⁰, como discute e admite Godoy (2009); contudo, esta não será uma discussão contemplada neste trabalho, visto que o objetivo é fazer uma descrição sócio-histórica da variação dos IVs *mais/com*. Esse será o tema da continuidade deste trabalho.

Em sentenças como *namorar com alguém*, o PP subordinado não expressa a noção comitativa/de companhia, mas de reciprocidade entre o argumento externo e o argumento interno. Nessas construções é possível a variação com o IV *mais*, talvez porque haja a ideia de acréscimo/adição de mais um elemento ao evento. Para ser recíproco, necessariamente, mais de um participante tem que estar envolvido no evento.

Um DP comitativo pode ser parafraseado por uma sentença em que aparece coordenado a um sujeito, como afirmam Mira Mateus *et. al.*, assemelhando-se sintaticamente às formas em que os recíprocos ocorrem. Os exemplos em (47) e (48) mostram a semelhança estrutural na superfície sintática das construções com DP comitativo e recíproco. Contudo, neste trabalho, considera-se, como Godoy (2009), que as estruturas em (47b) e (48b) são semanticamente diferentes.

⁵⁹ Lakoff e Peters (1969) e Kayne (1994).

⁶⁰ Dimitriadis (2005) e Ilari (1987).

- (47) a. João e Maria chegaram.
b. João chegou [com Maria]_{COM}
- (48) a. João e Maria casaram.
b. João casou [com Maria]_{REC}

A fim de classificar as ocorrências selecionadas do *corpus* como comitativo ou recíproco, foram aplicados os seguintes testes propostos de Godoy (2009).

	DP recíproco ⁶¹	DP comitativo
Teste 1: acarretamento sentencial (comparar sentença simples com descontínua)	A informação da sentença simples está contida na informação da sentença descontínua.	A informação da sentença simples não está contida, de fato, na informação da sentença descontínua.
Teste 2: sentença simples com argumento singular	Quando apagado, o DP faz falta para a sentença (Ex: <i>*João comparou</i>) ou ocorre com o sentido de que há um argumento implícito (João casou [com Maria]).	O DP é apagado sem fazer falta para a sentença (Ex: <i>João jantou</i>).
Teste 3: substituição de COM por NA COMPANHIA DE	O DP é agramatical (Ex: <i>*João casou na companhia de Maria</i>).	O DP é gramatical (Ex: <i>João jantou na companhia de Maria</i>).

Quadro 9: Testes para DPs comitativos e recíprocos.

Conforme apresentado na Tabela 11, a seguir, a frequência de uso da variante *mais* introduzindo DP comitativo é de dezenove pontos percentuais acima da média de 30%, ou seja, os DPs comitativos condicionam o uso dessa variante. O uso do IV *com*, com DPs comitativos, fica dezenove pontos abaixo da média de 70% para essa variante. Em construções com verbos recíprocos, os resultados apontam 28% de frequência para a variante *mais*, e 72% para a variante *com*. Esse resultado mostra que a estratégia comitativa com o IV *mais* prevalece na fala dos indivíduos das comunidades investigadas.

⁶¹ Intui-se que o papel temático dos DPs é o mesmo, tendo em vista que na estrutura profunda, assumindo o modelo da Regência e Ligação (Chomsky, 1981, 1986), nessas sentenças, os DPs sejam coordenados, embora isso fira o Critério Theta de Chomsky, quando se observa a estrutura superficial. É necessário investigar que operações sintáticas são requeridas na derivação desse tipo de coordenação. Como essa discussão ultrapassa o limite deste trabalho, serão investigadas posteriormente as possibilidades sintáticas no modelo Minimalista (CHOMSKY, 1993, 1995) que permitam confirmar ou refutar tal intuição.

Tabela 11: Distribuição das variantes *mais/com* segundo a *função semântica do PP*

	MAIS		COM		TOTAL
	Ocorr.	%	Ocorr.	%	
Agente	-	0	2	100	2
Paciente	-	0	6	100	6
Alvo	-	0	79	100	79
Fonte	-	0	7	100	7
Tema	-	0	17	100	17
Conteúdo	-	0	2	100	2
Comitativo	103	49	109	51	212
Recíproco	85	28	221	72	306
TOTAL	188	30	443	70	631

Foram considerados para análise apenas os números referentes às duas últimas linhas da tabela, que totalizam 518 ocorrências. De forma geral, os dados não incluídos na quantificação final dos dados são listados a seguir.

a) Desconsiderados por critérios prévios de seleção

- Expressões *junto com* e *juntamente com*;
- Outras expressões em geral;
- *Com* coordenador;
- *Com* na construção “estar com” significando posse.

b) Desconsiderados por *knockouts*

- Traço semântico [-animado];
- Papel temático de agente, paciente, alvo, fonte, tema e o fator “conteúdo”.

Nas subseções seguintes, apresentam-se os resultados obtidos, considerando as variáveis independentes sociais e linguísticas selecionadas como estatisticamente relevantes pelo VARBRUL, na seguinte ordem: *comunidade*, *estatuto sintático do PP*, *referencialidade do DP*, *faixa etária*, *estada fora da comunidade* e *gênero*.

3.2.1 Condicionamentos sociais

Foram selecionadas como estatisticamente relevantes pelo VARBRUL todas as variáveis sociais fixadas: *comunidade*, *faixa etária*, *estada fora da comunidade* e *gênero*. Os

valores referentes ao número de ocorrências, percentual e peso relativo, apresentam-se a seguir por ordem de relevância.

3.2.1.1 Comunidade

Considerando-se a condição de isolamento das quatro comunidades investigadas e relacionando-a com a conservação de marcas procedentes do contato linguístico, esperava-se que a variante *mais* fosse favorecida entre os falantes das comunidades com maior grau de isolamento, Helvécia e Cinzento, e a variante culta *e* fosse favorecida nas comunidades de Rio de Contas, por ser frequentada por muitos turistas e pessoas de fora da comunidade, e de Sapé, por estar localizada mais próxima à capital baiana (cf. seção 2.4.1). Os resultados obtidos são apresentados na Tabela 12 a seguir.

Tabela 12: Distribuição das variantes *mais/com* segundo a comunidade

	MAIS			COM			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Cinzento	74	63	.84	43	37	.16	117
Rio de Contas	39	45	.59	47	55	.41	86
Sapé	52	23	.33	173	77	.67	225
Helvécia	23	26	.33	67	74	.67	90
TOTAL	188	36	-	330	64	-	518

Nível de significância: .015

Como se pode observar na Tabela 12, a variante *mais* é favorecida primeiramente na comunidade de Cinzento, com .84 de peso relativo, e 63% de frequência. Esses dados ratificam a hipótese inicial, já que esta comunidade está entre as que teriam mantido um grau de isolamento relativamente alto e, portanto, guardariam vestígio de um processo de aquisição por TLI (LUCCHESI, 2000; BAXTER E LUCCHESI, 1999). A variante *mais* é também favorecida na comunidade de Rio de Contas, com peso relativo igual a .59 e 45% de frequência de uso. Entre os informantes de Sapé a hipótese inicial também se confirma; a tendência maior de uso é a variante urbana culta *e* com frequência de 77% e peso relativo de .67.

Entretanto, o resultado referente à variável comunidade não se confirma em relação à Helvécia. Diversos estudos sobre outros aspectos da gramática vêm apontando-a como a comunidade mais crioulezante no âmbito do Projeto Vertentes, porém, o uso do *mais*

subordinador não parece ser favorecido na comunidade, apresentado .33 de peso relativo e frequência de 26%, diferentemente do *mais* coordenador, que é favorecido com .87 de peso relativo. Uma hipótese levantada é a de que, nessa comunidade, a variante *com* seria uma concorrente do *mais*, no que diz respeito à possibilidade de atuar como coordenador e subordinador, conforme exemplo em (49), semelhante ao que ocorre em algumas línguas crioulas (cf. seções 1.3.1 e 1.3.2). Porém, deixou-se para realizar a verificação desta hipótese posteriormente⁶².

(49) “Era tempo de minha mãe *com* (= e) meu pai, tempo de cativêro, qu'eles morava lá”. (HV-13)

3.2.1.2 Faixa etária

Considerando o modelo laboviano de estudo linguístico em *tempo aparente*, fez-se uma comparação dos diferentes momentos, representados pela fala dos informantes das três faixas etárias, a fim de verificar se o fenômeno variável *mais/com* reflete um cenário de variação estável ou uma mudança em progresso nas comunidades, com a variante *com* substituindo a variante *mais* que emergiu do contato. Esperava-se que a variante *mais* fosse favorecida na fala dos adultos e idosos, que permite visualizar estágios anteriores da língua.

Tabela 13: Distribuição das variantes *mais/com* segundo a *faixa etária*

	MAIS			COM			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Faixa I	34	29	.37	84	71	.63	118
Faixa II	75	35	.52	137	65	.48	212
Faixa III	79	42	.57	109	58	.43	188
TOTAL	188	36	-	330	64	-	518

Nível de significância: .015

Os números da Tabela 13 revelam que a tendência de uso da variante culta *com* é maior entre os informantes com menos idade. Na Faixa I, tal variante é favorecida com .63 de peso relativo e 71% de frequência. Observa-se um padrão decrescente de uso à medida que a faixa etária do informante aumenta. Entre os mais idosos, o IV *com* é desfavorecido com .43 de peso relativo e frequência de 58%.

⁶² Agradeço ao Prof. Dr. Alan Baxter por ter sugerido a investigação da hipótese.

Por outro lado, a estratégia comitativa com o IV *mais* é favorecida entre os indivíduos mais velhos das comunidades, ou seja, em estágios anteriores da língua. Nesse caso, o peso relativo é de .57 e a frequência é de 42%; entre os indivíduos mais novos, da Faixa I, esse IV possui frequência de 29% e peso relativo de .37, o que aponta para seu desfavorecimento.

Tal cenário permite inferir que a variante culta *com* está substituindo a variante *mais* nessas comunidades. O resultado apresentado na Tabela 13 parece confirmar a hipótese de que a função do IV *mais* como subordinador em construções comitativas e recíprocas parece resultar do contato entre a LP com línguas do substrato africano no Brasil mediante um processo de relexificação (cf. seção 2.2.2)

O resultado também é relevante para verificar a trajetória do *mais* subordinador nas comunidades analisadas. O padrão apresentado na Tabela 13, ilustrado no Gráfico 4, a seguir, aponta para um processo de mudança em progresso, em que a variante *com*, a forma culta, ganha espaço entre os indivíduos mais novos das comunidades como principal estratégia de subordinação em construções comitativas e recíprocas, à medida que a variante *mais* tende a diminuir, confirmando o nivelamento linguístico proposto por Lucchesi (2001 e 2006).

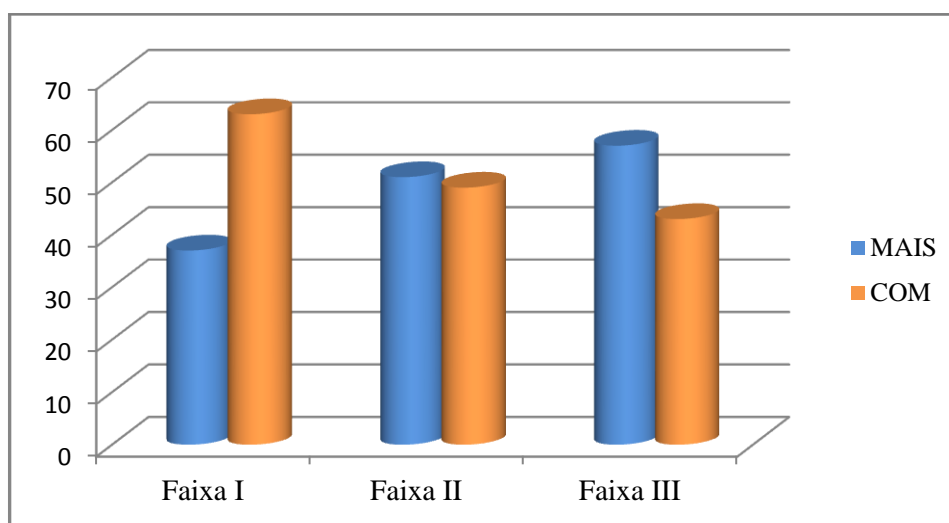


Gráfico 4: Distribuição das variantes *mais/com* segundo a *faixa etária*

3.2.1.3 Estada fora da comunidade

Como já foi dito na seção 2.4.3.2, esta variável busca verificar a possibilidade de interferência linguística de outras normas nas comunidades investigadas. A hipótese inicial é que o uso da variante conservadora *mais* fosse favorecido entre os informantes que nunca saíram de sua comunidade ou não se ausentaram dela por mais de seis meses, na perspectiva

de que tais indivíduos tendem a manter os padrões linguísticos próprios de seu dialeto. Os resultados obtidos constam na Tabela 14, a seguir.

Tabela 14: Distribuição das variantes *mais/com* segundo a *estada fora da comunidade*

	MAIS			COM			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Sim	86	35	.42	162	65	.58	248
Não	102	38	.57	168	62	.43	270
TOTAL	188	36	-	330	64	-	518

Nível de significância: .015

Tais resultados, conforme a Tabela 14, ratificam a hipótese inicial. Percebe-se que o maior uso da variante *mais* está na fala daqueles que não se ausentaram de sua comunidade. Nesse caso, o IV *mais* é favorecido com .57 de peso relativo, e frequência de 38%; a variante *com* é favorecida com .58 de peso relativo e 65% de frequência na fala dos informantes que se mantiveram fora da comunidade por mais de 6 meses.

Uma comparação dos resultados apresentados nesta seção com os resultados do fenômeno variável *mais/e* mostra que são semelhantes. O uso do IV *mais* como coordenador entre DPs e subordinador em construções comitativas e recíprocas parece ser o padrão linguístico nas comunidades investigadas.

3.2.1.4 Gênero

Considerando a hipótese de que o uso do *mais* subordinador resulta do contato entre línguas, esperava-se que esse IV fosse favorecido na fala das mulheres, considerada mais conservadora no âmbito das comunidades investigadas, por conta de uma vida social mais voltada para o âmbito doméstico. Na Tabela 15, a seguir, são apresentados os números obtidos.

Tabela 15: Distribuição das variantes *mais/com* segundo o *gênero*

	MAIS			COM			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Feminino	148	40	.54	225	60	.46	373
Masculino	40	28	.39	105	72	.61	145
TOTAL	188	36	-	330	64	-	518

Nível de significância: .015

Tais resultados confirmam a hipótese inicial. A estratégia comitativa com o IV *mais* é favorecida na fala das mulheres com .54 de peso relativo, e 40% de frequência; o uso de *com* é favorecido na fala dos homens com .61 de peso relativo e frequência de 72%.

Com exceção do gênero, as demais variáveis sociais (comunidade, faixa etária e estada fora da comunidade) também foram selecionadas como relevantes para a variação *mais/e*. O uso do IV *mais* como coordenador e como subordinador é condicionado por fatores sociais que apontam para o contato linguístico, uma vez que a norma linguística das comunidades se verifica mais claramente na fala dos indivíduos que não se ausentaram e na fala dos mais velhos, conforme argumentado na seção 2.4.3.2. Para ambos os fenômenos variáveis os resultados apontam para uma mudança em progresso nas comunidades investigadas.

Esses resultados são bastante relevantes para a investigação da hipótese proposta na introdução deste trabalho, de que a ampliação das funções do IV *mais*, como coordenador e como subordinador, resulta do contato da LP com o substrato africano no Brasil por meio de um processo de relexificação.

Na seção seguinte, apresentam-se os resultados referentes às variáveis independentes linguísticas.

3.2.2 Condicionamentos linguísticos

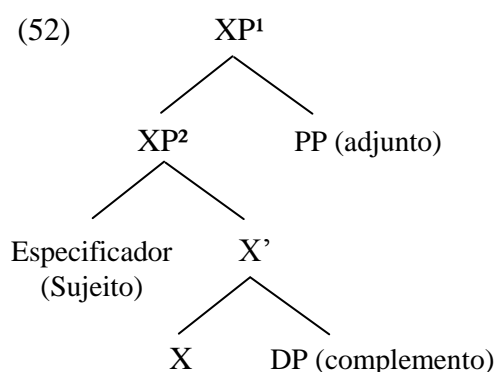
Esta subseção trata das variáveis linguísticas, apresentadas juntamente com as hipóteses norteadoras, dos resultados estatísticos e da interpretação desses resultados, considerando a ordem de relevância apontada pelo VARBRUL. Foram selecionadas como relevantes duas variáveis: o *estatuto sintático do PP* e a *referencialidade do DP*.

3.2.2.1 Estatuto sintático do PP encabeçado por *mais/com*

Em relação a essa variável, os dados foram codificados considerando-se duas funções: complemento verbal, como em (50), e adjunto de VP, como em (51), considerando-se as seguintes hipóteses: i) a variante *com* seria favorecida quando o PP fosse complemento, visto que, nessa construção, a preposição é funcional; ii) a variante *mais* seria favorecida ao introduzir PP adjunto, por funcionar, nesse caso, como uma preposição lexical. Imaginou-se que o valor semântico de acréscimo/adição seria mantido nessas construções, não sendo utilizado como preposição lexical, esse traço semântico do IV *mais* teria sido o motivo da escolha do substrato para relexificar uma entrada copiada de sua(s) língua(s) (cf. seção 2.2.2).

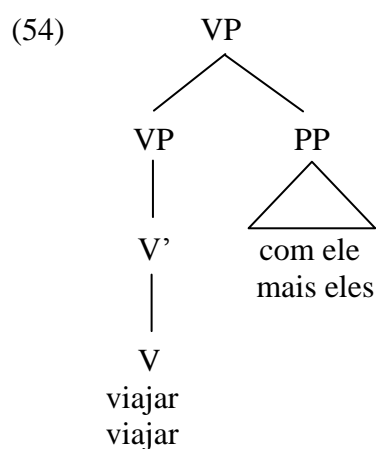
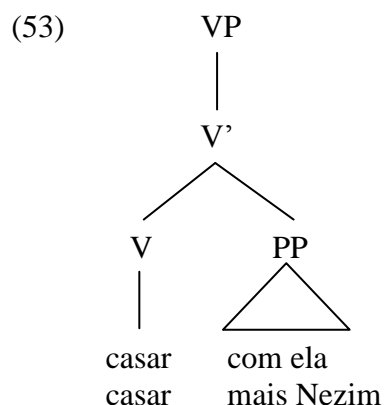
- (50) a. “qu'eu casei [*com* ela]” (HV-22)
 b. “Denestina (...) é casado [*mais* Nezim]” (HV22)
- (51) c. “Aí eu saía, viajava [*com* ele], ia direto.” (CZ-07)
 d. “poque ninguém nunca fez esse modo de...de viaja [*mais* eles] não...” (CZ-09)

O verbo *casar*, como em (50), exige complemento preposicionado, e os PPs [*com ela*] e [*mais Nezim*] são gerados internamente ao verbo. Em (51), entretanto, os PPs [*com ele*] e [*mais eles*] não são complementos, já que o verbo *viajar*, nesses contextos, é inergativo. Tais PPs, por não serem selecionados pelo verbo, não fazem parte de sua estrutura argumental, são adjungidos a VP e analisados pela tradição gramatical como adjuntos adverbiais de companhia. O papel temático dos DPs é atribuído pelas preposições *com/mais* que os introduzem. As posições de complemento e de adjunto no modelo da Teoria X-Barra⁶³ de Chomsky (1970) estão representadas em (52):



De acordo com Mito *et. al.* (2007, p. 84-85), dentro do nível sintagmático máximo, demonstrado em (52) por XP², não é prevista uma posição para adjunto, apenas para complemento e especificador, já que o adjunto é um constituinte autônomo e está contido em uma projeção máxima. Para representá-lo, dobra-se a categoria a qual o constituinte PP é adjungido, como o XP¹ em (52), e sua posição na estrutura é a de irmão de XP². A representação arbórea de (50) pode ser observada em (53) e a representação de (51) em (54):

⁶³ Sobre a teoria X-Barra ver nota 55 da seção 3.1.2.3.



Entre as variáveis independentes estabelecidas para análise, *estatuto sintático do PP* foi considerada a mais relevante estatisticamente pelo programa a que foram submetidos os dados. Na Tabela 16, a seguir, apresentam-se os resultados obtidos:

Tabela 16: Distribuição das variantes *mais/com* segundo o *estatuto sintático do PP*

	MAIS			COM			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
Adjunto	157	44	.62	197	56	.38	354
Complemento	31	19	.25	133	81	.75	164
TOTAL	188	36	-	330	64	-	518

Nível de significância: .015

A leitura da Tabela 16 revela que, quando o PP tem o estatuto sintático de adjunto de verbo, a variante *mais* é favorecida com .62 de peso relativo e frequência de 44%. Por outro lado, o PP complemento de verbo é favorecido pela variante *com*, apresentando peso relativo de .75 e frequência de 81%.

Essa variável é fundamental para a análise da natureza do IV *mais* na função de preposição. De acordo com os pressupostos teóricos da Gramática Gerativa, as preposições

possuem natureza híbrida, podendo ser funcionais ou lexicais/predicadoras. A preposição funcional encabeça um DP que é argumento de outro predicador e atribui a esse DP apenas caso, o papel temático é atribuído pelo predicador que seleciona esse DP. Já a preposição lexical encabeça um DP atribuindo-lhe caso e papel temático, uma vez que se trata de um argumento da própria preposição; o predicador do adjunto é a preposição lexical (MIOTO *et. al.*, 2007).

Os resultados, conforme a Tabela 16, mostram que a variante *mais* é favorecida quando introduz DP adjunto, na função de preposição lexical, evidenciando a relevância do traço semântico do IV *mais*. Esse resultado está em consonância com a hipótese proposta na seção 2.2.2, de que, principalmente, o valor semântico de acréscimo/adição tenha licenciado a relexificação de uma entrada copiada do substrato com o IV *mais* no PB.

3.2.2.2 Referencialidade do DP⁶⁴ encabeçado por *mais/com*

A fixação desta variável objetivou investigar se a escolha das variantes *mais/com* é condicionada pelo nível de referencialidade do DP. Esperava-se que os DPs [+referenciais] favorecessem a variante *mais*. O DP subordinado foi codificado como genérico ou referencial, sendo a referencialidade [+definida] ou [-definida], nos mesmos termos analisados para a variável *mais/e* (cf. seção 3.1.2.2). A seguir, apresentam-se exemplos do *corpus* segundo os fatores que constituem essa variável: DP genérico, em (55), DP [+definido], em (56), e DP [-definido], em (57).

- (55) a. “no terno a gente num sai *cum* muié, só sai *com* home, né?” (CZ-04)
 b. “quande ‘cê está num lugá *mais* gente de cidade” (HV-13)
- (56) a. “Ana Isidora mora *com* a filha dela” (CZ-05)
 b. “ele trabaia aqui *mais* eu.” (CZ-12)
- (57) a. “compade Jacó tá conversano *com* ôtro fazendêro de... criação de gado” (HV-20)
 b. “Eu fui *mais* um prim’meu que já tinha ido.” (RC-06)

⁶⁴ Refere-se à categoria DP mesmo que a categoria subordinada seja PP, visto que aquele é que carrega traços de referencialidade.

A quantificação dessa variável demonstrou que os DPs mais referenciais condicionam a escolha do informante pela variante *mais*; de outro modo, a variante *com* é condicionada pelos DPs menos referenciais e genéricos, confirmando a hipótese inicial, como mostram os dados da Tabela 17, a seguir. Esse resultado assemelha-se ao obtido para o fenômeno variável *mais/e*.

Tabela 17: Distribuição das variantes *mais/com* segundo a *referencialidade do DP*

	MAIS			COM			TOTAL
	Ocorr.	%	P. R.	Ocorr.	%	P. R.	
[+definido]	176	39	.54	271	61	.46	447
[-definido]	7	21	.32	27	79	.68	34
Genérico	5	14	.21	32	86	.79	37
TOTAL	188	36	-	330	64	-	518

Nível de significância: .015

Os resultados da Tabela 17 mostram que DPs com traço [+definido] favorecem o uso do *mais* com .54 de peso relativo, enquanto DPs [-definidos] e genéricos favorecem o IV *com*, como apontam os pesos relativos de .68 e .79, respectivamente. É possível que o tipo de entrevista favoreça a ocorrência de DPs mais definidos, visto que são narrativas de fatos ocorridos com o informante ou na comunidade.

Comparando esse resultado com os obtidos para a variação *mais/e*, observa-se cenário semelhante: a variante *mais* é favorecida em contextos [+definidos] com .58 de peso relativo, enquanto os [-definidos] desfavorecem-na, com peso relativo de .15. Na pouca quantidade de dados para a variação *mais/e*, no *corpus* investigado, a possibilidade de DP genérico introduzido pelo IV *mais*, como em (55b), não se verificou (cf. seção 3.1.2.2). Para a variação *mais/com*, há uma pequena quantidade de ocorrências, 37, das quais apenas 5 são com a variante *mais*. Tais resultados para o IV *mais* coordenador e subordinador parecem apontar para um traço [+referencial] desse IV.

Semelhante ao que ocorre com a variação *mais/e*, os resultados da Tabela 17 são ratificados pelos resultados obtidos para a variável *estrutura do DP encabeçado por com/mais*. Verificou-se que DPs realizados por pronome, como em (58), favorecem o uso do *mais*, com três pontos percentuais acima da média obtida, que é de 36% para a variante. O IV *com* é favorecido entre DPs constituídos por: artigo indefinido + nome, pronome indefinido + nome, numeral + nome. Essas estruturas foram agrupadas de acordo com a semelhança entre traços dos determinantes e o resultado quantitativo desse fator revelou quatorze pontos

percentuais acima da média obtida, que é de 64% para a variante. Em (59), apresentam-se alguns exemplos do *corpus* referentes a essas estruturas indefinidas.

- (58) a. “qu'eu casei *com ela*” (HV-22)
 b. “Tem um só que fica *mais eu*” (SP-02)
- (59) a. “já sabia e casô *com ôta mulhé...*” (SP-11)
 b. “cansemo de sai(r) daqui...eu *mais uma pessoa véia...*” (CZ-12)

Na seção seguinte, apresentam-se os condicionamentos sociais e linguísticos para as variantes *mais/com*.

3.2.3 Síntese dos fatores condicionantes da variável *mais/com*

Os fatores analisados mostram que os contextos mais favoráveis ao uso do *mais* são:

- a. As comunidades de Cinzento e Rio de Contas;
- b. A fala dos indivíduos adultos e idosos;
- c. A fala dos indivíduos que permaneceram na comunidade;
- d. A fala mais conservadora das mulheres;
- e. O PP com função de adjunto do verbo;
- f. O DP com referencialidade [+definida];
- g. O DP pronominal.

Por sua vez, a variante *com* é favorecida pelos seguintes fatores:

- a. As comunidades de Sapé e Helvécia;
- b. A fala dos indivíduos com menos idade, das faixas 1 e 2;
- c. A fala dos indivíduos que se ausentaram da comunidade por pelo menos seis meses;
- d. A fala mais inovadora dos homens;
- e. O PP com função de complemento do verbo;
- f. O DP com referencialidade [-definida] e o DP genérico;
- g. O DP formado com determinante menos definido (artigo/pronome indefinido e numeral).

CONSIDERAÇÕES FINAIS

Nesse trabalho, investigou-se o uso do IV *mais* como coordenador de DPs, alternando com o IV *e*, e como subordinador em estruturas comitativas e recíprocas, alternando com a preposição *com*, no português rural afro-brasileiro, buscando apresentar novas evidências para relacionar a realidade linguística do português popular do estado da Bahia ao contato entre línguas, em consonância com o objetivo central do Projeto Vertentes ao qual essa pesquisa está vinculada. No português popular do estado da Bahia, o contato linguístico parece se sentir, dentre outros, pelo fenômeno que aqui se investigou.

No capítulo 1, apresentou-se uma descrição do uso do IV *mais* como coordenador e como subordinador no português popular do estado da Bahia e no português rural afro-brasileiro, bem como de fenômeno semelhante em diferentes pidgins, crioulos e outras línguas africanas. No crioulo de Cabo Verde, nas variedades de Barlavento, o IV *ma* (do português *mais*) também exerce as funções de coordenador de DPs e subordinador comitativo; nas variedades de Sotavento é o IV *ku* (do português *com*). Os DPs com traço [+ referencial] também são coordenados pela estratégia comitativa, enquanto os demais são coordenados pelo IV *i*. Vale ressaltar que o padrão que se observa no português rural afrodescendente para o IV *mais* como coordenador de DPs e não de frases e orações assemelha-se ao padrão das línguas nigero-congolesas e não ao padrão das línguas europeias, em que um mesmo IV é utilizado para coordenar DPs e frases/orações, um IV diferente da partícula comitativa. Esse cenário parece apontar para a confirmação da hipótese defendida ao longo desse trabalho, de que o contato linguístico da LP no Brasil com outras línguas que realizam a coordenação comitativa permitiu a inserção do IV *mais* no contexto dessas novas funções.

No capítulo 2 discutiu-se a sócio-história do PB e o cenário de multilinguismo no Brasil desde suas origens. Apesar de a LP ter se tornado a língua oficial da Colônia e de ter sido implementada em todas as regiões do país por conta de diversos fatores sócio-históricos, é possível observar que o contato entre línguas está na origem da formação do PB, sobretudo das variedades populares. No Brasil, as línguas indígenas, que inicialmente chegavam a mais de mil (RODRIGUES, 1993) e representavam a principal forma de comunicação no primeiro século de colonização, principalmente em sua forma de *línguas gerais*, praticamente desapareceram por conta das condições a que foram submetidas as populações indígenas. Os índios foram dizimados ou recuaram para territórios cada vez mais remotos do país. Já as línguas africanas conviveram por mais tempo entre seus falantes nativos e a população que se

formava no Brasil, a exemplo do quimbundo, falado em Salvador por africanos oriundos de Angola, como atesta o documento histórico-linguístico de Dias (1697).

As marcas do contato entre a LP e outras línguas da colonização do Brasil, principalmente as línguas africanas, verificam-se no PB atual, seja superficialmente, no léxico, seja de forma mais profunda, como afirmam estudiosos e pesquisadores, que atribuem diversas estruturas originais do PB ao contato linguístico. Assumiu-se que, no Brasil, o contato da LP com as línguas do substrato africano que realizam a coordenação por meio da estratégia comitativa, tenha desencadeado um processo de relexificação, em que a matriz fonética do IV *mais*, da língua portuguesa, a língua-alvo, foi assimilada servindo como rótulo para um “novo” vocábulo do PB, com propriedades sintáticas e semânticas do IV original copiado da(s) língua(s) do substrato, um IV que pode coordenar DPs e expressar comitativo. O uso do IV *mais* em construções recíprocas parece ampliar ainda mais a funcionalidade desse IV, e pode ter sido licenciado pela leitura de *acréscimo* de uma estrutura recíproca.

No capítulo 3, analisaram-se os condicionamentos sociais e linguísticos que favorecem a variante *mais* nas comunidades rurais afrodescendentes. A hipótese de que o uso do IV *mais* como coordenador e subordinador emergiu do contato linguístico parece se confirmar por meio da investigação que se fez dos condicionamentos sociais, cuja análise é bastante relevante para a investigação do contato linguístico. De modo geral, tem-se o seguinte cenário nas comunidades investigadas: i) para o fenômeno variável *mais/e*, o IV *mais* é favorecido nas comunidades de Helvécia e Cinzento, na fala dos indivíduos que permaneceram na comunidade e na fala dos indivíduos adultos e idosos; ii) para o fenômeno variável *mais/com*, o IV *mais* é favorecido nas comunidades de Cinzento e Rio de Contas, na fala dos indivíduos que permaneceram na comunidade, na fala dos indivíduos adultos e idosos e na fala das mulheres.

Os resultados obtidos mostram que a variante *mais*, tanto na coordenação entre DPs quanto na subordinação comitativa/recíproca, é favorecida entre os indivíduos que permaneceram em sua comunidade. Esse cenário aponta mais uma vez para a confirmação da hipótese desta pesquisa, tendo em vista que a fala desses indivíduos tende a manter os padrões linguísticos próprios de sua norma. Em relação à variável gênero, para o fenômeno variável *mais/com* a estratégia comitativa com o IV *mais* é favorecida na fala mais conservadora das mulheres.

Nas quatro comunidades afrodescendentes investigadas a variante *mais* foi favorecida como estratégia coordenativa e comitativa/recíproca na fala dos indivíduos com mais idade, aqueles que apontam para os usos linguísticos mais antigos de suas comunidades. A análise

desses fatores sociais parece apontar o IV *mais* como a forma que emergiu do contato linguístico, como já havia sido considerado por Baxter (1987).

Por meio do recurso do *tempo aparente*, é possível perceber que ambos os fenômenos passam por um processo de mudança em progresso. Os resultados apontam para uma frequência crescente de uso das formas padrão *e/com*, na norma linguística das comunidades, na fala dos indivíduos mais jovens, que por sua vez, transmitirão tais formas para seus descendentes. Tal cenário parece confirmar o nivelamento linguístico que estaria fazendo desaparecer as marcas mais acentuadas do contato entre línguas por conta da difusão/penetração do padrão linguístico dos grandes centros urbanos no seio de comunidades outrora isoladas (LUCCHESI, 2006, 2009).

Em relação aos condicionamentos linguísticos, foram considerados relevantes, para ambos os fenômenos variáveis, a *referencialidade* e o *estatuto sintático*. O IV *mais* coordenador e subordinador é favorecido em contextos [+ referenciais], enquanto os contextos genéricos apontaram para seu desfavorecimento, como subordinador, ou mesmo para a sua inexistência no *corpus* investigado, como coordenador. Em relação ao *estatuto sintático*, o *mais* subordinador é favorecido na função de preposição lexical, quando introduz adjunto de VP, o que parece estar relacionado ao seu valor semântico de *acréscimo/adição*; enquanto o *mais* coordenador é condicionado pelas funções de sujeito, uma posição mais alta na hierarquia temática, e aposto de sujeito. Os DPs envolvidos em tais funções possuem papel temático de agente, experienciador e causativo, que exigem, em geral, DPs com traço [+ animado].

Para a variação *mais/e* a *estrutura do primeiro DP coordenado* também foi selecionada como relevante. A variante *mais* é condicionada por DPs constituídos de pronomes pessoais e possessivos, determinantes com traço [+ referencial], conforme a hierarquia de referencialidade proposta por Cyrino, Duarte e Kato (2000). Embora a *estrutura do DP* não tenha sido selecionada para a variação *mais/com*, é um fator que ratifica a referencialidade, pois o DP pronominal favorece a variante *mais*, enquanto a variante *com* é condicionada por DPs com estruturas indefinidas, como artigo/pronome indefinido e numeral.

Os resultados obtidos, de maneira geral, apontam para a confirmação da hipótese defendida ao longo desta dissertação. Entretanto, essa discussão não se encerra nessas linhas. É necessário dar continuidade à investigação de diversos aspectos que não puderam ser abordados nesta etapa da pesquisa, bem como ampliar a investigação para outros *corpora*.

REFERÊNCIAS

ALMEIDA, Napoleão Mendes de. *Gramática Metódica da Língua Portuguesa*. 45. ed. São Paulo: Saraiva, 2005.

ARAÚJO, P. J. P. Domínios conceituais das construções locativas, existenciais, comitativas e possessivas em línguas bantas. 2013. Tese (Doutorado) - Faculdade de Filosofia, Letras e Ciências humanas. Universidade de São Paulo, São Paulo.

ARENDS, J.; KOUWENBERG, S.; SMITH, N. Theories focusing on the non-European input. In: ARENDS, J.; MUYSKEN, P.; SMITH, N.(eds). *Pidgins and Creoles: an introduction*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1994.

AZEREDO, José Carlos de. *Gramática Houaiss da Língua Portuguesa*. 2. ed. São Paulo: Publifolha, 2008.

BAPTISTA, Marlyse. *The Syntax of Cape Verdean Creole: the Sotavento varieties*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2002.

BAXTER, A. Creole-like traces in rural Brazilian Portuguese dialects. Department of Spanish. La Trobe University, Bundoora, Australia. 47p, MS, 1987.

BAXTER, A. A Contribuição das Comunidades Afro-brasileiras Isoladas para o Debate sobre a Crioulização Prévia: um exemplo do Estado da Bahia. In: D'ANDRADE, E. & KIHM, A. (Orgs) *Actas do colóquio sobre Crioulos de Base Lexical Portuguesa*. Lisboa: Colibri, p.07-35, 1992.

BAXTER, A. Transmissão Geracional Irregular na História do Português Brasileiro: divergências nas vertentes afro-brasileiras. *Revista Internacional de Língua Portuguesa*, v. 14, p. 72-90, 1995.

BAXTER, Alan. Línguas pidgin e crioulas. In: FARIA, I. et. al. (Orgs). *Introdução à Linguística Geral e Portuguesa*. Lisboa: Caminho, 1996, p. 535-549.

BAXTER, Alan; LUCCHESI, Dante. Processos de descrioulização no sistema verbal de um dialeto rural brasileiro. *Papiá - Revista de Crioulos de Base Ibérica*, Brasília, Universidade de Brasília, v.2, n.2, p.59-71, 1993.

BAXTER, Alan; LUCCHESI, Dante. A relevância dos processos de pidginização e

crioulização na formação da língua portuguesa no Brasil. *Estudos Linguísticos e Literários*, Salvador, n.19, p.65-83, 1997.

BAXTER, Alan; LUCCHESI, Dante. Un paso más hacia la definición del pasado criollo del dialecto afro-brasileño de Helvécia (Bahia). In: ZIMMERMANN, Klaus (Ed.). *Lenguas criollas de base lexical española y portuguesa*. Madri: Iberoamericana, 1999. p.119-141.

BAXTER, Alan; LUCCHESI, Dante. Processos de crioulização na história sociolinguística do Brasil. In: CARDOSO, Suzana Alice; MOTTA, Jacyra; MATOS e SILVA, Rosa Virgínia (Org.). *Quinhentos anos de história lingüística do Brasil*. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006.

BECHARA, Evanildo. *Moderna gramática portuguesa*. 37. ed. rev. e ampl. Rio de Janeiro: Nova Fronteira, 2009.

BESTEN, H.; MUYSKEN, P.; SMITH, N. Theories focusing on the European input. In: ARENDS, J.; MUYSKEN, P.; SMITH, N.(eds). *Pidgins and Creoles: an introduction*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1994.

BICKERTON, Derek. Creole languages and the Bioprogram. In: NEWMAYER, Frederick (Ed.). *Linguistics: the Cambridge survey*. v.2. Cambridge: Cambridge University Press, 1988. p.268-284.

BONVINI, E. Línguas africanas e português falado no Brasil. In: PETTER, M; FIORIN, J. L. (orgs). *África no Brasil: a formação da língua portuguesa*. São Paulo: Contexto, 2008.

BORBA, F. S. et. al. *Dicionário gramatical de verbos do português contemporâneo do Brasil*. 2. ed. São Paulo: Editora Unesp, 1991.

BORBA, F.S. et al. *Dicionário de usos do português do Brasil*. São Paulo, Ática, 2002.

BRITO, A. M. Categorias sintáticas. In: In: MIRA MATEUS, Maria Helena. et al. *Gramática da língua portuguesa*. 5.ed. Lisboa: Caminho, 2003.

CALLOU, D.; BARBOSA, A.; LOPES, C. O português do Brasil: polarização sociolinguística. In: CARDOSO, Suzana Alice; MOTTA, Jacyra; MATOS e SILVA, Rosa Virgínia (Org.). *Quinhentos anos de história lingüística do Brasil*. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006.

CARDEIRA, Esperança. *História do português*. Lisboa: Caminho, 2006. p. 87-96.

CASTRO, Yeda Pessoa de. *A língua mina-jeje no Brasil: um falar africano em Ouro Preto do século XVIII*. Belo Horizonte: Secretaria de Estado da Cultura, 2002.

CEGALLA, Domingos Paschoal. *Novíssima Gramática da Língua Portuguesa*. 13. ed. São Paulo: Companhia Editora Nacional, 1973.

CHOMSKY, Noam. Remarks on Nominalization. In: JACOBS, R.; ROSEMBAUM, P. (Eds.). *Readings in English Transformational Grammar*. Waltham: Ginn, 1970.

CHOMSKY, Noam. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.

CHOMSKY, Noam. *Barriers*. Cambridge, Massachusetts: MIT Press, 1986.

CHOMSKY, Noam. A Minimalist Program for Linguistic Theory. In: HALE, K.; KEYSER, S. (eds). *The View from Building 20*. Cambridge, Massachusetts: MIT Press, 1993.

CHOMSKY, Noam. *The Minimalist Program*. Cambridge, MA: MIT Press, 1995.

COUTO, Hildo Honório do. *Introdução ao estudo das línguas crioulas e pidgins*. Brasília: Editora da UNB, 1996.

COUTO, Hildo Honório do. Hipótese da relexificação na gênese dos crioulos e pidgins. *Revista da ABRALIN*, Florianópolis, v. 1, n. 1, p. 221-250, 2002.

CUNHA, Celso Ferreira da. *Gramática da Língua Portuguesa*. 11. ed. Rio de Janeiro: FAE, 1986.

CUNHA, C. F. *Gramática do português contemporâneo*. Belo Horizonte: Bernardo Álvares, 1970.

CYRINO, S. M. L. Mudança sintática e português brasileiro. In: CASTILHO, A.T. et.al. (Org). *Descrição, história e aquisição do português brasileiro - estudos em homenagem à Mary Kato*. Campinas: Pontes/FAPESP, 2007, p. 361-374.

CYRINO, S.; DUARTE, M.; KATO, M. Visible subjects and invisible clitics in brazilian portuguese. In: KATO, M.; NEGRÃO, E. (Ed.). *Brazilian portuguese and the null subject*

parameter. Madri: Iberoamericana; Frankfurt am Main: Verveurt, 2000, p. 55-73.

DIMITRIADIS, A. The event structure of irreducibly symmetric reciprocals. In: DOLLING, J.; HEYDE-ZYBATOW, T. (Eds.) *Event structures in linguistic form and interpretation*. Berlin: De Gruyter, 2005.

DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português brasileiro. In: ROBERTS, I.; KATO, M. (Org.). *Português brasileiro: uma viagem diacrônica*. Campinas, São Paulo: Editora da UNICAMP, 1993, p. 107-128.

DUARTE, I.; OLIVEIRA, F. Referência nominal. In: MIRA MATEUS, Maria Helena. et al. *Gramática da língua portuguesa*. 5.ed. Lisboa: Caminho, 2003.

DUARTE, I. Relações gramaticais, esquemas relacionais e ordem de palavras. In: MIRA MATEUS, Maria Helena. et al. *Gramática da língua portuguesa*. 5.ed. Lisboa: Caminho, 2003.

FARACO & MOURA. *Gramática*. São Paulo: Ática, 1999.

FISHER, J. Social influences on the choice of a linguistic variant. *Word*, 1958, p. 47-56.

FIGUEIREDO, C. A comunidade de Sapé. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza. *O Português Afro-Brasileiro*. Salvador: EDUFBA, 2009.

FREYRE, Gilberto. *Casa-Grande & Senzala*. Rio de Janeiro: Nova Aguilar, 2002

GALVES, C. *Ensaio sobre a Gramática do Português*. Campins, São Paulo: UNICAMP, 2001.

GODOY, L. Os verbos recíprocos no PB e a hipótese da determinação semântico-lexical sobre a sintaxe. *Alfa*, v. 53, n. 1, p. 49-65, 2009.

GOODMAN, M. *A Comparative Study of Creole French Dialects*. The Hague: Mouton, 1964.

GUY, Gregory R. *Linguistic variation in Brazilian Portuguese: aspects of phonology, syntax and language history*. 1981a. Tese (Doutorado em Linguística) - University of Pennsylvania, Pennsylvania.

GUY, Gregory R. On the nature and origins of vernacular Brazilian Portuguese. In: *Estudios sobre Español de América y Lingüística Afroamericana*. Bogotá: Instituto Caro y Cuervo, 1989. p. 226-244.

GUY, Gregory R. A questão da crioulização no português do Brasil. In: ZILLES, Ana Maria Stahl (Org.). *Estudos de variação linguística no Brasil e no Cone Sul*. Porto Alegre: Editora da UFRGS, 2005. p.15-62.

GUY, G.; ZILLES, A. *Sociolinguística Quantitativa: instrumental de análise*. São Paulo: Parábola Editorial, 2007.

HASPELMATH, M. Coordinating constructions: An overview. In: HASPELMATH, M. (Ed.). *Coordinating constructions*. Typological Studies in Language 58. Amsterdam: Benjamins, 2004.

HEINE, Bernd. *Possession. Cognitive Sources, Forces and Grammaticalization*. Cambridge: Cambridge University Press, 1997.

HEINE, Bernd; KUTEVA, Tania. *World Lexicon of Grammaticalization*. Cambridge: CUP, 2002.

HOLM, J. Creole influence on Popular Brazilian Portuguese. In: GILBERT, G. G. (Ed.). *Pidgin and Creole languages. Essays in Memory of John E. Reinecke*. Honolulu: University of Hawaii Press, 1987. p.406-429.

HUTTAR, G. Sources of creole semantic structures. *Language* 51. 1971, p. 684-695.

IGNÁCIO, S. E. O caso comitativo. *Revista CON(TEXTOS) Linguísticos*, v. 3, n. 3, p. 11-25, 2009.

ILARI, R. Dos problemas de imperfeita simetria. *Cadernos de estudos lingüísticos*, v. 13, p. 49-65, 1987.

KATO, M.A. O Português são dois...ou três? In: LOBO, T. et. al. (Org.). *ROSAE -Linguística histórica, história das línguas e outras histórias*. Salvador: EDUFBA, 2012, p. 114-131.

KAYNE, R. *The antisymmetry of syntax*. Cambridge: MIT Press, 1994.

LABOV, William. *The social stratification of English in New York*. Washington, D. C.: Center for Applied Linguistics, 1966.

LABOV, William. *Principles of linguistic change*. Oxford; Cambridge: Blackwell, 1994.
LABOV, William. *Padrões Sociolinguísticos*. Tradução de Marcos Bagno, Marta Scherre e Caroline Cardoso. São Paulo: Parábola, 2008 [1972].

LAKOFF, G.; PETERS, S. (1969) Phrasal conjunction and symmetric predicates. In: REIBEL, D. A.; SCHANE, S. A. (Eds.). *Modern studies in English*. New Jersey: Prentice-Hall, 1969. p. 113-142.

LEFEBVRE, C. *Creole genesis and the acquisition of grammar: The Case of Haitian creole*. Cambridge Studies in Linguistics 88. New York: Cambridge University Press, 1998.

LEFEBVRE, C. Relexification in creole genesis and effects on the development of the creole. In: SMITH, N.; VEENSTRA, T. (eds). *Creolization and Contact*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2001.

LEFEBVRE, C. On the principled nature of the respective contributions of substrate and superstrate languages to a creoles' lexicon. In: MICHAELIS, S. (ed.). *Roots of Creole Structures: Weighing the contribution of substrates and superstrates*. 2008, p. 197-223.

LEFEBVRE, C.; LUMSDEN, J. The central role of Relexification in Creole Genesis: The Case of Haitian Creole. In: LEFEBVRE, C.; LUMSDEN, J. (orgs.). *La genèse du créole haïtien: un cas particulier d'investigation sur la forma de la grammaire universelle*. UQAM, 1994.

LINS, A. B. Três hipóteses e alguns caminhos para melhor compreender o processo constitutivo do português brasileiro. In: OLIVEIRA, K. et. al. (Orgs). *Do português arcaico ao português brasileiro: outras histórias*. Salvador: EDUFBA, 2009.

LUCCHESI, Dante. Variação e norma: elementos para uma caracterização sociolinguística do português do Brasil. *Revista Internacional de Língua Portuguesa*, Lisboa, n.12, p.17-28, 1994.

LUCCHESI, Dante. A constituição histórica do português brasileiro como um processo bipolarizado: tendências atuais de mudança nas normas culta e popular. In: GROßE, S.; ZIMMERMANN, K. (Ed.). *"Substandard" e mudança no português do Brasil*. Frankfurt am Main: TFM, 1998. p.73-100.

LUCCHESI, D. *A variação na concordância de gênero em uma comunidade de fala afro-brasileira: novos elementos sobre a formação do português popular do Brasil*. 2000a. Tese (Doutorado em Linguística) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

LUCCHESI, Dante. As duas grandes vertentes da história sociolinguística do Brasil. *DELTA*, São Paulo, v.17, n.1, p.97-130, 2001a.

LUCCHESI, Dante. Norma linguística e realidade social. In: BAGNO, Marcos (Org.). *Linguística da norma*. São Paulo: Loyola, 2002b. p. 63-92.

LUCCHESI, Dante. O conceito de transmissão linguística irregular e o processo de formação do português do Brasil. In: RONCARATI, C.; ABRAÇADO, J. (Org.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7Letras, 2003b. p.272-284.

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

LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza. *O Português Afro-Brasileiro*. Salvador: EDUFBA, 2009.

LYONS, J. Referência, sentido e denotação. In: _____. *Semântica*. Trad. Wanda Ramos. Lisboa: Martins Fontes, 1977, 145-87.

MATOS, Gabriela. Estruturas de coordenação. In: MATEUS, Maria Helena Mira, et. al. *Gramática da língua portuguesa*. 5.ed. Lisboa: Caminho, 2003.

MATTOS E SILVA, Rosa Virgínia. De fontes sócio-históricas para a história social linguística do Brasil: em busca de indícios. In: MATTOS E SILVA, Rosa Virgínia. *Para a história do português brasileiro*. v. 2, t. 2, São Paulo: Humanitas, 2001, p. 275-301.

MATTOS E SILVA, Rosa Virgínia. *Ensaio para uma sócio-história do português brasileiro*. São Paulo, SP: Parábola, 2004.

MATTOS E SILVA, Rosa Virgínia. Uma compreensão histórica do português brasileiro: velhos problemas repensados. In: CARDOSO, Suzana Alice; MOTTA, 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.

MATTOSO, Katia. *Ser escravo no Brasil*. 3.ed. São Paulo: Brasiliense, 2003.

MICHAELIS, Susanne Maria, et. al. (eds.). *The Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Instituto Max Planck de Antropologia Evolutiva. Disponível em: <<http://apics-online.info/>>.

MICHAELIS, Susanne Maria, et. al. (eds.). *The Atlas of Pidgin and Creole Language Structures*. United Kingdom: Oxford, 2013.

MIOTO, C.; SILVA, M. C. F.; LOPES, R.E.V. 3. ed. *Manual de Sintaxe*. Florianópolis: Insular, 2007.

MIRA MATEUS, Maria Helena. et al. *Gramática da língua portuguesa*. 5.ed. Lisboa: Caminho, 2003.

MOLLICA, M.; BRAGA, M. (Orgs). *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003.

MOLLICA, M.; PAIVA, M.; PINTO, I. Relação entre [l] → [r] r [r] → [0] em grupos consonantais em português. In: *Relatório final do Projeto Mecanismos Funcionais do Uso Linguístico*. Rio de Janeiro: UFRJ, 1989.

MOURA NEVES, Maria Helena de. *Gramática de usos do português*. São Paulo: UNESP, 2000.

MUFWENE, S. The founder principle in creole genesis. *Diachronica*, v. 13, n. 1, 1996, p. 83-134.

MÜHLHÄUSLER, Peter. *Pidgin & Creole Linguistics*. Londres: Basil Blackwell, 1986.

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

MUYSKEN, P. Half-way between Quechua and Spanish: The case for relexification. HIGHFIELD, A.; VALDMAN, A. (eds.) *Historicity and variation in creole studies*. Ann Arbor: Karoma, 1981a, p. 52-79.

MUYSKEN, P.; VEENSTRA, T. Universalist approaches. In: ARENDS, J.; MUYSKEN, P.; SMITH, N.(eds). *Pidgins and Creoles: an introduction*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1994.

NARO, Anthony Julius; SCHERRE, Marta Maria Pereira. O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In: RONCARATI, C.; ABRAÇADO, J. (Org.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7Letras, 2003, p. 285-302.

NARO, Anthony Julius; SCHERRE, Marta Maria Pereira. *Origens do Português Brasileiro*. São Paulo: Parábola, 2007.

PAIVA, M. A variável gênero/sexo. In: MOLLICA, M.; BRAGA, M. (Orgs). *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003.

PARKVALL, M. Da África para o Atlântico. Campinas, São Paulo: Editora da Unicamp, 2012.

PETTER, M. Línguas africanas no Brasil. In: CARDOSO, Suzana Alice; MOTTA, 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.

ROBERTS, Ian. *Diacronic syntax*. Oxford: Oxford University Press, 2007.

ROCHA LIMA, Carlos Henrique da. *Gramática Normativa da língua portuguesa*. 45. ed. Rio de Janeiro: José Olympio, 2006.

ROCHA LIMA, C. H. *Gramática normativa da língua portuguesa*. Rio de Janeiro: José Olympio, 1973.

RODRIGUES, A. D. Línguas Indígenas: 500 anos de descobertas e perdas. *DELTA*, v. 9, n.1, p. 83-103, 1993a.

RODRIGUES, Aryon. 1996. As línguas gerais sul-americanas. *Papia*, 4(2), p. 6-18. Acesso em: 22.01.14 Disponível em: <<http://www.etnolinguistica.org/artigo:rodrigues-1996>>.

RODRIGUES, A. As outras línguas da colonização do Brasil. In: CARDOSO, Suzana Alice; MOTTA, 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.

SAPIR, E. *Language – an introduction to the study of speech*. New York: Harcourt, Brace and World, 1949 [1921].

SCHERRE, M. M. Sobre a influência das variáveis sociais na concordância nominal. In: OLIVEIRA E SILVA, G. M.; SCHERRE, M. M. *Padrões sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro, 1996, p. 239-264.

SILVA, J. A. A comunidade de Cinzento. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza. *O Português Afro-Brasileiro*. Salvador: EDUFBA, 2009.

SINGLER, J. Nativization and pidgin/creole genesis: a reply to Bickerton. *Journal of Pidgin and Creole Languages* 7, 1992, p. 319-333.

STASSEN, Leon. AND-languages and WITH-languages. *Linguistic Typology*, v. 4, n. 1, p. 1-54, 2000.

STASSEN, Leon. Noun Phrase Conjunction. In: Haspelmath, M. (eds). *The World Atlas of Language Structures Online*. Leipzig: Instituto Max Planck de Antropologia Evolutiva. Disponível em: <<http://wals.info/feature/63A#2/25.5/153.5>>

STASSEN, Leon. *Predicative Possession*. Oxford: Oxford University Press, 2009.

STEWART, W. Creole languages in the Caribbean. In: RICE, F. (Org). *Study of the role of second languages*. Washington D. C.: Center of Applied Linguistics of the Modern Language Association of America, 1962, p. 34-53.

STOLZ, Thomas; STROH, Cornelia; URDZE, Aina. *On Comitatives and related categories: a typological study with special focus on the languages of Europe*. Berlin/New York: Mouton de Gruyter, 2006.

TARALLO, Fernando. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In: ROBERTS, Ian; KATO, Mary (orgs). *Português Brasileiro: uma viagem diacrônica*. 2. ed. Campinas: Editora da UNICAMP, 1993.

TARALLO, Fernando. *A pesquisa sociolinguística*. 8 ed. São Paulo: Ática, 2007.

TAYLOR, D. Language contact in the West Indies. *Word* 12,3, 1956.

TOMPSON, R. A note on some possible affinities between creole dialects of the Old World and those of the New. In: LE PAGE, R. (Org). *Proceeding of the conference on creole language studies*. Londres: MacMillan, 1961.

VOORHOEVE, J. Historical and linguistic evidence in favor of the relexification theory in the formation of creoles. *Language in society* 2.1973, p. 133-145.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. Trad. Marcos Bagno, revisão técnica Carlos Alberto Faraco. São Paulo: Parábola, 2006 [1968].

WHINNOM, K. *Spanish contact vernaculars in the Philippine Islands*. Hong Kong: Hong Kong University Press, 1956.

ANEXOS

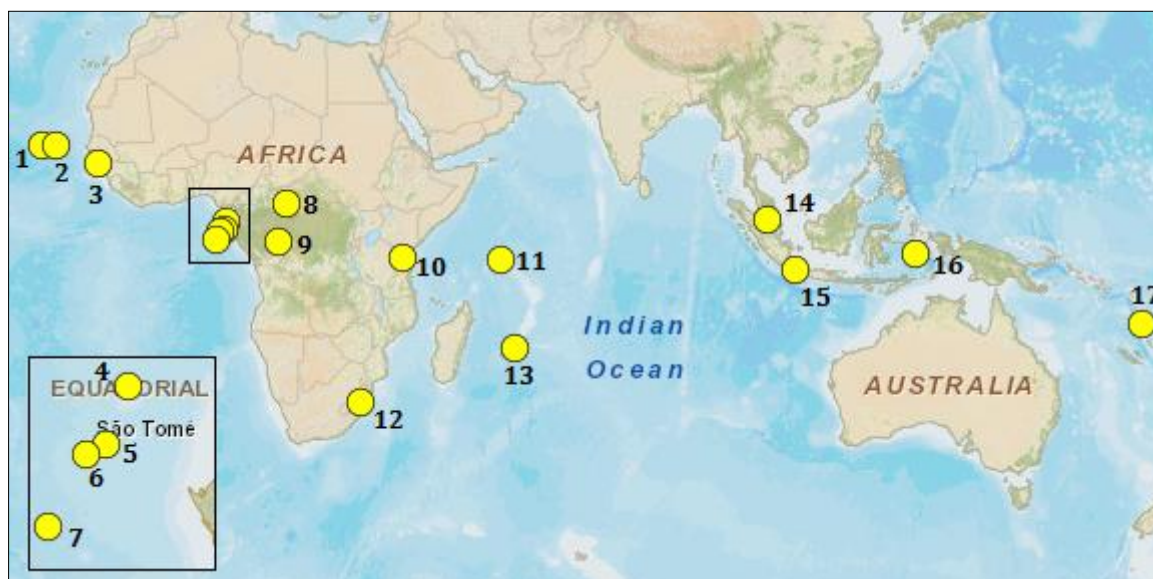
Anexo A: Línguas com o valor *identidade* de formas na América (APiCS)

Fonte: WALs On-line. Disponível em: <<http://wals.info/feature/63A#2/16.6/153.5>>

LEGENDA

- (1) Palenquero
- (2) Guadeloupean
- (3) Martinica
- (4) Berbice
- (5) Sranan
- (6) Saramaccan
- (7) Nengee
- (8) Guyanais

Anexo B: Línguas com o valor *identidade* de formas nos demais continentes (APiCS)

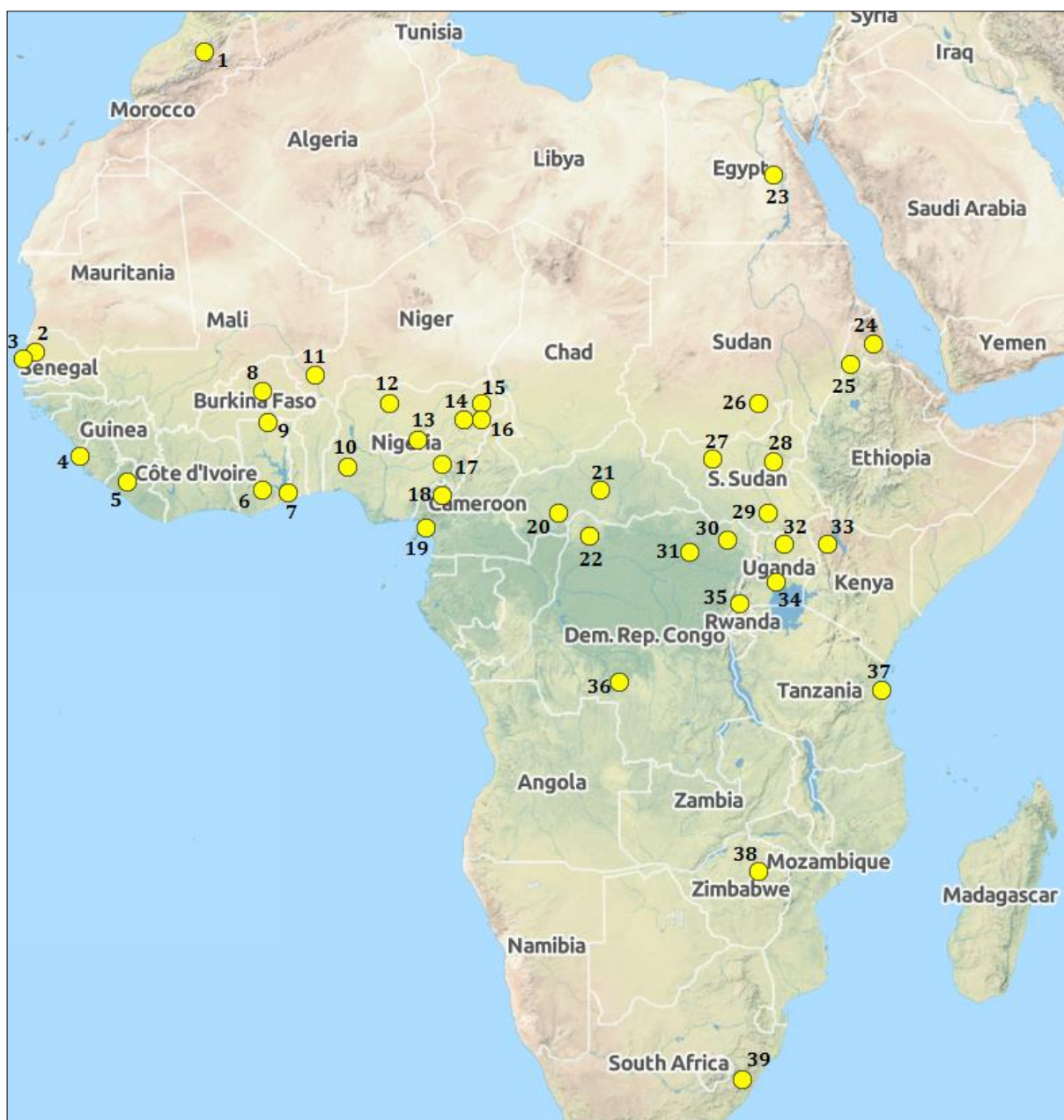


Fonte: WALS On-line. Disponível em: <<http://wals.info/feature/63A#2/16.6/153.5>>

LEGENDA

- | | | |
|--------------------------------------|-------------------------|-----------------------|
| (1) Crioulo Caboverdiano de Brava | (7) Fa d'ambo | (13) Maurícia crioulo |
| (2) Crioulo Caboverdiano de Santiago | (8) Sango | (14) Papia Kristang |
| (3) Guiné-Bissau Kriyol | (9) Lingala | (15) Batavia crioulo |
| (4) Princiense | (10) Mbugu | (16) Ambon Malay |
| (5) Santome | (11) Seychelles crioulo | (17) Bislamá |
| (6) Angolar | (12) Fanakalo | |

Anexo C: Línguas com o valor *identidade* de formas na África (WALS)



Fonte: WALS On-line. Disponível em: <<http://wals.info/feature/63A#2/16.6/153.5>>

LEGENDA

(1) Berber	(11) Zarma	(21) Linda	(31) Mangbetu
(2) Wolof	(12) Hausa	(22) Ngbaka	(32) Acholi
(3) Palor	(13) Birom	(23) Coptic	(33) Turkana
(4) Temne	(14) Tera	(24) Bilin	(34) Luganda
(5) Kpelle	(15) Kanuri	(25) Kunama	(35) Knore-Kiga
(6) Akan	(16) Margi	(26) Nubian	(36) CiLuba
(7) Ewe	(17) Fula	(27) Dinka	(37) Swahili
(8) Mooré	(18) Babungo	(28) Nuer	(38) Shona
(9) Gurenne	(19) Duala	(29) Bari	(39) Zulu
(10) Yoruba	(20) Sango	(30) Mamvu	

Anexo D: Línguas indígenas com o valor *identidade* de formas no Brasil (WALS)

Fonte: WALS On-line. Disponível em: <<http://wals.info/feature/63A#2/16.6/153.5>>

LEGENDA

- (1) Apalaí
- (2) Urubú-Kaapor
- (3) Canela-Krahô
- (4) Tupi
- (5) Kaingang