



UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA– UEFS
Departamento de Letras e Artes
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS
MESTRADO EM ESTUDOS LINGUÍSTICOS – MEL

DAYANE MOREIRA LEMOS

**PORTUGUÊS BRASILEIRO E PORTUGUÊS ANGOLANO:
VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO**

Feira de Santana, BA

2014

DAYANE MOREIRA LEMOS

**PORTUGUÊS BRASILEIRO E PORTUGUÊS ANGOLANO:
VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO**

Dissertação apresentada ao Programa de Pós-graduação em Estudos Linguísticos, da Universidade Estadual de Feira de Santana, como requisito parcial para obtenção do título de Mestre em Estudos Linguísticos.

Orientadora: Prof^a. Dr^a. Eliana Sandra Pitombo Teixeira

Feira de Santana, BA

2014

DAYANE MOREIRA LEMOS

**PORTUGUÊS BRASILEIRO E PORTUGUÊS ANGOLANO:
VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO**

Dissertação apresentada ao Programa de Pós-graduação em Estudos Linguísticos, da Universidade Estadual de Feira de Santana, como requisito parcial para obtenção do título de Mestre em Estudos Linguísticos.

Aprovada em 17 de março de 2014.

Prof^a. Doutora Eliana Sandra Pitombo Teixeira
Orientadora - UEFS

Prof^o. Doutor Alan Norman Baxter
UFBA

Prof^a. Doutora Norma Lucia Fernandes de Almeida
UEFS

A Deus pela sabedoria e discernimento para elaboração deste trabalho; Aos meus pais, Domingos e Lourdes, meu namorado, Rômulo Miranda, familiares e amigos que diretamente contribuíram na finalização de mais uma etapa.

AGRADECIMENTOS

Agradeço primeiramente à Deus pela vida, sabedoria e paciência a me concedida.

Aos meus QUERIDOS pais, minhas pedras preciosas, orgulho da minha vida, agradeço pela base, incentivo e exemplo de amor.

A Rômulo Miranda, amor meu, pelo companheirismo, amor e carinho, que nunca me faltaram nestes 2 anos. Por ter sempre um abraço amigo nos momentos de desespero, em que tudo parecia se perder.

Aos meus familiares, em especial meu irmão Danilo Lemos e minha prima Nilvia Lemos, que realizaram minha inscrição neste Programa de pós-graduação e hoje acompanham a finalização de mais uma etapa. Obrigada pela torcida!

Aos meus antigos e novos mestres, que sabiamente me conduziram ao mundo do conhecimento científico. Em especial, a Zenaide Carneiro, que em meio às turbulências sempre esteve ao meu lado, me acompanhando sempre com muito carinho. Obrigada pelas palavras de conforto e amizade.

A Eliana Pitombo que aceitou ser minha orientadora e sempre esteve disposta a auxiliar nessa árdua jornada de finalização de dissertação.

A minha banca de defesa, Alan Baxter e Norma Lucia, pelo saber científico e pelas indicações que engrandeceram meu trabalho.

À CAPES, pelo incentivo a pesquisa através da bolsa a me concedida.

As minhas amigas Shirley Guedes e Tércia Dória, por deixarem esta caminhada mais leve. Sem vocês tudo seria mais difícil! Valéria, Nilce e Quelle pela companhia agradável e incentivos.

À família e amigos pelo apoio e confiança.

A todos, peço desculpas pelas infinitas ausências, muitas vezes incompreendidas... a vontade de está pertinho de vocês era enorme.

A todos que, direta ou indiretamente, contribuirão para realização de mais uma etapa.

“As palavras são finitas pra dizer o infinito”

(Antônio Cândido).

“Tudo que é seu, encontrará uma maneira de chegar até você.”

Chico Xavier

RESUMO

Este trabalho busca descrever e analisar dados do processo de variação da concordância de número no sintagma nominal, nas variedades do português popular de Mato Grosso, amostra representativa do Português Brasileiro (PB), em comparação ao português popular falado em Luanda, amostra representativa do Português Angolano (PA), na tentativa de identificar convergências/divergências entre essas duas variedades do português, tendo como base de análise os padrões canônicos estabelecidos pelo português europeu. Consideramos o estudo sob uma perspectiva sócio-histórica, supondo que as características atuais da fala popular resultam de mudanças produzidas no contexto de contato do português com outras línguas. O material analisado foi extraído dos *corpora* dos projetos *A língua portuguesa no semiárido baiano* e do projeto *Em busca das raízes do português brasileiro*, respectivamente, amostras do Português popular do Brasil e do Português popular de Luanda, o que se justifica pela proposta, assumida neste trabalho, de que, embora as variedades cultas e populares do PB e do PA se oponham ao Português Europeu (PE), o português popular tem sua origem no processo de transmissão linguística irregular, diferente do que ocorreu, a princípio, com a variante culta. As amostras de língua oral são de informantes analfabetos e de baixo nível de escolaridade (até a quarta série) das comunidades de Mato Grosso e Luanda. Dessa forma, buscamos verificar se o comportamento linguístico dos falantes que compõem as amostras apresentam semelhanças e/ou divergências, no que se refere ao fenômeno gramatical. Os resultados da análise mostraram simplificação da concordância nominal de número nas duas variedades estudadas com índices muito semelhantes de não concordância.

Palavras-chave: Português popular. Contato linguístico. Concordância nominal de número.

ABSTRACT

This study aims to describe and analyze data on number agreement variation in nominal phrase in popular Portuguese variety used in Mato Grosso – representing Brazilian Portuguese (BP) – compared to Luanda’s – representing Angolan Portuguese (AP), searching to find any convergences/deviations between both varieties. The comparison takes the canonical patterns of European Portuguese (EP) as its base. We have undertaken the study from a socio-historical viewpoint, assuming that the characteristics of popular speech of today are the result of changes brought about in the context of contact of Portuguese with other languages. The material analyzed was extracted from the *corpora* of the projects *A Língua Portuguesa no Semiárido Baiano* and *Em busca das Raízes do Português Brasileiro*, respectively, Brazilian and Luanda’s popular speech samples. The choice of popular speech is justified by the approach we have adopted in this study that although popular and standard varieties of BP and AP are different from that of EP, popular BP and AP originate from irregular linguistic transmission process, differently from what has occurred with standard Portuguese. The oral language samples are from illiterate people as well as from people with only four years of schooling from Mato Grosso and Luanda. Therefore, we have aimed to verify if the linguistic performance of the informants in the *corpora* presents similarities and/or differences related to this grammatical phenomenon. The results of the analysis show agreement simplification in both varieties with very similar rates of non-agreement.

Keywords: Brazilian Portuguese. Language contact. Nominal number agreement.

LISTA DE ILUSTRAÇÕES

FIGURA 1 - Demografia histórica do Brasil	24
FIGURA 2 - Organização social de Angola no período colonial	34
FIGURA 3 - Distribuição de Greenberg: troncos linguísticos africanos	36
FIGURA 4 - Panorama dos estudos sobre variação na concordância nominal desenvolvidos no Brasil	64
FIGURA 5 - Panorama dos estudos sobre variação na concordância nominal desenvolvidos na África	68
FIGURA 6 - Mapa da região semiárida da Bahia	87
FIGURA 7 - Mapa territorial de Angola	93
FIGURA 8 - Frequência da variação na concordância no SN em amostras representativas do PB e PA	102
FIGURA 9 - Efeito da variável gênero na variação na concordância de número nos SNs do PB	113
FIGURA 10 - Efeito da variável faixa etária na variação na concordância de número nos SNs do PB	116
FIGURA 11 - Efeito das variáveis Gênero e Faixa etária, associadas, na concordância ..	118
FIGURA 12 - Efeito da variável gênero na variação na concordância de número nos SNs do PA	120
FIGURA 13 - Efeito da variável do número de constituintes flexionáveis na variação na concordância de número nos SNs no PB	123

LISTA DE QUADROS

QUADRO 1 - Línguas subsaarianas	22
QUADRO 2 - Descrição dos volumes da <i>Coleção Amostras da Língua Falada no Semiárido Baiano</i>	88
QUADRO 3 - Descrição da população amostra do PB	91
QUADRO 4 - Descrição da população amostra do PA	94
QUADRO 5 - Descrição da população amostra	96

LISTA DE TABELAS

TABELA 1-	Estimativas de população no Brasil	22
TABELA 2-	Estimativas de imigração para o Brasil	23
TABELA 3-	População de falantes das diversas línguas africanas que chegaram ao Brasil	31
TABELA 4-	População escrava	32
TABELA 5-	Taxas de uso da concordância no SN em amostras representativas do PB e PA	102
TABELA 6-	Taxas de uso da concordância no SN em função do gênero no PB	112
TABELA 7-	Dados comparativos: uso da concordância no SN em função do gênero no PB	114
TABELA 8-	Taxas de uso da concordância no SN em função da faixa etária no PB	114
TABELA 9-	Dados comparativos: uso da concordância no SN em função da faixa etária no PB	115
TABELA 10-	Taxas de uso da concordância no SN em função do gênero e da faixa etária no PB	117
TABELA 11-	Taxas de uso da concordância no SN em função do gênero no PA	119
TABELA 12-	Dados comparativos: uso da concordância no SN em função do gênero	119
TABELA 13-	Taxas de uso da concordância no SN em função da origem no PA	121
TABELA 14-	Taxas de uso da concordância no SN em função do número de constituintes flexionáveis no PB	123
TABELA 15-	Dados comparativos: uso da concordância no SN em função do número de constituintes flexionáveis	124
TABELA 16-	Taxas de uso da concordância no SN em função da saliência fônica no PB	125
TABELA 17-	Dados comparativos: uso da concordância no SN em função da saliência fônica no PB	126

SUMÁRIO

CONSIDERAÇÕES INICIAIS	13
1 LÍNGUAS E SOCIEDADE: O PORTUGUÊS E SUAS VARIEDADES	17
1.1 PORTUGUÊS DO BRASIL: ASPECTOS SÓCIO-HISTÓRICOS	19
1.1.1 População africana no Brasil	29
1.2 PORTUGUÊS DE ANGOLA: ASPECTOS SÓCIO-HISTÓRICOS	32
1.2.1 Angola: situação linguística	35
1.2.1.1 Línguas Banto em Angola	37
2 ORIGEM DO PORTUGUÊS BRASILEIRO: BREVES DISCUSSÕES	40
2.1 DERIVA SECULAR DAS LÍNGUAS ROMÂNICAS	41
2.2 CONTATO LINGUÍSTICO	44
2.2.1 Conceituando pidgins e crioulos	46
2.2.2 Transmissão linguística irregular	51
3 AS DIVERSAS ABORDAGENS DA CONCORDÂNCIA NOMINAL DE NÚMERO	55
3.1 CONCORDÂNCIA NOMINAL: PERSPECTIVA TRADICIONAL E FORMAL	59
3.2 CONCORDÂNCIA NOMINAL DE NÚMERO VARIÁVEL NO PORTUGUÊS DO BRASIL E DE ANGOLA	62
3.2.1 Concordância nominal de número variável no português do Brasil	62
3.2.2 Concordância nominal de número variável no português de Angola	67
3.4 A VARIAÇÃO NA CONCORDÂNCIA NOMINAL E A POLÊMICA SOBRE A FORMAÇÃO DO PORTUGUÊS DO BRASIL	70
4 PRESSUPOSTOS TEÓRICOS METODOLÓGICOS	73
4.1 CAMINHOS DA SOCIOLINGUÍSTICA: TEORIA E MÉTODO	74
4.1.1 A Sociolinguística quantitativa	78

4.2 TEORIA FUNCIONALISTA	83
4.3 TEORIA DOS 4M	83
4.4 ABORDAGEM METODOLÓGICA	86
4.4.1 Amostragem	86
4.4.1.1 Amostras da comunidade de fala de Mato Grosso	86
4.4.1.2 Amostras da comunidade de fala de Luanda	92
4.4.2 Coleta e análise de dados	95
5 A CONCORDÂNCIA NOMINAL DE NÚMERO SOBRE O OLHAR DA SOCIOLINGUÍSTICA QUANTITATIVA	100
5.1 ANÁLISE DOS DADOS	101
5.2 VARIÁVEIS SOCIAIS	111
5.2.1 Português brasileiro: variáveis sociais	112
5.2.1.1 Gênero	112
5.2.1.2 Faixa etária	114
5.2.1.2 Gênero <i>versus</i> Faixa etária	117
5.2.2 Português angolano: variáveis sociais	118
5.2.2.1 Gênero	118
5.2.2.2 Origem	121
5.3 VARIÁVEIS LINGUÍSTICAS	122
5.3.1 Português brasileiro: variáveis linguísticas	122
5.3.1.1 Número de constituintes flexionáveis do SN	122
5.3.1.2 Saliência fônica	124
CONSIDERAÇÕES FINAIS	128
REFERÊNCIAS	131

CONSIDERAÇÕES INICIAIS

Nesta dissertação, analisamos a concordância nominal no panorama da variedade do Português Brasileiro (doravante PB) em comparação a variedade do Português Angolano (doravante PA). Entre os diversos aspectos da variação no PB, talvez a concordância de número nominal seja um dos que mais chama a atenção de qualquer observador, mesmo não muito preocupado com a linguagem, nem profundo conhecedor da gramática normativa.

Refletindo sobre as questões inerentes ao processo de variação nos sintagmas nominais, temos como pressuposto o fato de que o amplo quadro de variação existente no sistema de concordância, nas normas populares do PB, seria decorrente da história sociolinguística brasileira, marcada pelo contato entre línguas, com processos de transmissão linguística irregular do português, impulsionados pela população africana e indígena, principalmente pela primeira, já que esta esteve na base da demografia do Brasil durante mais de três séculos.

Partindo do pressuposto segundo o qual “[...] a difusão do português brasileiro pela nossa população de origem africana envolve [...] uma das questões centrais na compreensão histórica do português brasileiro” (MATTOS E SILVA, 2001a, p. 276), vimos a importância de trazer dados do PA para que pudessem servir de base comparativa, na identificação de convergências/divergências entre essas duas variedades do português, tendo como base de análise os padrões canônicos estabelecidos pelo português europeu.

Como poderão observar no *Capítulo 3* desta dissertação, no âmbito do PB já se realizaram diversos estudos sobre a variação na concordância, que permitem delinear, com razoável segurança, os fatores estruturais condicionantes da marca de número plural no SN. O mesmo, no entanto, não se pode dizer em relação a outras variedades do português, a especificar o PA. No que concerne ao PB, a importância atribuída à variável decorre do fato de a não marcação canônica de plural no SN não só se revestir de forte cunho estigmatizante, mas também fornecer elementos para os debates sobre a constituição dessa variedade e sobre os aspectos que a diferenciam do Português Europeu (doravante PE). No que se relaciona a algumas das variedades africanas, temos uma grande lacuna a ser preenchida, mas estudos já indicam possíveis convergências com os padrões variáveis observados no PB.

Compreendendo a importância de um trabalho que busque descrever e analisar dados do processo de variação da concordância de número no sintagma nominal, nas variedades populares do PB e PA, tenta-se identificar convergências/divergências entre essas duas variedades do português, tendo como base de análise os padrões canônicos estabelecidos pelo português europeu.

Dessa forma, analisamos a variável linguística constituída da presença e da ausência de marca de flexão de plural entre os elementos do SN, através do material extraído dos *corpora* dos projetos *A língua portuguesa no semiárido baiano* e do projeto *Em busca das raízes do português brasileiro*, respectivamente, amostras do Português popular do Brasil (doravante PPB) e do Português popular de Angola (doravante PPA). Quanto ao fato de as amostras serem voltadas ao português popular, justifica-se pela proposta, assumida nesta dissertação, de que, embora as variedades cultas e populares do PB e do PA se oponham ao PE, o português popular tem uma origem no processo de transmissão linguística irregular, diferente do que ocorreu, a princípio, com a variante culta, o que pode apresentar fatos linguísticos de maior distanciamento do PE.

Para cumprimento do objetivo dessa dissertação tomamos com base teórica os pressupostos teóricos metodológicos da Teoria da Variação e Mudança Linguística (WEINREICH, LABOV, HERZOG, 2006; LABOV, 2008), em que a língua é assumida como um sistema heterogêneo e, portanto, sujeito a variação, razão pela qual não pode ser analisada isoladamente, sem se levar em conta o contexto social no qual se processa. Também trazemos em evidência os aspectos da hipótese funcionalista de Kiparsky (1972), numa releitura de Scherre (1988), em que se acredita que uma teoria linguística adequada que “[...] pretende ter caráter explanatório deve incorporar também motivações ou princípios gerais externos, ou seja, princípios não estritamente gramaticais determinando o uso e, conseqüentemente, a forma da língua” (SCHERRE, 1988, p. 17). Ainda sobre aporte teórico, trazemos à cena a Teoria dos 4M (quatro tipos de morfemas) de Myers-Scotton e Jake (2000), numa releitura de Lopes (2001), acreditando que a variação da concordância nominal de número entre os elementos do SN, no português do Brasil, tem a ver com o tipo de aquisição por que passaram negros e índios e seus descendentes.

Dessa forma, pretendemos apontar os fatores que podem motivar a presença das marcas de plural nos sintagmas nominais, além de colaborar para a apreensão de convergências/divergências entre as duas variedades e, em última instância, para os debates sobre as origens do Português do Brasil (LUCCHESI, 2009, 2008, 2003, 1998 e NARO e SCHERRE 2007).

Visando dar a este trabalho a melhor organização possível, o mesmo apresenta-se dividido da seguinte forma:

- (i) Capítulo intitulado *Línguas e sociedade: o português e suas variedades*, em que buscamos compreender as motivações sócio-históricas que deram origem a cada uma das variedades em análise, português brasileiro e angolano, para a partir de então,

poder traçar os possíveis paralelismos entre essas duas variedades do português europeu. Para tanto, foi feita uma descrição da configuração sócio-histórica do Brasil e de Angola, iniciando pelos primeiros anos de colonização europeia, pelas trocas estabelecidas entre colonizados e colonizadores, a escravização, até aos dias atuais, em que se questiona as origens do português brasileiro e angolano;

- (ii) Capítulo intitulado *Origem do português brasileiro: breves discussões*, em que discutimos as explicações para a origem do Português Brasileiro. Para tanto, discute-se sobre os supostos pidgins e crioulos, e de forma mais consistente aborda-se as duas hipóteses para as origens do português: Deriva Secular e Transmissão Linguística Irregular;
- (iii) Capítulo intitulado *As diversas abordagens da concordância nominal de número*, em que discutimos sobre o fenômeno em análise, *a variação na concordância de número no sintagma nominal*, com base nas abordagens formalista e funcionalista. De forma concomitante, faz-se um levantamento de trabalhos já realizados na área a fim de atestar resultados referentes ao fenômeno, que possam contribuir para as análises a serem realizadas no capítulo V desta dissertação;
- (iv) Capítulo intitulado *Pressupostos teóricos metodológicos*, apresentamos os aportes teóricos e metodológicos que subsidiarão as posteriores análises da concordância gramatical de número entre os elementos flexionáveis do sintagma nominal no PB e PA. Por tanto, adota-se uma perspectiva teórica que une aspectos da Teoria da Variação Linguística Laboviana ou Sociolinguística Quantitativa e aspectos da Teoria Funcionalista. Neste capítulo também tecemos considerações acerca da Teoria dos 4M, de Myers-Scotton e Jake (2000 *apud* LOPES, 2001) e apresentamos a abordagem metodológica utilizada no tratamento e análise dos dados. Para tanto, delinearemos a área em que a pesquisa se insere e modelo quantitativo que a sustenta;
- (v) Capítulo intitulado *A concordância nominal de número sobre o olhar da Sociolinguística quantitativa* procederemos à análise do objeto de estudo, consoante os pressupostos teórico-metodológicos já citados, considerando variáveis preestabelecidas para esta investigação, buscando ainda dialogar com outras pesquisas que trataram da temática da concordância;
- (vi) Por fim, apresentaremos nossas *Considerações finais*, em que mostramos nossa visão sobre o fenômeno gramatical em estudo, fazendo inferências acerca dos resultados

obtidos ao comparar variedades do português e sua relevância no processo de contato linguístico.

Com afirma Brandão (2011, p. 164), as questões relativas à perda de morfologia flexional e de regras de concordância constituem importantes parâmetros para a definição do *status* de variedades que emergem de contato entre populações distintas, dos pontos de vista linguístico e cultural. Nesse sentido, estudos sobre a concordância nominal têm servido de base para a formulação de diferentes interpretações sobre o surgimento e desenvolvimento de variedades do português.

No entanto, apesar dos muitos estudos comparativos do português brasileiro com o português europeu, só muito recentemente os linguistas brasileiros têm-se preocupado em comparar a variedade brasileira com as das ex-colônias portuguesas na África. Portanto, esta dissertação assume essa “nova” proposta de análise comparativa, no campo da variação morfofonológica, e apresenta-se como fruto do projeto *Em busca das raízes do português brasileiro*, concebido e em desenvolvimento por minha orientadora Eliana Pitombo.

1 LÍNGUAS E SOCIEDADE: O PORTUGUÊS E SUAS VARIEDADES

Última flor do Lácio, inculta e bela,
És, a um tempo, esplendor e sepultura:
Ouro nativo, que na ganga impura
A bruta mina entre os cascalhos vela...

Amo-te assim, desconhecida e obscura.
Tuba de alto clangor, lira singela,
Que tens o trom e o silvo da procela,
E o arrollo da saudade e da ternura!

Amo o teu viço agreste e o teu aroma
De virgens selvas e de oceano largo!
Amo-te, ó rude e doloroso idioma,

Em que da voz materna ouvi: "meu filho!",
E em que Camões chorou, no exílio amargo,
O gênio sem ventura e o amor sem brilho!

Língua Portuguesa – Olavo Bilac

Para iniciarmos as reflexões sobre o português, em sua modalidade brasileira e angolana, é preciso perpassar, ainda que superficialmente, pela formação da língua portuguesa, considerando um pouco sobre a própria história de Portugal, e os impactos que momentos históricos causaram na formação do idioma oficial falado no Brasil e em Angola. Como sempre dizia a eminente linguista Rosa Virgínia Mattos e Silva, os aspectos sócio-históricos são fundamentais para compreensão de uma língua, e acrescento, principalmente ao trata-se do português brasileiro e do angolano.

Em meados do século XII a Europa, nascida das ruínas do Império Romano e da presença dos chamados povos bárbaros, começa pouco a pouco a se modificar, pela expansão da agricultura e do comércio (FAUSTO, 1996). Após esse período, se inicia, no século XIII, as muitas batalhas, as quais definiam as fronteiras da Europa e faziam surgir o Estado com uma organização política centralizada. Faraco (2008) ainda afirma que na metade do século XIII, Portugal tinha já suas fronteiras atuais claramente definidas.

Com a formação do Estado, em 1217, o rei D. Dinis I, instaura o Português como a língua oficial de Portugal. A partir de meados do século XV, o português – na esteira da expansão marítima de Portugal – se tornou uma língua internacional, com seus falantes estabelecendo contato ao longo da costa do continente africano, alcançando a Índia em 1498, a América em 1500, a China por volta de 1515 e o Japão em 1543 (FARACO, 2008, p. 14), o

que confirma que a nação portuguesa foi protagonista de uma longa história de contato com povos que não dominavam a sua língua.

A trajetória histórica de Portugal fez com que a língua portuguesa se espalhasse nos continentes, muitas vezes tornando-se instrumento de comunicação, principalmente no que tange à colonização do Brasil, em que a língua portuguesa é imposta a grande parte da população, os colonizados. A transplantação para outros territórios, o contato com outros povos e com outras culturas resultou em uma grande diversidade de falares, embora, tal fato, não tenha impedido a intercomunicação entre os falantes europeus, africanos, asiáticos e americanos.

Entre os séculos XIV e XVI, com a construção do império português de ultramar, a língua portuguesa faz-se presente em várias regiões da Ásia, África e América. Portanto, a relação entre metrópole e colônia, estabelecida, entre essas regiões, pela coroa portuguesa, fez com que a língua portuguesa se espalhasse pelo mundo. Segundo dados fornecidos pelo site *Ethnologue: language of the World*, hoje temos 11 países que utilizam o português como veículo de comunicação, perfazendo uma estimativa de 202 milhões de falantes que constituem o mundo lusófono.

A língua portuguesa, transportada assim para o ultramar, se expandiu por vastos territórios, mas política e administrativamente, nada resta hoje do antigo Império, uma vez que os países colonizados já se tornaram independentes, a exemplo do Brasil que se tornou independente em 1822 (TEYSSIER, 1997, p. 42). Porém, a língua portuguesa permaneceu, como língua oficial, no Brasil e em diferentes países da África e da Ásia, a citar: Angola, Cabo Verde, Guiné Bissau, Moçambique, Portugal, São Tomé e Príncipe, Timor Leste.

As grandes conquistas da coroa portuguesa, os contatos com povos de diferentes culturas, costumes e línguas, deixaram muitos legados ao português falado por esses povos. Hoje, a Linguística Moderna toma como objeto de estudo o “resultado” desses contatos entre línguas, considerando seu caráter heterogêneo. No entanto, não se pode deixar de ressaltar que as consequências linguísticas decorrentes da transplantação do português para outros territórios dependem de diferentes fatores, o número de falantes portugueses que foi para um mesmo destino; as relações entre colonizadores portugueses e populações locais; o interesse dos portugueses em relação às regiões (colonização, exploração, comércio etc.), no sentido de que as variedades do português constituídas além-mar podem se apresentar mais ou menos distanciadas em relação à norma europeia (CASTRO, 1991), fatores que buscaremos discorrer ao longo do capítulo.

Portanto, este capítulo trará indícios de que as descrições sócio-históricas do português partem do pressuposto, já afirmado por Serafim da Silva Neto (1986), de que:

[...] a língua, longe de ser um organismo, é um produto social, é uma atividade do espírito humano. Não é, assim, *independente da vontade do homem*, porque o homem não é uma folha seca ao sabor dos ventos veementes de uma fatalidade desconhecida e cega. Não está obrigada a prosseguir na sua trajetória, de *acordo com leis determinadas*, porque as línguas seguem o destino dos que a falam, são o que delas fazem as sociedades que as empregam. (p. 18 - grifos do autor)

Ao citar Silva Neto (1986), destaca-se sua importância no campo da sócio-história do português brasileiro, uma vez que, na primeira metade do século XX, no ano de 1950, lança seu clássico livro *Introdução ao estudo da língua portuguesa no Brasil*, continuamente reeditado. Portanto, tem-se um trabalho que abre caminhos para estudos posteriores, no intuito de reconstituir/compreender a história do português brasileiro.

1.1 PORTUGUÊS DO BRASIL: ASPECTOS SÓCIO-HISTÓRICOS

No contexto de formação do português brasileiro, buscar-se-á dar lugar e voz aos milhares de índios, que já viviam em terras brasileiras, com culturas e línguas autônomas, como bem descreve o professor Ayron Rodrigues, e também aos negros africanos que por muito tempo tiveram sua voz e história ocultadas na formação da nossa língua. É através desse viés que buscaremos romper com o ideal eurocêntrico, impregnado nas abordagens tradicionais da história linguística do Brasil, que põe à margem toda a história criada/vivida pelos autóctones e africanos, e colaborar, mesmo que de forma descritiva, para se aproximar da compreensão – ou a tão sonhada, por Rosa Virgínia Mattos e Silva, reconstituição – do português brasileiro.

Dentre os avanços teóricos, destacam-se as análises desenvolvidas recentemente no campo linguístico, as quais buscam conhecer o passado da sociedade brasileira, informações consideradas por Mattos e Silva (2001) essenciais para uma aproximação menos generalizante da história social linguística do Brasil e, conseqüentemente, para a história do português brasileiro. Nesse sentido, entende-se que

a matéria de história da língua portuguesa no Brasil há de investigar-se na etnografia e na evolução histórico-social do povo brasileiro. Esse assunto encontra símiles na história externa do português asiático e africano, do inglês norte-americano, do espanhol argentino, chileno etc.

Mas é preciso frisar a muito maior complexibilidade do caso brasileiro, em vista das particularidades de nossa formação *étnico-social*. (SILVA NETO, 1986, p. 14 - grifos do autor)

O processo de colonização do Brasil se deu efetivamente a partir de 1532, com a implantação das 15 capitanias hereditárias, iniciando-se pelo litoral. Sendo assim, de início, só o litoral era colonizado, mas com a fundação de São Paulo e a devastação das matas litorâneas para produção da cana-de-açúcar e extração de lenha, foi aumentando a penetração em direção ao interior do Brasil, o que foi intensificado com a descoberta do ouro nas Minas Gerais e a ação dos bandeirantes (TEYSSIER, 1997; CASTRO, 1991; SILVA NETO, 1986). Sendo assim, podemos afirmar que os pontos de partida para a colonização do Brasil foram as zonas litorâneas e arredores, o que teve consequências iminentes para formação do português brasileiro (NOLL, 2010, p. 105).

Nos dois primeiros séculos de colonização, o português sofreu, no Brasil, algumas transformações que o estreparam do português lusitano (SILVA NETO, 1986, p. 28). Tal fato se explica pelo contato do português com diferentes grupos linguísticos, oportunizado pelas entradas e bandeiras, o que fez a língua portuguesa transplantada para o Brasil, no século XVI, aos poucos, assumir uma feição peculiar face ao português de Portugal. Porém tem surgido, desde as primeiras manifestações sobre a natureza da língua do Brasil, ainda no século XIX, muita polêmica sobre a existência ou não de uma nova língua, como discutiremos no capítulo que segue.

De certa forma, sabemos que os colonizadores que foram trazidos ao nosso continente, vinham das mais diversas regiões da metrópole, do campo ou da cidade. No que tange à língua desse grande contingente de falantes de língua portuguesa, acredita-se que tenham vindo de diferentes regiões de Portugal e portanto havia heterogeneidade linguística mesmo entre os portugueses, mas que, pertencendo ao mesmo domínio linguístico, eram capazes de se entender uns aos outros (SILVA NETO, 1986, p. 42).

Ao chegarem ao Brasil, os portugueses encontraram a terra povoada por índios que, do ponto de vista linguístico, apresentavam uma grande diversidade. A comunicação entre as diversas tribos era feita através de uma espécie de língua franca, “fácil, e elegante, e suave, e copiosa, a dificuldade está em ter muitas composições” (RODRIGUES, 1983, p. 23).

É a partir da década de 30 do século XVI que temos os principais contatos dos portugueses com as línguas indígenas, no Brasil, quando o número de colonos ainda era pequeno e esses dependiam dos índios para conhecer e conseguir sobreviver no Brasil (RODRIGUES, 2010, p. 30-31). Os primeiros contatos foram estabelecidos com falantes das

variedades tupi e tupinambá¹ e progressivamente os portugueses foram aprendendo a usar a língua desta comunidade indígena e aos poucos foram introduzindo em suas conversas, em português, palavras de origem indígena.

Ciente da complexibilidade do português falado no Brasil, Venâncio (2007) discute, de forma reflexiva, sobre a inviabilidade de não se reconhecer a língua portuguesa, como idioma oficial do Brasil, uma vez que o Brasil foi colonizado por portugueses. Porém, o mesmo autor afirma que essa concepção/afirmação pode e deve ser relativizada, pois sabemos que em 1500 o atual território brasileiro era povoado por vários povos indígenas, falantes de idiomas distintos, além disso, ocorreu a vinda, forçada, de milhões de africanos em “[...] 90% dos casos provenientes da África Atlântica, compreendida *grosso modo* entre Angola e Senegêmbia” (VENÂNCIO, 2007, p. 359 – grifos do autor). Assim, além das línguas gerais e das inúmeras línguas indígenas, o português concorreu com as línguas dos africanos de diferentes grupos étnicos, que desde o início foram trazidos ao Brasil.

Como assinala Mattos e Silva (2004), os negros também fizeram parte desse cenário sócio-histórico, uma vez que, em substituição ao trabalho indígena, foi crescente a imigração forçada dos negros africanos para o Brasil. Esse processo de tráfico de africanos, em que eram submetidos ao trabalho escravizado, se inicia nos primeiros anos de colonização, oficialmente em 1549, por Dom João III, e perdura por um longo período. Segundo Américo Venâncio Machado Filho (2008), esses milhões de negros africanos pertenciam, principalmente, a dois troncos linguísticos: o afro-asiático e o congo-cordofaniano. O autor afirma que o primeiro representa uma minoria e se fez presente, sobretudo na região do Nordeste, especialmente na Bahia; ao segundo, pertence a maioria dos escravos que chegaram ao Brasil, e integrava dois principais ramos de uma única família linguística – Níger-Congo.

Castro (2009, p. 179) apresenta o seguinte quadro de línguas subsaarianas documentadas no Brasil durante o período da escravidão, de acordo com a proposta de Joseph Greenberg, em 1966.

¹ Os contatos se deram especialmente com falantes das variedades do tupi e tupinambá. O mais amplo dicionário desses empréstimos, excetuados os topônimos, é o de Antônio Geraldo da Cunha (1978), que sob o nome tupi engloba as palavras oriundas do tupi, tupinambá e das duas línguas gerais. (RODRIGUES, 2010, p.46)

Família	Grupo linguístico	Língua	Século de registro	Localização	
Níger-congo	Banto	Congo-Angola	XVII ao XIX	Salvador (Bahia)	
	Oeste-africanas	KWA	Ewe-fon ou mina-jeje	XVIII	Zona de mineração – Minas Gerais
			Nagô-iorubá	XIX	Centro urbano – Salvador (Bahia)
			Tapa ou nupe	XIX	Centro urbano – Salvador (Bahia) – minoritária
	Atlântico ocidental	Fulani	XIX	Centro urbano – Salvador (Bahia) – minoritária	
	Gur ou Voltáico	Grunce	XIX	Centro urbano – Salvador (Bahia) – minoritária	
Afro-asiática	Afro-asiático	Huaçá	XIX	Centro urbano – Salvador (Bahia) – minoritária	

Quadro 1: Línguas subsaarianas

Fonte: Castro (2009, p. 179)

O *Quadro 1* nos permite afirmar que, apesar da diversidade de línguas transplantadas para o Brasil, é possível traçar uma origem comum, uma vez que, as línguas, com exceção da Huaçá, da família afro-asiática, pertencem à única família linguística Níger-congo, o que nos permite afirmar que as línguas africanas vindas ao Brasil, eram línguas aparentadas.

Mattos e Silva (2004) e outros autores, também apresentam argumentos veementes que comprovam que não só a língua portuguesa esteve presente no processo de formação da sociedade brasileira, mas também as indígenas e africanas. Porém, afirmam que os dados demográficos, apresentados em seus trabalhos, representam apenas estimativas demográficas, pois poucos registros se têm para as primeiras décadas da colonização no Brasil. Na *Tabela 1* se pode observar valores relativos a população indígena, em comparação a população europeia e africana:

Tabela 1: Estimativas de população no Brasil

Período	População indígena	População europeia e africana	População total	% de população indígena
c. 1500	4.000.000	-	4.000.000	100,0
c. 1822	800.000	3.596.132	4.396.132	18,1
c. 1889	215.000	13.733.915	13.948.915	1,5

Fonte: Venâncio (2007, p. 361)

As estimativas demonstram que nos primeiros anos de colonização, a população indígena era majoritária, mas aos pouco esse cenário sofre alterações, com a inserção da população europeia e o início do tráfico negreiro, em terras brasileiras. Em 1889, se pode visualizar uma pequena parcela da população indígena, que corresponde a apenas 1,5% da população geral, o que traduz o pronunciamento enfático de Mattos e Silva (2004, p. 14 – grifo meu) ao afirmar que no processo de colonização do Brasil “[...] se inicia a trajetória dizimadora dos índios brasileiros e de suas línguas, *percurso etnocida e glotocida* [...]”. Essas poucas e sábias palavras conseguem retratar todos os números apresentados por Venâncio (2007).

Com isso, se pode considerar que a historiografia sobre a colonização brasileira demonstra que não foi apenas o elemento português que entrou na demografia do Brasil, mas também os autóctones e os africanos, bem como outros europeus que imigraram para o Brasil na condição de trabalhadores assalariados desde a primeira metade do século XIX. Para maior precisão se faz necessário apresentar estimativas que demonstrem a imigração para o Brasil entre 1651 e 1800.

Tabela 2: Estimativas de imigração para o Brasil

Período	Portugueses	Africanos	Total de imigrantes
1651-1700	100.000	360.000	460.000
1701-1800	800.000	1.700.300	2.500.300

Fonte: Venâncio (2007, p. 361)

Sabe-se que, na medida em que imigrantes europeus chegavam ao Brasil e africanos eram introduzidos através do tráfico, a quantidade de indígenas ia sendo reduzida, porém, durante um longo período, apresentado na *Tabela 2*, esses povos com etnias, línguas e culturas diversas conviviam em um mesmo território, configurando, ou melhor, intensificando um cenário de multilinguismo, uma vez que, como afirma Lucchesi (2009), o português brasileiro apresentava-se nesse contexto não apenas um caráter heterogêneo e variável, mas também plural e polarizado.

Os dados/números apresentados nas *Tabelas 1 e 2* são essenciais, pois “informações históricas desse tipo são preciosas para a reconstituição histórica do passado linguístico do PB, [...] já que não se pode negar que a história das línguas passa

necessariamente pela história demográfica de seus falantes” (MATTOS E SILVA, 2004, p. 17).

No que tange ao caráter plural e polarizado² do PB, citado por Lucchesi (2003), sabe-se que

[...] o processo do encontro entre a língua portuguesa, língua de dominação, com muitas línguas autóctones e as diversas línguas aqui chegadas, primeiro as africanas, depois as línguas de imigrantes, que tornaram esta área americana, multilíngüe de origem, ainda mais complexa lingüisticamente. (MATTOS E SILVA, 2004, p. 11)

Silva Neto (1986, p. 19 – grifos do autor), já nos dizia que “[...] o ponto fundamental desta centenária querela de *língua brasileira* ou *dialeto brasileiro*, residiu no erro palmar de encarar-se o nosso português como *um bloco, uma massa uniforme*”. Com dados mais detalhados sobre a demografia histórica do Brasil, Mussa (1991, p. 163) apresenta as porcentagens étnicas presentes nas diversas fases de colonização:

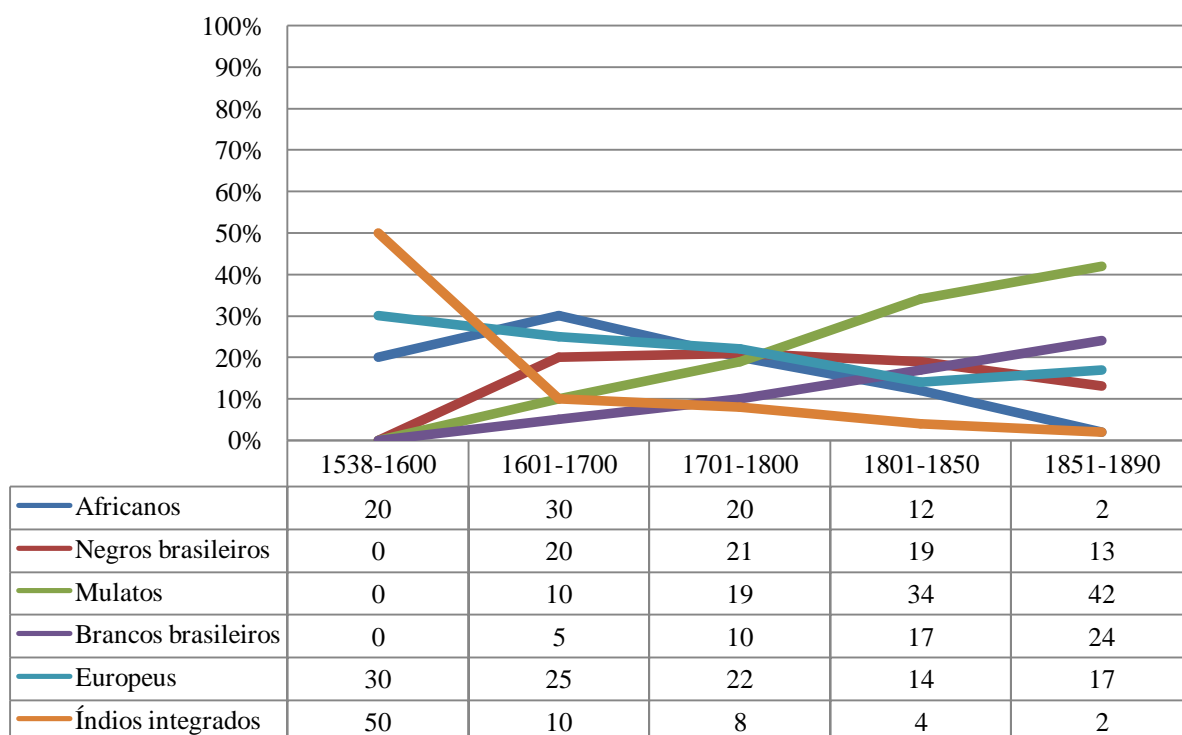


Figura 1: Demografia histórica do Brasil

Fonte: Adaptado de Mussa (1991, p. 163)

² Segundo Lucchesi (2009, p. 42) temos motivos suficientes para pensar a realidade linguística brasileira como um sistema polarizado, constituído por dois grandes subsistemas - uma norma dita culta e uma norma popular -, cada um deles com sua lógica própria, com suas respectivas tendências de mudança linguística e seu sistema particular de avaliação subjetiva.

Nos séculos XVII e XVIII já se observa um cenário pluriétnico, com população “equilibrada” entre europeus, africanos, indígenas e a primeira geração desses povos (mulatos, negros brasileiros e brancos brasileiros).

Comparando as primeiras décadas de colonização com as últimas décadas do século XIX, fica visível um drástico declínio da população indígena (de 50% para 2%), dando espaço ao aumento de mulatos (de 0% a 42%), e de brancos brasileiros (de 0% a 24%). Mattos e Silva (2004) afirma que o aumento populacional de europeus e brancos brasileiros, na 2ª metade do século XIX, é reflexo da extinção do tráfico, coincidindo com a chegada de imigrantes da Europa e da Ásia. Em contraponto, tem-se uma ascensão da população africana (de 20% para 30% - século XVI e XVII) e em seguida, com a extinção do tráfico negreiro, um declínio progressivo, chegando ao final do século XIX com apenas 2% da população geral.

Diante das pesquisas realizadas no campo da etnolinguística, já se pôde registrar as muitas línguas indígenas (RODRIGUES, 1985) e africanas (CASTRO, 2001) que se fizeram presentes nos primeiros séculos. Porém, esses povos, índios e africanos, representam, no século XIX, uma pequena parcela da população (*cf.* Figura 1), o que nos faz levantar as seguintes questões: a) se já se tem certeza do cenário plurilíngue montado em terras brasileiras, como se caracterizou o processo de intercâmbio entre índios, africanos e europeus? b) o que restou desses povos, das suas respectivas línguas, no português falado hoje, no século XXI, no Brasil? e c) como o português poderia se manter “puro” diante de um quadro multilíngue e pluriétnico, que caracterizou a implantação e difusão do português em terras brasileiras? Essas são algumas das perguntas que nos norteiam, quando se pensa em uma reconstituição do português brasileiro.

Essas questões ecoam a indagação levantada por Teyssier (1997, p. 97) no livro *História da Língua Portuguesa*, “Como explicar as particularidades do português do Brasil?”; porém sabemos que a resposta a essa indagação constitui um dos problemas cruciais na busca de resultados empíricos para se alcançar o conhecimento referente à efetiva formação do português brasileiro.

No que tange ao processo de colonização e à difusão do português no Brasil, Mattos e Silva (2004, p. 14) conta-nos, através da narrativa de Pero Vaz de Caminha, que quatro portugueses, dois degradados e dois grumetes que fugiram da frota, ficaram em solo brasileiro, e são responsáveis por semear o português no Brasil, mas antes, estes foram “obrigados” a aprender a língua dos índios, para só então, ensinar-lhes o português. Negrão e Viotti (2012, p. 14) afirmam que, até aproximadamente 1530, a coroa portuguesa enviava navios para o Brasil com dois objetivos: guardar a costa, protegendo a nova colônia da

invasão de piratas; e extrair o tão valioso pau-brasil. Nesses navios, vinham basicamente homens, a maioria de origem portuguesa, mas não todos – havia europeus de outras nacionalidades e línguas.

Negrão e Viotti (2012) refletem sobre as pessoas envolvidas na expansão imperial europeia e deduzem que esses homens certamente deviam estar acostumados a situações de contato linguístico, e provavelmente já deviam ter desenvolvido eficientes estratégias de comunicação com falantes de línguas diferentes. Sendo assim, o português que chegou ao Brasil

não se tratava de um português único, mas de uma multiplicidade de dialetos portugueses; não se tratava de um português “puro”, mas de um português que já carregava marcas do contato com uma variedade de línguas africanas e com as línguas de outros europeus que participavam da expansão mercantilista. Uma vez em território brasileiro, as relações entre os europeus e as diversas populações indígenas que viviam na costa brasileira certamente devem ter causado o *surgimento de um sistema de comunicação funcional que permitisse a interação entre eles*. (NEGRÃO; VIOTTI, 2012, p. 15 – grifo meu)

O que as autoras definem como “sistema de comunicação funcional”, pode ser entendido, dentro do contexto de colonização brasileira, como a *língua geral*³.

Essa língua geral nasceu quando os primeiros povoadores portugueses, no início do século XVI, vieram para o Brasil e entraram em contato com as tribos indígenas que habitavam a costa e, por falarem línguas muito aparentadas do tronco tupi, eram capazes de se comunicarem entre si utilizando uma espécie de *kioné*, da língua tupinambá. Essa língua franca, que viria a ser denominada *língua geral da costa brasileira* – ou simplesmente *língua geral* – foi o instrumento de comunicação adotado pelos portugueses para integrar a força de trabalho indígena inicialmente na extração do pau-brasil e posteriormente no cultivo da cana-de-açúcar, do tabaco e do algodão (LUCCHESI, 2009, p. 43). O que pode ser ratificado no trecho que Serafim da Silva Neto (1986) declara que a *língua geral* foi uma “[...] língua comumente falada na costa, que chegou até nós com o nome de ‘geral’. Ela era precisa para comunicação”.

Aryon Rodrigues (2010, p. 37) acredita que

O número de mestiços, chamados então de mamelucos, cresceu tanto usando a língua tupi (ligeiramente modificada pelo convívio com a portuguesa) que esta foi difundindo-se e passou a ser chamada de *língua geral*, hoje mais especificada como

³ A denominação *língua geral* recobre uma grande diversidade, conforme alerta Mattos e Silva (2001, p. 286) “só podemos idealizar essa língua geral como heterogênea desde o século XVI, heterogeneidade que se tornará mais complexa ao longo da diacronia da colonização [...]. Para a compreensão do que se chama genericamente língua geral, é fundamental o estudo de Aryon Rodrigues (1986: 99-109), em que, com precisão, apresenta uma caracterização diatópica e diacrônica das línguas gerais, já no plural, a paulista e a amazônica”.

língua geral paulista (LGP). Os mamelucos passaram a integrar as expedições ou bandeiras [...] levando consigo o uso das duas línguas, português e língua geral, e atribuindo nome ora numa, ora na outra aos lugares por onde passavam [...]. Como os bandeirantes criavam povoados nos lugares onde se deixavam ficar, a LGP ficou sendo falada nesses povoados em concorrência com a língua portuguesa [...].

Neste contexto de língua geral temos a contribuição dos jesuítas que, por considerarem que a variedade de línguas impedia a conversão, procuraram aprender o tupi, um tupi simplificado, despojado de seus traços fonológicos e gramaticais mais típicos. Essa língua foi estudada, fixada em catecismos, dicionários e gramáticas e institucionalizada como língua de contato entre colonizadores e índios. Portanto, a língua geral foi essencial no processo de evangelização nos séculos XVI e XVII, foi através dessa estratégia que os portugueses, de forma progressiva, invadiam as comunidades linguísticas indígenas. Com isso, se entende que

o processo colonizador e evangelizador dos séculos XVI e XVII teve de utilizar, como instrumento fundamental para a dominação, línguas indígenas brasileiras. Embora homogeneizadora da atividade catequética – construiu um ‘tupi jesuítico’ –, a língua geral da costa, de base tupi, chegou a ser um risco para a hegemonia do português no Brasil, juntamente com outras línguas gerais indígenas que foram veículos de intercomunicação entre brancos, negros e índios não só no litoral brasileiro, mas nas entradas paulistas [...]. (MATTOS E SILVA, 2004, p. 14-15)

O que Mattos e Silva (2004) define como um risco a hegemonia do português no Brasil diz respeito à amplitude de uso da língua geral entre os habitantes das terras brasileiras, uma vez que foi utilizada como principal veículo de comunicação entre brancos, negros e índios. Diante desse contexto, em 1757, a coroa proibiu que a língua portuguesa coexistisse com outras línguas, instaurando um decreto, pelo Marquês de Pombal, que fixava o português como língua oficial brasileira, proibindo o uso da língua geral. Dessa forma,

[...] se define uma política linguística e cultural que fez mudar de rumo a trajetória que poderia ter levado o Brasil a ser uma nação de língua majoritária indígena, já que os dados históricos informam que uma língua geral de base indígena ultrapassara de muito as reduções jesuíticas e se estabelecia como língua familiar no Brasil eminentemente rural de então. (MATTOS E SILVA, 2004, p. 20-21)

De fato, a partir da segunda metade do século XVIII, com o decreto do Marquês de Pombal, proibindo o uso da língua geral nas escolas, e com a expulsão dos jesuítas, em 1754, a língua geral perdeu seus principais protetores, verificando-se, então, o predomínio da língua portuguesa.

Cabe relatar, portanto, que até o estabelecimento da política linguístico-cultural de Pombal o agente escolarizador generalizado no Brasil foi a Companhia de Jesus. Mattos e Silva (2004) discorre sobre a educação na Capitania da Bahia com base na Carta VIII, do

professor Luís dos Santos Vilhena, que descreve como se processava o ensino na Companhia de Jesus:

No tempo em que existiam os jesuítas, incumbido então de todas as escolas menores, tanto em Portugal, como por todos os seus domínios, havia nos gerais do Colégio desta cidade sete classes em que se instruí a mocidade, não só da capital, como de todas as vilas... na primeira das sete mencionadas classes se ensinava gramática portuguesa, desta passavam os meninos a aprender, na segunda os primeiros rudimentos da língua latina, estudavam a sintaxe e a sílaba na terceira classe, da qual passavam para a quarta, onde aprendiam a construção da mesma língua, a retórica, tal como se ensinava; na sexta filosofia e na sétima teologia moral (VILHENA, 1969, p. 273-273 *apud* MATTOS E SILVA, 2004, p. 39-40)

Mattos e Silva (2004, p. 40) continua comentando que através da Carta VIII de Vilhena (1969) é possível inferir que de fato a língua portuguesa não era objetivo prioritário no currículo das “escolas menores” dos jesuítas. Portanto, com o decreto Pombalino, criam-se “as escolas das primeiras letras”. A autora ainda relata que através da Carta V de Vilhena (1969) temos um *flash* do iletramento, então generalizado: “[...] são uns mulatos ou negros tão estúpidos, que eu não conheci ainda um que soubesse ler, ou escrever o seu nome, e se algum branco exercita a arte, nada difere daqueles, quanto à instrução.” (VILHENA, 1969, p. 189 *apud* MATTOS E SILVA, 2004, p. 40).

Com a interrupção do trabalho nas escolas jesuíticas no Brasil, o esforço de escolarização da população local – indígena, mestiça, ou descendente de portugueses – não é retomado senão mais de uma década depois, de forma bastante intermitente e irregular. Ao se tratar dessa temática no Brasil, as fontes ainda são escassas, mas podemos citar o consistente trabalho de Houaiss (1985) que apresenta dados que demonstram que ainda no século XIX, apenas cerca de 0,5% da população era letrada, ou seja, em uma população com 3,5 a 4 milhões de habitantes, esse número correspondia a aproximadamente 18 a 20 mil letrados, os quais pertenciam à classe dirigente e dominante, ou eram “assessores” intelectuais dessa classe – como escritores, agregados, escrivães, escriturários, correspondentes, comerciantes e clero (HOUAISS, 1985, p. 141).

Ao analisar o panorama da escolarização no Brasil, Mattos e Silva (2004) conclui que os dados sobre a escolarização permitem-nos entrever que são as variantes orais dialetais espaciais e verticais que dominam o português brasileiro. Já em relação à normatização da língua, desencadeada no século XIX, a autora acredita que não houve uma generalização para o território brasileiro, uma vez que, por ser de caráter lusitanizante, restringiu-se apenas à elite brasileira.

Porém, mesmo com a implantação dos agentes promotores da normatização linguística - a escola - não temos a exclusão das centenas de línguas que vigoravam no país. Nesse sentido, Lucchesi (2001, p. 102) afirma que

[...] a língua portuguesa tinha que lutar para se impor, em primeiro lugar, contra às variedades da língua franca de base tupi falada na costa brasileira, chamada *língua geral*. Por outro lado, a língua portuguesa era adquirida nas situações as mais precárias pelos escravos, que muitas vezes preferiam se comunicar entre si, usando uma língua franca africana. Sob essas ásperas condições, a língua portuguesa foi se disseminando entre a população pobre, de origem predominantemente indígena e africana, nos três primeiros séculos da história do Brasil.

Sendo assim, Serafim da Silva Neto (1986, p.36) afirma que

[...] há muitas semelhanças entre o português dos índios e o português dos negros. Isso é, aliás, bem natural, pois tanto o índio como o negro, em atraso de civilização, aprenderam o português como língua de emergência, obrigados pela necessidade.

Até porque, o autor continua comentando, a educação não compete exclusivamente ao professor, mas encontra sólido apoio no convívio com as pessoas que nos rodeiam, nas coisas que vemos e até nas maneiras que nós observamos nos outros (SILVA NETO, 1986, p. 40).

O contato entre portugueses, índios e africanos, designado por Silva Neto (1986, p. 40) como co-educação, gerou influências mútuas, no que tange a fusão das culturas, dando origem a uma sociedade híbrida, de caráter diverso da européia. Ao tratar do vocabulário dos povoadores portugueses acredita-se que esse “restringia-se por força, porque no Brasil não se encontravam dezenas e dezenas de coisas ou fatos do Reino. Mas, por outro lado, enriquecia-se-lhes a expressão com o adotar de numerosíssimos vocábulos designativos de coisas e fatos americanos” (SILVA NETO, 1986, p. 40). Diante de tal contexto o autor questiona: enriqueceu-se no Brasil a linguagem corrente portuguesa ou, pelo contrário, diminuíram-lhe os recursos? Busca-se até hoje a resposta para esse questionamento.

Diante das discussões podemos afirmar que os debates sobre a formação do português brasileiro têm proporcionado grandes reflexões, mas muitas histórias sobre o passado do Brasil ainda precisam ser reconstituídas para que seja possível compreender a história do diversificado português brasileiro.

1.1.1 População africana no Brasil

Como já relatado, a formação da nossa língua parece estar inteiramente interligada ao processo de colonização do Brasil, quando ocorre o encontro plurilinguístico entre o

português europeu e as muitas línguas indígenas e africanas, principalmente quando se trata das línguas africanas, cujos falantes foram os principais difusores do português vernacular brasileiro, como afirma Mattos e Silva (2004).

Uma África plural, dotada de uma grande diversidade de povos, culturas e línguas, com suas correspondentes variedades dialetais, configura-se como um dos caminhos que nos faz entender mais sobre o português brasileiro. Já que, pesquisas (CASTRO 2005; LUCCHESI 2003, 2009; MATTOS E SILVA, 2004; MUSSA 1991; dentre outros) demonstram, que nas primeiras décadas de colonização os negros africanos representavam parcela considerável da população brasileira.

Bonvini (2009) retrata quatro grandes ciclos de importação de escravos para o Brasil:

- (i) no século XVI, o ciclo da Guiné, sendo os escravos principalmente sudaneses, originários da África situada ao norte do equador;
- (ii) no século XVII, o ciclo do Congo e de Angola, que trouxe ao Brasil negros da zona banta;
- (iii) no século XVIII, o ciclo da costa de Mina, que atingiu de novo negros sudaneses. A partir da metade do século XVIII, esse ciclo desdobrar-se-á num ciclo propriamente baiano: o ciclo da baía do Benim;
- (iv) no século XIX, os escravos vieram um pouco de cada lugar, mas com predominância de negros vindos de Angola e de Moçambique. (BONVINI, 2009, p. 26 – grifo meu)

O autor afirma que a repartição em ciclo só deve ser entendida dentro das suas grandes linhas, porque o fim de um ciclo não implica o interrompimento do tráfico de negros vindo da região representada no ciclo anterior (BONVINI, 2009, p. 27). Neste cenário – brasileiro – Angola se caracteriza como o grande centro de tráfico negreiro e se torna incontestável que a escravidão implicou na transferência de milhões de africanos, os quais endossaram o cenário sócio-histórico.

Albuquerque e Fraga Filho (2006, p. 45 – grifo meu) comentam que o

[...] tráfico transatlântico promoveu o povoamento do Brasil por gente vinda de diversas regiões do continente africano. A metrópole portuguesa adotou a política de misturar escravos de diferentes regiões e etnias para impedir a concentração de negros da mesma origem na colônia, os quais, solidários na cultura e falando a mesma língua, podiam se rebelar mais facilmente. Essa política nem sempre era seguida à risca, pois a depender das relações comerciais na África, os traficantes tendiam a transportar escravos que em sua maioria vinham duma mesma região. Foi por isso que, no século XVI, a maioria dos escravos trazidos para o Brasil vinha da região da Senegâmbia, denominada Guiné pelos portugueses [...]. Mas, no decorrer daquele século até a primeira metade do século XVIII, os chefes políticos e mercadores do território presentemente ocupado por Angola forneceram a maior parte dos escravos utilizados em todas as regiões do Brasil [...]. *Luanda, Benguela e Cabinda eram os principais portos de embarque.*

No que tange às línguas africanas vindas para o Brasil, Bonvini (2009) e Petter (1998) estimam entre 200 e 300 línguas e Mendonça (2012) relata que na linguística africana só se

projetam duas individualidades fortes: línguas bantu e línguas não bantu. Diante da afirmação de Mussa (1991, p. 146) segundo a qual “[...] o percentual de falantes bantu foi sempre superior, e quase sempre maciçamente, em todo o período do tráfico” no Brasil, pode-se deduzir que a influência bantu, é muito mais profunda em razão da sua distribuição humana em território brasileiro (CASTRO, 2005, p.6). Dessa forma, dentre os negros falantes de língua africana, habitantes do Brasil, temos:

Tabela 3 - População de falantes das diversas línguas africanas que chegaram ao Brasil

	XVI	XVII	XVIII	XIX
Oeste-atlântico	20%	7%	1%	3%
Mande	20%	7%	1%	3%
Kru	14%	5%	1%	3%
Gur	1%	3%	8%	9%
Benue-Kwa (não bantu)	7%	10%	20%	24%
Bantu	35%	65%	64%	50%
Outros	3%	3%	5%	8%

Fonte: Mussa (1991, p. 163)

No que tange aos dados demográficos, Alkmim (2007) relata que ao longo do período colonial, o tráfico de escravos se centralizou nas regiões da África, porém, pouco se sabe sobre o contingente de negros traficados para o Brasil, devido à ausência de documentos históricos, mas estima-se entre 3.500.000 e 3.600.000 escravos africanos, até 1850. Mattos e Silva (2004, p. 105) ressalta que Robert Conrad - em sua pesquisa sobre os últimos anos de escravidão no Brasil - demonstra que às vésperas da abolição da escravidão, em 1884, a população escrava se concentrava em Minas Gerais, Rio de Janeiro e São Paulo, graças à migração do Nordeste para a área cafeeira. Vejamos as estimativas:

Tabela 4: População escrava

PROVÍNCIA	1884 / POPULAÇÃO ESCRAVA
Minas Gerais	310.000
Rio de Janeiro	260.000
São Paulo	160.000
Bahia	140.000
Pernambuco	80.000
Rio Grande do Norte	60.000
Maranhão	45.000

Fonte: Conrad (1978, p. 81 *apud* MATTOS E SILVA, 2004, p.105)

A presença maciça de africanos e afro-descendentes na demografia histórica do Brasil demonstra não apenas a atuação dos escravos frente à economia da colonização, a mobilidade geográfica, mas também o significado social e linguístico dos espaços ilegítimos de escravidão, os quais nos tem permitido inferir que esse segmento numeroso e operante – os africanos e afro-descendentes – é o agente principal da difusão do português no território brasileiro, na sua majoritária, a popular ou vernácula (MATTOS E SILVA, 2004, p. 106).

Porém, a história do negro brasileiro é uma história anterior à escravidão nas Américas, à vida de cativo no Brasil. Trata-se de uma saga que se cruza com a aventura dos navegadores europeus, principalmente os portugueses, e com a formação do Brasil como país (ALBUQUERQUE; FRAGA 2006, p. 13). Por isso, é essencial conhecer a história da África para entendermos um pouco mais sobre a triste trajetória dos negros trazidos para o Brasil.

1.2 PORTUGUÊS DE ANGOLA: ASPECTOS SÓCIO-HISTÓRICOS

No período das conquistas marítimas lusitanas, no século XV, muitas áreas da costa africana foram conquistadas, enquanto Portugal buscava estabelecer seu comércio de especiarias com a Índia, dentre as conquistas temos o território onde hoje fica Angola, que foi ocupado pelos portugueses somente algumas décadas depois da ocupação efetiva do Brasil.

Segundo Mingas (2000), a República Angolana é resultado de um conjunto de antigos reinos africanos, organizados em fronteiras não naturais, criadas pela Conferência de Berlim. Dessa forma, se entende que o processo de colonização da África foi resultado de um jogo de interesses europeu, o que gerou a partilha do continente, na Conferência de Berlim, em 19 de novembro de 1884 e 26 de fevereiro de 1885, que hipoteticamente buscou organizar a ocupação do continente africano e consagrou Angola como colônia dos portugueses, que inicialmente, limitaram-se ao litoral.

Os estudos sobre a colonização do continente africano afirmam que a Conferência resultou numa divisão arbitrária que desrespeitou as características étnicas e culturais de cada tribo que vivia naquele território, muitas tribos rivais foram obrigadas a conviver em um mesmo espaço geográfico, e outras unidas foram obrigadas a se separar. Tudo isso culminou em muitos conflitos no continente africano, uma vez que toda a história desses povos foi anulada diante do etnocentrismo europeu. Entretanto, “[...] a resistência do povo angolano foi tal, que, apesar da sua presença no território desde o século XV, somente em 1926, puderam os colonialistas portugueses dominar totalmente o país” (MINGAS, 2000, p. 30).

O maior interesse dos portugueses, nas terras africanas, decorria do rentável comércio de escravos, que se inicia no século XV e rende muitos desembarques além-mar, realizados por europeus. Sem dúvidas o tráfico de negros africanos foi um dos maiores movimentos populacionais da história e certamente a maior migração por mar antes da grande emigração europeia, também originariamente para as Américas, que se desenvolveu à medida que o tráfico de escravos no Atlântico conhecia o seu fim (FAGE, 2010, p. 269).

No que tange, especificamente, a Angola, Fage (2010, p. 305) acredita que os

[...] interesses dos Portugueses na África ocidental banto, em Angola, não estavam ainda totalmente definidos antes 1680, e só nessa data é que outros comerciantes europeus interessados na aquisição de escravos para as Américas começaram a dedicar muita atenção ao litoral entre Angola e os Camarões.

Segundo Teyssier (1997) a descolonização do continente africano, após a revolução de 25 de abril de 1974, levou à constituição de 05 (cinco) repúblicas independentes: Cabo Verde, Guiné-Bissau, São Tomé e Príncipe, Angola e Moçambique, as quais constituem a África lusófona. No que tange a língua oficial adotada pelas cinco repúblicas, Teyssier (1997) afirma que o português é a língua oficial desses territórios, e era utilizada na administração, no ensino, na imprensa, assim como nas relações com o mundo exterior. Assim, sobrepondo-se às línguas locais, os falares africanos, mas isso não implica a exclusão das muitas línguas

faladas em território africano, pois na comunicação corrente, nas mesmas comunidades linguísticas, as línguas locais ainda eram o principal instrumento de comunicação.

Essa situação de plurilinguismo, instaurada na costa da África, certamente fica mais evidente quando se trata de Angola, uma vez que os dados de Teyssier (1997) comprovam que dentre as repúblicas citadas, Angola era a maior em dimensão territorial (1.246.700 km²) e a segunda maior em números de habitantes (5.840.000).

No que tange ao contexto sócio-histórico de Angola, Mingas (2000) apresenta informações importantes sobre a organização social dos angolanos no período colonial, afirmando que a sociedade era dividida em dois grupos principais – os portugueses e os angolanos – com suas respectivas subdivisões. Vejamos:

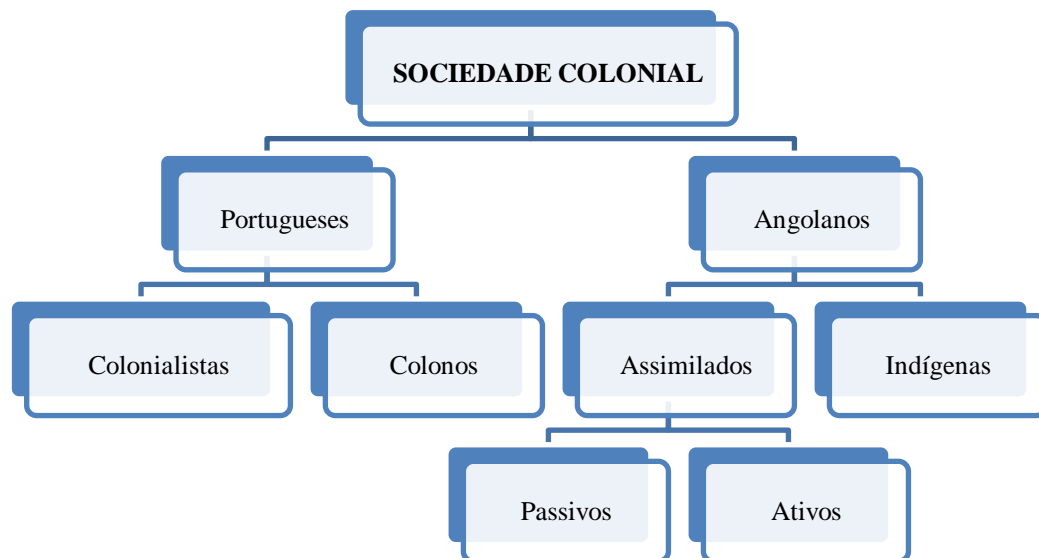


Figura 2: Organização social de Angola no período colonial
Fonte: Mingas (2000)

Segundo Mingas (2000, p. 44-45), no período colonial a cidade de Angola era dividida em três zonas habitacionais: a) a costa ou centro, onde habitavam os colonialistas (agentes da metrópole); b) a zona dos bairros arenosos, onde habitavam os colonos (instrumentos de colonização) e uma parte dos assimilados; c) os arredores (musseques), onde habitavam a maioria dos assimilados e indígenas. Chegados em grupos numerosos, os colonos e colonialistas portugueses consolidavam sua posição na colônia, organizando um grupo à parte, distinto da maioria angolana, porém, “sendo na sua quase maioria proscritos e condenados, rejeitados pelos seus, uma parte dos colonos misturou-se com a população autóctone daí a percentagem muito elevada de mestiços em Angola” (MINGAS, 2000, p. 45).

A autora segue afirmando que aos poucos a população de portugueses nascidos em Angola começou a superar o dos nascidos em Portugal, o que fez com que os colonialistas decidissem diferenciar entre:

- (i) Portugueses ou brancos de primeira – nascidos em Portugal, os quais detinham algum privilégio, assim como ajuda econômica e cargos de alta chefia na colônia;
- (ii) Portugueses ou brancos de segunda – nascidos em Angola.

No que tange aos privilégios, podemos destacar que os angolanos assimilados (*cf.* Figura 2) possuíam o direito de seus filhos terem acesso a escola conjuntamente com as crianças portuguesas. Na subdivisão em passivos e ativos, podemos afirmar que os primeiros “eram alienados, não assumindo a sua cultura, o seu povo, a sua história nem a sua condição de seres humanos. Eram o que os colonialistas designavam como ‘negros diferentes’, ‘negros que não eram como os outros, porque tinham a pele negra mas a alma era branca’” (MINGAS, 2000, p. 46); já os ativos representavam os angolanos que aprendiam tudo o que fosse possível, no intuito de libertar o seu povo e país da dominação colonial portuguesa, esses “utilizaram a língua do colonizador para denunciar o colonialismo português através dos jornais [...]” (MINGAS, 2000, p. 47).

Ao tratar de língua não poderemos deixar de refletir, através das palavras de Mingas (2000), sobre o processo de ensino em Angola. Com a política portuguesa de ensino, tem-se a supremacia da língua portuguesa em detrimento da muitas línguas locais e a

aplicação prática deste tipo de política linguística foi apoiada por uma vigilância cada vez mais acentuada da polícia política portuguesa, a famosa PIDE, que via na utilização de qualquer uma das línguas locais ou mesmo de vestuários locais pelos ‘assimilados’, atitudes subversivas. (MINGAS, 2000, p. 48)

Portanto, a língua portuguesa era a única língua ensinada na escola primária, exceto as escolas dirigidas por missionários, onde a língua local podia ser usada como meio de ajuda para facilitar a aprendizagem do português.

1.2.1 Angola: situação linguística

Com a política linguística, a língua portuguesa, em Angola, era uma espécie de passaporte à ascensão social e aos “assimilados” era negado o direito de falar outra língua que não fosse português. Contudo Mingas (2000, p. 49) afirma que cada grupo social correspondia a um nível linguístico:

- (i) Portugueses – monolinguísmo minoritário (português);
- (ii) Assimilados – bilinguismo (kimbundo/português);
- (iii) Indígenas – monolinguísmo maioritário (kimbundo).

Parte dos falantes bilíngues, que tinham o português como segunda língua, falavam o que Mingas (2000) denomina de ‘pretoquês’, que seria um português morfo e sintaticamente simplificado. Porém, como estratégia de assegurar a imposição do português, como língua oficial, este “pretoquês” passou a ser ridicularizado em peças teatrais, o que levou os “assimilados” a contribuírem para a política linguística dos colonizadores evitando falar sua língua materna e proibindo o uso dessa língua por seus filhos (MINGAS, 2000, p. 50).

Este contexto sociolinguístico criou condições propícias ao surgimento de falantes do português como língua primeira, mas o contexto familiar de pais bilingues e o fato de ser a população constituída, na sua grande maioria, de analfabetos resultou na transmissão de um português com interferências das línguas locais, fato que trataremos no capítulo a seguir.

Ao se tratar das línguas de Angola e as demais línguas da África tem-se suas distribuições, graças às relações de proximidade ou de afastamento que apresentam, em famílias de línguas. Nesse viés, Castro (2005), a partir da classificação de Greenberg, apresenta a seguinte distribuição em troncos linguísticos:

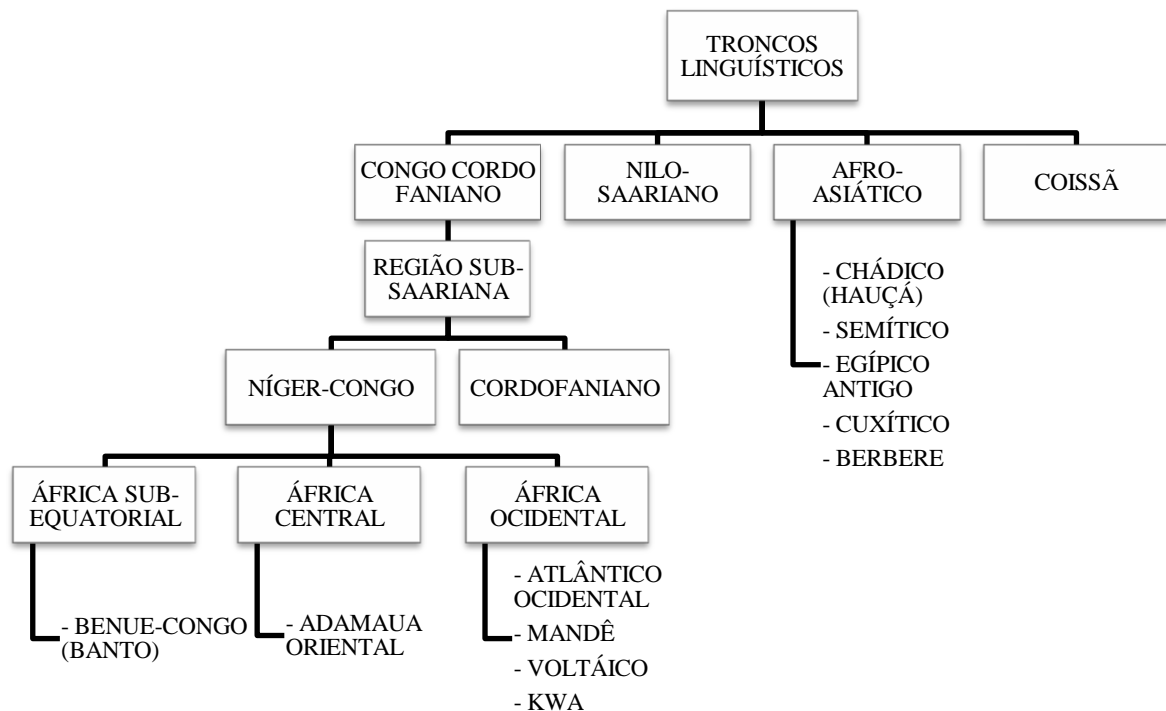


Figura 3: Distribuição de Greenberg: troncos linguísticos africanos
Fonte: Castro (2005, p. 27)

O esquema nos permite visualizar que Greenberg subdivide as línguas africanas em quatro troncos linguísticos, sendo um deles o *Congo Cordofaniano*, que representa a região subsaariana e se subdivide em duas grandes famílias, a *Níger-Congo* e a *Cordofaniano*, a primeira, se faz presente em três regiões - *África Subequatorial, Central e Ocidental* - sendo o grupo banto representado pela grande família *Níger-Congo*, no ramo *Benue-Congo*. A repartição, proposta por Greenberg, foi adotada por grande parte dos linguistas africanistas, com algumas revisões ligadas à reorganização interna de cada tronco (BONVINI, 2009 p. 23).

Assim como Mingas (2000) acreditamos que o fenômeno da colonização em Angola pôs de um lado, uma comunidade lusófona e, de outro, uma comunidade majoritária, de expressão banto, o que nos faz subentender a existência de uma forte influência das línguas banto no português falado pelos angolanos.

1.2.1.2 Línguas Banto em Angola

Pela vastidão territorial, Angola é o segundo maior país de língua oficial portuguesa. A realidade linguística do país é, porém, bem diferente da brasileira. A população, majoritariamente concentrada nas maiores cidades costeiras, tem como língua materna uma ou mais das 40 línguas banto existentes no país (INVERNO, 2004, p. 02). Dentre essas línguas, as mais significativas são: o umbundo (4.000.000), o quimbundo (3.000.000) e o quicongo (1.144.000 falantes), resumindo-se o português como língua materna a apenas 57600 falantes, majoritariamente elites e jovens dos centros urbanos do litoral, numa população total de 12.479.000 de habitantes (GRIMES, 1988; ASHER, 1994, p. 127 *apud* INVERNO, 2004, p. 02).

Castro (2005) afirma que o termo banto foi proposto por W. Bleek, em 1862, na primeira gramática comparativa do banto, no intuito de denominar a família linguística que havia descoberto, “[...] composta de várias línguas oriundas de um tronco comum, o protobanto, falado há três ou quatro milênios atrás” (CASTRO, 2005, p. 25). A autora ainda afirma que só posteriormente pesquisadores se utilizaram deste termo para denominar os indivíduos que habitavam a extensão abaixo da linha do equador. Dessa forma, a região banto, engloba: República Centro-Africana, Camarões, Guiné Equatorial Gabão, Angola, Namíbia, República Popular do Congo (Congo-Brazzaville), República Democrática do Congo (DRC ou Congo-Kinshasa), Zâmbia, Burundi, Ruanda, Uganda, Qênia, Malauí, Zimbábue, Botsuana, Lesoto, Moçambique, África do Sul (CASTRO, 2005).

Estima-se a 400 mil, ou mais, línguas banto faladas em toda extensão territorial da África. Pesquisas comprovam suas semelhanças, tanto na estrutura gramatical quanto no vocabulário. No século XIX, com o avanço dos estudos científicos sobre as línguas banto, se definem estas línguas como uma família linguística, com muitas similiaridades, dentre as muitas outras famílias linguísticas presentes na África (*cf.* Figura 3). Dessa forma, “[...] qualquer língua banto podia ser parcialmente inteligível para os falantes de línguas próximas, e talvez não totalmente ininteligível para os falantes de qualquer outra língua banto, mesmo uma que fosse falada a uma distância considerável” (FAGE, 2010, p. 31), o que leva alguns autores a considerarem as línguas banto como uma família distante/separada das outras famílias linguísticas africanas.

Fage (2010), em um percurso cronológico sobre as origens da sociedade africana, afirma que “[...] os povos hoje chamados Bantos se tornaram o tronco humano dominante em toda a metade meridional de África, à exceção do extremo sudoeste, dentro e à volta do deserto do Calaári [...]” (p. 31) e que o termo banto corresponde a uma classificação linguística das línguas faladas, atualmente, pelos povos negros, em sua grande maioria, habitantes da metade meridional da África. Nesse sentido, se entende que

[...] alguma forma do radical *nto*, como sentido genérico de <<ser humano>>, e o de estas línguas apresentarem um sistema de categorias de substantivos em que uma forma do prefixo *ba-* significa o plural da categoria que denota pessoas. (FAGE, 2010, p. 31)

Mingas (2000) relata que em Angola existem muitas línguas pertencentes à família linguística banto, como o umbundo, kimbundu, cokwe, kikongo (variantes), o helelo, oxindonga, ngangela e o nhaneka e outras não pertencentes a família banto como o khoisan e o vatwa. No entanto, em um contexto similar ao de colonização do Brasil, todas essas línguas vêm sendo “substituídas” pelo português, que se torna a única língua oficial de Angola. Vale a ressalva, que essa suposta substituição não implica no desaparecimento das muitas línguas nativas.

Com a imposição do português, as línguas locais foram reduzidas e se desenvolveram separadamente, dentro das suas comunidades linguísticas, uma vez que todos eram obrigados a falar a língua do colonizador. Diante deste contexto de imposição e de plurilinguismo, Mingas (2000) esboça uma interessante analogia, comparando essas línguas a um arquipélago linguístico, onde cada língua representa uma ilha, com suas especificidades linguísticas.

Numa rápida trajetória sobre a situação linguística de Angola torna-se claro o contato entre duas sociedades, uma de origem europeia, falando uma língua neolatina, e outra de

origem africana, falando uma língua banto, no contexto de Luanda, o kimbundu. Mingas (2000) ressalta que a situação de bilinguismo não era restrita aos falantes angolanos, portugueses, principalmente os comerciantes, também foram obrigados a adquirir a língua do colonizado. Cabe, no momento, avaliar todo o contexto sócio-histórico para refletir sobre até que ponto houve interferência da língua oficial sobre as nacionais, e vice-versa.

2 ORIGEM DO PORTUGUÊS BRASILEIRO: BREVES DISCUSSÕES

Da América à Ásia, cada povo que fala hoje o português o modelou, o recriou à sua imagem. Nenhum exemplo é mais relevante do que o do Brasil.

(EDUARDO LOURENÇO)

Desde o século XIX, a questão da formação do português brasileiro vem sendo discutida. Tal discussão foi iniciada pelo filólogo português Adolpho Coelho e continuada por Serafim da Silva Neto e, mais recentemente, por linguistas tais como Gregory Guy, John Holm, Fernando Tarallo, Dante Lucchesi, Anthony Naro, Marta Scherre, Volker Noll, Mikael Parkvall, só para citar alguns. Portanto, o objetivo desse capítulo é apresentar, a partir dos pressupostos dos autores acima citados, os pontos convergentes e divergentes de seus posicionamentos e, de forma mais consistente, refletir sobre as duas vertentes que têm sido objeto de acaloradas discussões: deriva secular e transmissão linguística irregular.

Como foi possível observar no *Capítulo I*, o grande desafio que se coloca à sócio-história da língua portuguesa do/no Brasil é justamente desvendar o mistério que envolve a formação dessa língua. Acreditamos que o caminho que tem sido percorrido em trabalhos recentes sobre a formação do português brasileiro em que se buscam informações, tais como a constituição da população brasileira, o nível de letramento do seu povo, o contínuo rural/urbano (BORTONI-RICARDO, 2011), as relações sociais que se estabeleceram entre os diversos segmentos da sociedade, especialmente entre escravos e não-escravos, têm possibilitado desenhar a imagem do povo brasileiro. Portanto, debruçaremos sobre a formação do português que se constituiu e se constitui no Brasil, afinal, retomando nossa epígrafe, “[...] cada povo que fala hoje o português o modelou, o recriou à sua imagem” (LOURENÇO, 2001, p. 132).

Como explicar a história de uma língua tão diversa em suas especificidades? Seria mesmo fruto de um contato linguístico? Ou um processo natural de mudança linguística? Essas são questões que têm estado presentes nos grandes debates sobre a constituição histórica do português brasileiro e têm impulsionado, desde as últimas décadas do século XX, pesquisas que apresentam as especificidades do português brasileiro.

Pesquisas desse cunho, normalmente, são conduzidas sob os “olhares” investigativos da Sociolinguística e da Linguística Histórica, ciências que se debruçam sobre os fenômenos linguísticos e tentam explicá-los a partir de seus processos sincrônicos e diacrônicos, respectivamente. Através do aparato teórico-metodológico dessas ciências, é possível a

realização de diversas pesquisas sobre a sócio-história do português do Brasil, o que tem fortalecido os debates sobre a sua origem. É nessa linha que se insere o nosso trabalho que, ao mesmo tempo que considera fatores externos (sociais) numa tentativa de projeção da imagem da comunidade estudada, também leva em conta fatores linguísticos que podem influenciar na escolha da manutenção da marca de plural no sintagma nominal ou na sua eliminação, conservando apenas a marcação no primeiro elemento do sintagma.

2.1 DERIVA SECULAR DAS LÍNGUAS ROMÂNICAS

Quando se fala em deriva, pensa-se no trabalho pioneiro de Anthony Naro, o qual há muito tempo vem apresentando, juntamente com a pesquisadora Marta Scherre, evidências de que o PB é resultado das nossas origens românica e lusitana, considerando que especificidades do PB representam modificações pelas quais o PE passou no Brasil, as quais já seriam previstas no sistema, entendendo que a mudança transparece na frequência dos usos e que a gramática da língua já pressupunha tais variações/mudanças.

No livro *Origens do português brasileiro*, Naro e Scherre (2007, p. 47) apresentam uma descrição geral do modelo de desenvolvimento do português popular do Brasil (PPB), afirmando que a língua portuguesa falada em Portugal, antes da colonização do Brasil, já possuía uma deriva secular que a impulsionava ao longo de um vetor de desenvolvimento, o qual, no Brasil, se encontrou com outras forças que reforçavam e expandiam a direção original desta língua. Essas pressuposições baseiam-se na noção de deriva linguística nos termos de Edward Sapir (1949[1921]) de que a língua se move ao longo do tempo num curso que lhe é próprio, o que implica que as mudanças linguísticas estão, de certa forma, previstas no processo de evolução natural da língua⁴.

Portanto, a hipótese dos autores difere da hipótese da crioulização ou semicrioulização da língua portuguesa em terras brasileiras, uma vez que acreditam que os processos de variação e mudança do PB não advêm do contato com as línguas que vieram para o Brasil com seus povos, ou das línguas ameríndias, que aqui já se encontravam quando chegaram os colonizadores europeus (*cf.* Capítulo I).

Mesmo negando a influência das línguas africanas e indígenas, na formação do português popular do Brasil, Naro e Scherre (2007), por mais contraditório que pareça ser,

⁴ Trecho original: “Language moves down time in a current of its own making. It has a drift” (SAPIR, 1921).

reconhecem a importância desses povos na configuração social e cultural do país: “longe de negar a importância da influência africana e indígena para nossa cultura” (p. 17). Eis a grande questão: como desvincular língua e sociedade? Estariam os autores abrindo “fendas” na já consolidada Sociolinguística?

Para amenizar tal contradição, os autores passam a considerar as línguas africanas e indígenas como *força de atração*, as quais juntas se reforçam para produzir o português popular do Brasil (NARO; SCHERRE, 2007, p. 25).

Todas essas questões são minuciosamente explicadas, a partir de argumentações históricas, que favorecem a compreensão da proposta de deriva secular, defendida ainda por alguns pesquisadores que se debruçam sobre as origens do português brasileiro.

Sabe-se que a partir do trabalho de Silva Neto (1986) há referências à fala dos índios e dos brancos, porém Naro e Scherre (2007) ressaltam que, nas pesquisas desse autor, não se encontra nenhuma evidência documental do português falado pelos africanos trazidos para o Brasil, a não ser algumas menções a africanos que não dominavam a língua portuguesa, mas a língua geral tupi ou outras línguas africanas. Nesse sentido, Naro e Scherre (2007, p. 28) afirmam que “a documentação não transmite nem a mais leve suspeita de que a língua portuguesa falada pelos brasileiros descendentes dos primeiros cativos africanos fosse diferente da fala de brasileiros de outras origens étnicas”. Nesse sentido, Heliana Mello (2002, p. 344) afirma que

no estudo da mudança linguística, a maior dificuldade que se encontra, ao se tentar avaliar um período remoto na história de uma língua, é a falta de dados associados à variabilidade intrínseca de qualquer comunidade linguística. Portanto, normalmente o que se faz é descartarem-se os idioletos, quando encontrados, e enfocar-se em uma amostragem que se considere representante da comunidade linguística em estudo. Mas não se pode esquecer que a deriva é consequência justamente da escolha, por parte da comunidade linguística, de formas que terão predominância sobre outras dentro de uma ecologia da variabilidade linguística.

Dessa forma, a hipótese da deriva secular, apregoada por Naro e Scherre (2007), apresenta a ideia de que o PB é, na verdade, uma espécie de continuação do português arcaico, com alterações/variações. Assim, acreditam que

[...] o português moderno do Brasil é o resultado natural da deriva secular inerente na língua trazida de Portugal, indubitavelmente exagerada no Brasil pela exuberância do contato de adultos, falantes de línguas das mais diversas origens, e pela nativização desta língua pelas comunidades formadas por estes falantes. (NARO; SCHERRE, 2007, p. 69).

E argumentam em favor da ideia de que

no português do Brasil inexistente influência gramatical específica de qualquer língua africana, ou de língua de qualquer outra proveniência não portuguesa, como também não existe nenhuma forma ou estrutura inteiramente nova criada por um processo geral de simplificação durante a fase de aquisição da língua. (NARO; SCHERRE, 2007, p. 182)

Luchessi (2012, p. 269) critica essa perspectiva, adota por Naro e Scherre, afirmando que

os fatos aduzidos por Naro e Scherre poderiam, na melhor das hipóteses, servir para refutar a hipótese do contato, não para firmar a hipótese da deriva. Apresentar evidências que neguem a variação que se observa hoje no PB esteja historicamente relacionada ao contato entre línguas não significa demonstrar que essa variação tem a sua origem numa deriva secular. Tal relação não resulta automaticamente de refutação da hipótese do contato linguístico, ela carece igualmente de fundamentação teórica e sustentação empírica [...].

Importante relatar que, mesmo hoje, Naro e Scherre sendo referências ao se tratar da hipótese da Deriva Secular do PB, Joaquim Mattoso Câmara Jr. (1972) já fala nesses termos, quando procurou uma razão estrutural, interna, para explicar o uso brasileiro do *ele* acusativo, na expressão “eu vi ele”. Em seu trabalho ele exclui a possibilidade de um crioulo de base indígena, porque as línguas indígenas “foram substituídas no intercuro dos índios com os brancos por uma língua única – o chamado Tupi” (p. 75), restringindo-se aos empréstimos léxicos sua contribuição ao PB. Quanto às línguas africanas, o autor afirma que os escravos negros adaptaram-se ao português sob a forma de um falar crioulo, o que impossibilitaria possíveis mudanças fonológicas e gramaticais profundas sem correspondência com as próprias tendências estruturais da língua portuguesa (Cf. p. 75-77). Portanto, Câmara Jr. admite o impacto das línguas indígenas e africanas sobre o PB e, embora não admita a existência de um crioulo africano, em nenhum momento afasta a hipótese da mudança natural, ou hipótese da deriva.

Dessa forma, podemos inferir que os autores que tratam da Deriva secular acreditam que “toda a gama de traços variáveis evidentes e/ou abstratos, embora com porcentagens de uso talvez ínfimas, veio toda, em suas origens, com os portugueses” (NARO; SCHERRE, 2007, p. 182).

Portanto, conclui-se que na vertente da deriva secular, tenta-se provar a inexistência de traços linguísticos do português falado no Brasil que sejam diretamente ligados a etnia afro-brasileira, o que exclui hipóteses de uma possível língua pidgin ou crioula de base lexical portuguesa, uma vez que estudos têm evidenciado o fato de que características morfossintáticas e fonológicas do português brasileiro são heranças românicas e portuguesas arcaicas e clássicas, e não modificações advindas das línguas africanas, ou das línguas dos

povos ameríndios, como afirmam estudiosos que buscam explicar a origem do PB através do contato linguístico. Em contrapartida, Baxter e Lucchesi (1997), Lucchesi (2001, 2012) acreditam que o contato entre línguas, estabelecido no nosso território durante a colonização, poderia fornecer-lhes uma explicação mais palpável para a formação do PB (BAXTER; LUCCHESI, 1997; LUCCHESI, 2001, 2012).

2.2 CONTATO LINGUÍSTICO

O contato entre línguas na constituição da história do PB é um assunto que tem sido amplamente discutido desde o século XIX, quando já se observava que alguns usos pareciam ser característicos do português falado no Brasil, o que o distanciava do português europeu. Os estudos desde então têm avançado em torno deste assunto, e muito se tem discutido sobre a origem do PB, buscando-se comprovações em todos os níveis de análise linguística: semântico, fonético, morfológico e sintático.

A hipótese do contato linguístico, como já é possível inferir a partir de sua própria designação, acredita que as dinâmicas de contato linguístico tiveram um papel fundamental nas especificidades do português brasileiro, frente ao português europeu, especificamente nas variedades do português popular brasileiro. Diante das reflexões de Rosa Virgínia Mattos e Silva (2001, p. 298-299), sobre a formação do português brasileiro, acredita-se que esse resulta de vários processos sócio-históricos e linguísticos, marcados pelo multilinguístico/multidialealismo, a saber:

- a. *o português europeu* na sua dialeção diatópica, diastrática, que teria ao longo do período colonial um contingente de 30% da população brasileira; seria esse *português europeu*, base histórica do *português culto brasileiro* que começaria a elaborar-se a partir da segunda metade do século 18;
- b. *as línguas gerais indígenas*, que, plurais e dialetalizadas, poderiam até confundir-se com *o português geral brasileira* nas áreas geográficas delimitáveis em que se difundiram;
- c. *o português geral brasileiro*, antecedente histórico do *português popular brasileiro* que, adquirido na oralidade e em situações de aquisição imperfeita, é difundido pelo geral do Brasil sobretudo pela maciça presença africana e dos afro-descendentes que perfizeram uma média de mais de 60% da população por todo o período colonial. (MATTOS E SILVA, 2001, p. 298-299 – Grifo meu)

Na perspectiva do contato, temos o filólogo e linguística Serafim da Silva Neto, que já falava nesses termos, ao qual devemos os créditos às pesquisas linguísticas e filológicas desenvolvidas sobre o português do Brasil, já que desde 1960, em artigo publicado no livro *Língua, cultura e civilização*, no trabalho intitulado *Problemas do português da América*, dizia:

A linguagem falada em nossa terra, em virtude de múltiplos fatores, tomou cunho próprio [...] o modo de viver modifica-se e transforma-se. E a língua, instrumento social, foi-se adaptando à nossa sociedade pois, como sabemos, ela caminha lado a lado com a história social. (SILVA NETO, 1960, p. 248-258)

Neste cenário, retomamos e reafirmamos, mais uma vez, a importância de Silva Neto que, a partir do seu clássico livro *Introdução ao estudo da língua portuguesa no Brasil* (1950), inicia a tentativa de uma reconstrução do passado da língua portuguesa no Brasil e destaca o contato da língua da colonização com a língua dos *alóglotas*⁵.

Como já observado, no capítulo primeiro deste relatório de pesquisa, a partir dos dados de Mussa (1991), o fato de os africanos e afrodescendentes comporem a maior parte do contingente populacional em território brasileiro, entre os séculos XVII e XIX, é inegável, mas as informações em relação ao “nível” e a “forma” que estas línguas e estes povos exerceram na configuração singularizadora do português brasileiro é o grande enigma, a grande questão a ser respondida pelos que acreditam que as especificidades do PB resultam de um contato linguístico.

Muitas são as hipóteses que nos auxiliam na busca de evidências plausíveis em torno da reconstituição histórica do PB. Gregory Guy (2005) acredita na transferência de traços da língua materna (no caso dos africanos) para a segunda língua (o português), sendo que as variedades do PB resultam de um processo de criouliização, hipótese também defendida por Holm (1987). Já Petter (2009), reconhece o protagonismo dos africanos, mas não condiciona as variedades do PB a um processo de criouliização, uma vez que, em linhas gerais, defende que variedades do português constituiriam um *continuum* afro-brasileiro. Nesse viés, Lucchesi (2003; 2009) vem paulatinamente comprovando uma nova vertente que busca evidências de que o PB ultrapassa os limites delineados pelos que acreditam em mudanças previsíveis e contínuas inerentes a língua, já que acredita e defende que o contato linguístico resultou em um “processo de transmissão linguística irregular ou criouliização do tipo leve, que estaria na base da formação das atuais variedade populares do português do Brasil” (LUCCHESI, 2009, p. 71-72).

⁵ Termo utilizado por Serafim da Silva Neto (1950) ao se referir as línguas indígenas brasileiras e as línguas africanas trazidas para o Brasil.

2.2.1 Conceituando pidgins e crioulos

Para discutir, especificamente, sobre os conceitos de língua, tomamos como base os pressupostos da Crioulística, ciência que tem focado seus estudos nos processos históricos de contato entre línguas, dos quais são originadas línguas gramaticalmente distintas das envolvidas no processo de contato. Nesta perspectiva, de que a história externa da língua é uma evidência positiva em favor de hipótese crioula, temos trabalhos como os de Coelho⁶ (1880), Silva Neto (1950), Guy (1981), dentre outros, que têm atestado e/ou alegado uma origem crioula para o PB.

Nesse viés, temos os trabalhos de Renato Mendonça e Jacques Raimundo (1933 *apud* BONVINI; PETTER, 1998) que elaboram os debates sobre criouliização, afirmando que as diferenças ocorrentes no PB devem-se a influência das línguas africanas, principalmente o kimbundo e o yorubá (BONVINI; PETTER, 1998, p.1). No que tange as discussões sobre a relevância do contato entre línguas na formação histórica do português do Brasil, Lucchesi (2012, p. 248) afirma que essas questões voltaram à cena na linguística brasileira na década de 1980, com a intervenção de Gregory Guy⁷ (1981) e John Holm⁸ (1987; 1992).

Guy (1981) e Holm (1987) retomam a discussão sobre a hipótese crioula do PB apoiando-se em dados sócio-históricos, mais precisamente demográficos, e na evidência linguística. Para Guy (1981), a hipótese da criouliização apresenta duas evidências, uma histórica e outra linguística. Em relação à evidência histórico-social, ele acredita que o contexto do Brasil era propício para a formação de um crioulo. Assim, desse ponto de vista, acredita que a criouliização do PB existiu, mas que o sistema linguístico brasileiro está se descrioulizando e as forças dessas mudanças linguísticas estão retomando vagarosamente ao seu alvo original. Ainda nessa perspectiva, Guy (1981) acredita que a evidência linguística provém de duas variáveis fonológicas: apagamento do –s e desnasalização de vogais finais; e de duas variáveis morfológicas: concordância no sintagma verbal e concordância no sintagma nominal – objeto de nosso estudo a ser desenvolvido no *Capítulo V*.

⁶ Adolfo Coelho (1886) foi quem inaugurou os debates sobre criouliização, lançando a idéia da origem crioula no PB.

⁷ Guy trabalha com “[...] base em análises variacionistas da fala de indivíduos não escolarizados do Rio de Janeiro, postulando que as condições sociais dos primeiros séculos da colonização do Brasil eram muito favoráveis à criouliização e que o português popular do Brasil seria o resultado de um acelerado processo de descrioulização de um crioulo português que se teria formado no século XVII”. (LUCCHESI, 2012, p. 249)

⁸ Holm “resgatou o conceito de *semicrioulo*, utilizado por Serafim da Silva Neto na década de 1950, defendendo que o português teria passado no Brasil, devido ao contato linguístico, por um processo de *reestruturação parcial*.” (LUCCHESI, 2012, p. 249)

Para iniciarmos a discussão sobre os conceitos de *pidgin* e crioulo, se faz necessário entender como surge um *pidgin*. De forma geral, e baseando-se em pesquisas que tratam sobre o tema, se pode afirmar, ainda que de forma preliminar, que o *pidgin* surge do contato entre línguas mutualmente ininteligíveis e torna-se uma língua auxiliar para comunicação entre indivíduos que cujas línguas maternas são diferentes. Porém, estudiosos a definem de uma forma mais precisa. Naro e Scherre (2003, p. 286) acreditam que

o *pidgin* costuma surgir a partir de um contato entre grupos populacionais para propósitos bem definidos e delimitados, tais como trocas comerciais ou trabalho forçado em uma plantação, em um contexto em que as pessoas não dispõem de outro meio verbal comum de comunicação.

Lucchesi (2003, p. 273) também afirma que um *pidgin* surge

quando uma grande população de adultos – em muitos casos falantes de línguas diferenciadas e mutuamente ininteligíveis – é forçada a adquirir uma segunda língua emergencialmente em função de relações comerciais e/ou de sujeição [...].

Tentaremos explicar melhor este conceito através do processo de formação do português do Brasil. Como já discutido no Capítulo I deste trabalho, o processo de colonização europeia, entre os séculos XVI e XIX, ocasionou várias situações de contato entre povos com diferentes línguas, portanto, houve diversas situações propícias ao surgimento de *pidgin* e/ou línguas crioulas. No encontro entre esses povos/comunidades colonizadas, iniciou-se a busca pelo estabelecimento da comunicação entre os escravos e os senhores e entre os próprios escravos, que até então eram falantes de línguas diversas. Diante da necessidade de se comunicar, os falantes trazidos pelos colonizadores de comunidades/línguas diferentes, buscavam uma linguagem comum.

Esse contato interlinguístico tem sido a fonte para a infundável discussão sobre a origem do português brasileiro, até porque muitos são os conceitos adotados pelos pesquisadores, o que torna mais difícil o objetivo, deste subitem, de conceituar *pidgin* e crioulo.

Iniciaremos expondo que uma língua crioula nasce a partir do contato de comunidades linguísticas diferentes em meio a uma relação que se estabelece entre o grupo dominante com o grupo dominado. Antes de tornar-se uma língua crioula, o resultado do contato entre essas etnias será um *pidgin*, um meio de comunicação, simplificado.

Nessa interação verbal dos indivíduos dos grupos dominados, ocorre uma negociação em torno das soluções gramaticais empregadas para a formação de frases a partir do vocabulário básico da LA compartilhado por eles, e as estruturas gramaticais do *pidgin* vão se formando, na medida em que sejam consagradas pelo uso, dentro da rede de relações sociais que também se vão formando nessa comunidade emergente. Dessa forma, o *pidgin* vai se expandindo, na medida em que as regras gramaticais empregadas no seu uso são chanceladas dentro de uma rede de

relações sociais historicamente determinada. No plano sócio-histórico, os *pidgins* tendem a desenvolver-se nos casos em que os grupos do substrato são mantidos em seu universo cultural de origem e conservam o uso de suas línguas nativas. (LUCCHESI; BAXTER, 2009, p. 105).

Então podemos entender que a partir do momento que começam a nascer crianças na comunidade, elas passam a ter o *pidgin* que seus pais falam como *input* na fase de aquisição da linguagem; o *pidgin* passa então, por ação dessa nova geração de falantes, a ser um crioulo, ou seja, o meio de comunicação utilizado pelos adultos naquela comunidade passou a ser uma língua natural para as crianças daquele meio, que o dotam de uma gramática e assim temos uma língua natural.

Os autores seguem afirmando que a existência de *pidgins* expandidos comprova que a reestruturação gramatical da variedade que se forma na situação de contato pode ocorrer independentemente da sua nativização, o que é tradicionalmente definido como criouliização.

Nesse contexto Lucchesi (2003, p. 103) aponta que

a noção da criouliização variável foi proposta de maneira explícita por Baker (1982) e Bickerton (1984), que afirmam que a criouliização a partir de situações de contato com maior acesso aos modelos da língua alvo deram origem a crioulos gramaticalmente mais próximos da língua alvo.

Mikael Parkvall, em seu livro *Da África para o Atlântico*, livro traduzido no Brasil, em 2012, por Rodolfo Ilari, afirma que muito se tem discutido a respeito da distinção entre *pidgins* e crioulos, e nas discussões tem sido reconhecido

cada vez mais que a nativização não é uma condição *sine qua nom* da criouliização. Hoje em dia, costuma-se admitir que os *pidgins* podem expandir-se de modo a formar línguas plenas mediante apenas um uso frequente, e foi com base nisso que os ‘*pidgins*’ por exemplo da Nigéria, de Camarões e da Guiné-Bissau foram tratados como crioulos e não como *pidgin*. (PARKVALL, 2012, p. 33)

Assim entendemos que os *pidgins* e crioulos se formam a partir de mecanismos radicais de reestruturação da gramática original da língua alvo e de transferência, para esta, de estruturas provenientes das línguas dos segmentos dominados (LUCCHESI, 2003, p. 272). Porém, Baxter (1994) e Lucchesi (2003, p. 273) afirmam que os contextos sócio-históricos em que ocorre contato entre línguas nem sempre dão margem à implementação de mecanismos mais radicais de reconfiguração de uma língua. Para ele, as causas da redução na estrutura gramatical da língua de superstrato, advindas do contato com outras línguas, são três:

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

- b) o fato de os falantes dessas outras línguas serem, em sua grande maioria, adultos, não havendo, pois, o acesso aos dispositivos da *faculté du langage*, que atuam naturalmente no processo de aquisição da língua materna;
- c) a ausência de uma ação normatizadora, ou seja, de uma norma ideal que oriente e restrinja o processo de aquisição/nativização, já que esse processo tem como objetivo, fundamentalmente, a comunicação emergencial com os falantes da língua alvo. (LUCCHESI, 2003, p. 273):

Dessa forma, acredita-se que língua crioula é a evolução do código emergencial implementado no início do contato linguístico, ou seja, é a nativização deste, propiciada pela socialização do código emergencial entre os segmentos dominados. De acordo com Lucchesi (2003, p. 273), este último “vai se convertendo em modelo para a aquisição da língua materna dos descendentes dos falantes das outras línguas”, para cujo *status* ocorre uma expansão gramatical ou por meio da gramaticalização de itens lexicais originários da língua alvo ou por mecanismos gramaticais disponíveis nas línguas de substrato.

No que tange aos critérios linguísticos e estruturais Naro e Scherre (2003, p. 286) acreditam que

normalmente, o léxico do pidgin se baseia na língua do grupo socialmente dominante, mas, pelo menos no início do processo, as estruturas empregadas podem variar de falante para falante, de acordo com seus conhecimentos linguísticos anteriores e a experiência colhida em suas interações com outros falantes na situação de contato.

Parkvall (2012, p. 43) relata a dificuldade em compreender como funciona o contato linguístico e a reestruturação das funções. Ao discutir as características linguísticas que compõem um crioulo, o autor afirma que deveria ser consensual afirmar que essas características deveriam ter uma linha de princípio, com uma ou mais das quatro origens:

- (i) A língua lexificadora;
- (ii) Os substratos;
- (iii) Os universais da reestruturação;
- (iv) O desenvolvimento independente, incluindo mudanças posteriores à cristalização, quer as motivadas internamente, quer as que resultaram de influência do adstrato. (PARKVALL, 2012, p. 43-44)

Parkvall (2012, p. 56) ainda expõe que os crioulistas precisam considerar não só a língua lexificadora e o substrato, mas também as propriedades gerais dos *pidgins* e crioulos e as descobertas da linguística tipológica e da linguística histórica. Portanto, o autor propõe divisões que considera essencial para a compreensão da reestruturação linguística e os resultados de um contato linguístico. A saber:

- Que um traço linguístico de um P/C seja encarado como uma *conservação da língua lexificadora* se (= se e somente se) estiver presente na língua

lexificadora e ausente dos substratos, e se for incomum interlinguisticamente e não geralmente presentes em outros P/Cs não relacionados;

- Que estamos lidando com uma *transferência a partir do substrato* se o traço estiver presente nos substratos, ausente na língua lexificadora e se for translinguisticamente incomum e não geralmente presente em outros P/Cs não relacionados;
- Analogamente, um traço é um *universal de reestruturação* e se está ausente tanto na língua lexificadora quanto nos P/Cs não relacionados;
- *Um caso de desenvolvimento independente*, por fim, caracteriza-se por estar ausente de todos os componentes input, assim como de outros P/Cs não relacionados. (PARKVALL, 2012, p. 56-57 – grifos do autor)

Porém, ao tratar do processo de criouliização, muitos pesquisadores discordam veementemente, dentre esses Dante Lucchesi (2009, p. 62), de que, em um raciocínio lógico/matemático, diz que “o nível de criouliização é inversamente proporcional à facilidade de acesso aos modelos da língua alvo”. O que implica dizer que quanto menor o número de falantes da língua dominante, o português europeu, maior será a chance de ocorrer criouliização, mas o autor afirma que esse não é o contexto demográfico do Brasil colônia, uma vez que

[...] a população branca no Brasil nunca foi inferior a 30%, crescendo significativamente no século XIX, quando chega a quase a metade do total. Esse quadro está bastante distante do que deu ensejo às formas mais típicas de criouliização, ocorridas no Caribe. No Haiti e na Jamaica, a proporção de brancos nunca ultrapassou os 10% durante praticamente todo o período da colonização, e o nível de miscigenação entre brancos e negros foi muito mais baixo do que o observado no Brasil. Por tanto, [...] *o acesso aos modelos da língua-alvo no Brasil sempre foi maior do que o que se observa nas situações típicas de criouliização.* (LUCCHESI, 2009, p. 62-63 – grifos do autor)

Entre esses debates referentes a caracterização do português brasileiro, surge uma nova hipótese que se encontra no rol das discussões sobre a origem do português brasileiro, e reconhece a formação do PB como fruto de um processo de transmissão linguística irregular em que o contato do PE com diferentes línguas no Brasil, mas diretamente as africanas e indígenas, permitiu o surgimento de uma variedade da língua portuguesa, como discutiremos a seguir, em 2.2.2.

Lucchesi (2012, p. 52) acredita que desenvolvimentos teóricos mais recentes da *crioulística*, o ramo da linguística que focaliza as situações radicais de contato entre línguas, podem ter um papel decisivo para ampliar a compreensão acerca da mudança linguística, desde que devidamente articulados com os achados empíricos advindos da pesquisa de campo sobre variedades populares de línguas transplantadas no processo histórico da expansão colonial europeia. Dessa forma, as reflexões da crioulística não se circunscreveriam apenas às situações específicas e bem localizadas de pidginização e criouliização, abarcando também as

situações de contato massivo que estão presentes na história da grande maioria das línguas do mundo.

O autor (2012, p. 53) ainda afirma que é preciso compreender como os processos de mudança que atuam na pidginização/crioulização podem atuar também em situações de contato cujo resultado não é necessariamente a formação de uma língua pidgin ou crioula. Com esse intuito, Lucchesi (2012) afirma que tem desenvolvido nos últimos quinze anos o conceito de *transmissão linguística irregular* (LUCCHESI 2003, 2004 e 2008; LUCCHESI; BAXTER 2009; BAXTER; LUCCHESI, 1997) e acredita que através desta hipótese podemos prover a formação histórica do português brasileiro de uma explicação mais satisfatória.

2.2.2 Transmissão linguística irregular

Através do conceito de transmissão linguística irregular tenta-se delinear os parâmetros sócio-históricos e linguísticos que serviram de base para uma visão sistemática dos processos históricos de mudança induzidas pelo contato entre línguas que não resultam na formação de pidgins e crioulos típicos (LUCCHESI, 2003, p. 272). Nesse sentido, Lucchesi (2012, p. 254) acredita que a ideia de crioulização leve ou de uma semicrioulização, discutida no subitem anterior, encontra a sua formalização teórica no conceito de transmissão linguística irregular⁹, uma vez que pretende-se dar conta dos processos de mudança pelos quais a língua portuguesa teria passado no Brasil em função do contato entre línguas.

Antes de adentrarmos no conceito e caracterização da transmissão linguística irregular, é importante esclarecer que esse é um conceito considerado mais amplo do que o de pidginização e crioulização, uma vez que aborda tanto os processos de mudança resultantes do contato entre línguas que originam uma língua nova – *pidgin* ou *crioulo* –, quanto os processos nos quais uma língua sofre alterações decorrentes do contato com outras línguas, sem que essas alterações cheguem a configurar o surgimento de uma nova língua.

Portanto, para dar conta das especificidades de português popular brasileiro, entende-se o conceito de transmissão linguística irregular como algo que designa “[...] os processos históricos de contato massivo e prolongado entre línguas, nos quais a *língua do segmento, que detém o poder político é tomada como modelo ou referência para os demais segmentos*”

⁹ Lucchesi (2012, p. 254) afirma que “o conceito de *transmissão linguística irregular* guarda grandes similaridades e alguma complementariedade com o conceito de *reestruturação linguística parcial* de John Holm (2004)” (grifos do autor).

(LUCCHESI, 2003, p. 272 – grifo meu). A partir deste conceito preliminar poderemos iniciar as discussões em torno da caracterização do processo de transmissão linguística irregular do PB.

Talvez surjam dúvidas sobre como tal processo pode produzir variações na língua receptora. Nesse sentido, Lucchesi e Baxter (2009, p. 101) nos explicam que

tal processo de nativização da língua dominante ocorre de maneira irregular no sentido de que os dados linguísticos primários de que as crianças que nascem nessas situações dispõem para desenvolver a sua língua materna provêm praticamente de versões de segunda língua desenvolvidas entre os falantes adultos das outras línguas, que apresentam lacunas e reanálises em relação aos seus mecanismos gramaticais. Tal processo diferencia-se da situação de transmissão geracional normal das línguas humanas, em que as crianças dispõem de dados linguísticos mais completos fornecidos pela língua materna dos seus pais.

Nesse contexto, há duas possibilidades que podem ser geradas por uma transmissão linguística irregular, a formação de uma língua crioula ou uma nova variedade da língua de superstrato, que é representado pelos processos de variação e mudança induzidos pelo contato entre línguas (LUCCHESI; BAXTER, 2009).

No Brasil, a língua portuguesa, representava superioridade econômica e política, era o caminho para uma ascensão social. Nesse intuito, muitos colonizados viram na língua de “prestígio” uma chance para ascensão social, porém adquirir uma nova língua, naquele contexto social e multicultural, não seria tão simples. Dessa forma, a língua portuguesa chega até os colonizados/escravizados através das lentes deformadas, já apresentando algumas alterações morfossintáticas, resultando numa variedade da língua alvo, como atestam alguns trabalhos de Lucchesi (2001; 2003; 2009), que afirma que

prolongando-se a situação de contato, essa variedade segunda da língua alvo, por ser socialmente mais viável (em função de sua relação com os segmentos política e economicamente dominantes e de maior prestígio cultural), vai progressivamente assumindo novas funções na rede de interação linguística, ao tempo em que se vai convertendo em modelo para a aquisição da língua materna dos descendentes dos falantes das outras línguas. (LUCCHESI, 2003, p. 273)

Lucchesi (2009) afirma que na história sociolinguística do Brasil, não ocorreram processos que possam ser caracterizados como crioulação. Portanto, tal afirmação dá margem a entender que o contato entre línguas, no período colonial, deu origem a uma variedade de língua a ser explicada pelo processo de transmissão linguística irregular. Dessa forma, Lucchesi (2009, p. 71) explica que

a aquisição precária do português pelos escravos trazidos da África e pelos índios integrados na sociedade brasileira e a nativização desse modelo defectivo de português como língua segunda nas gerações seguintes de seus descendentes

endógamos e mestiços desencadearam um processo de transmissão linguística irregular que teve importantes consequências para a formação da atual realidade linguística brasileira, nomeadamente para as suas variedades populares.

O conceito de transmissão linguística irregular tomado por Lucchesi e Baxter (2009) busca analisar de forma ampla os processos históricos de “contato maciço entre povos falantes de línguas tipologicamente diferenciadas, entre os séculos XVI e XIX, em decorrência da ação do colonialismo” (p. 101). Portanto, torna-se inevitável retroceder ao contexto sociolinguístico no período colonial para que se entenda como se configura essa situação de contato e como o PB pode ter sido originado da transmissão linguística irregular.

Como já retratado no capítulo primeiro, o Brasil se apresenta inserido em um contexto plurilinguístico, em que havia a língua do dominador, o português europeu, denominado língua de *superstrato* ou *língua-alvo*, e as muitas línguas dos dominados (*cf.* subitem 1.1; 1.2), denominadas línguas de *substrato*. Como em todo processo de colonização, a língua de superstrato se impõe às demais línguas, de modo que “[...] os falantes das outras línguas, em sua maioria adultos, são forçados a adquiri-la em condições bastante adversas de aprendizado, em função de sua sujeição e marginalização” (LUCCHESI; BAXTER, 2009, p. 101). Os autores ainda afirmam que essas aquisições “imperfeitas” de segunda língua, e nas condições que foram adquiridas, acabam por fornecer modelos de aquisição da língua materna para as novas gerações de falantes. Isso se configura no conceito do que Lucchesi (2003; 2009) denominará de transmissão linguística irregular.

Essa variedade de língua alvo, que se forma de contato plurilinguístico apresenta simplificações em sua estrutura gramatical. Lucchesi (2009) acredita que da língua alvo se mantêm os elementos essenciais necessários ao preenchimento das funções comunicativas básicas, no mais apresentam altos índices de variação em relação ao modelo de língua alvo.

Em linhas gerais, Lucchesi e Baxter (2009) postulam duas fases presentes no processo de transmissão linguística irregular. Vejamos:

- (i) Fase inicial: variação, mais ou menos intensa, no uso dos mecanismos gramaticais, sendo mais rara a eliminação desses mecanismos; opacidade relativa nos estímulos-gatilhos que possibilitam a aquisição de valores diferenciados nos parâmetros sintáticos;
- (i) Fases seguintes: manutenção do quadro de variação, em que se observa a concorrência entre as formas gramaticais reintroduzidas a partir da influência da LA e potenciais processos de gramaticalização que emergem na heterogeneidade da fala; variação nas frequências de uso dos parâmetros sintáticos. (LUCCHESI; BAXTER, 2009, p. 123)

Lucchesi (2003, p. 276) apresenta três características fundamentais do processo de transmissão linguística irregular:

- (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.

As mais recentes pesquisas no campo da transmissão linguística irregular tem demonstrado que esse é um amplo conceito que tem sido capaz de abarcar e interpretar muitos processos históricos em que o português sofreu alterações/variações. Dessa forma, no campo da ciência da linguagem, torna-se essencial buscar evidências/fatos que venham a ratificar sua importância no contexto de origem do português brasileiro.

3 AS DIVERSAS ABORDAGENS DA CONCORDÂNCIA NOMINAL DE NÚMERO

*“Eu prefiro ser
Essa metamorfose ambulante
Do que ter aquela velha opinião
Formada sobre tudo.”*

(Metamorfose ambulante - Raul Seixas)

Neste capítulo busca-se discutir sobre o fenômeno em análise *a variação na concordância de número no sintagma nominal*, com base nas abordagens formalista e funcionalista. De forma concomitante, pretende-se fazer um levantamento de trabalhos já realizados na área a fim de atestar resultados referentes ao fenômeno, que possam contribuir para as análises a serem realizadas no *capítulo V* desta dissertação.

A variação na concordância de número no português falado no Brasil tem sido intensamente documentada e estudada no âmbito da sociolinguística, oferecendo, desse modo, a oportunidade de evidenciar as perdas morfológicas tão típicas do português brasileiro.

Iniciaremos este capítulo, na tentativa de alcançar o objetivo proposto, apresentando os mecanismos gramaticais de concordância da língua portuguesa como elencados abaixo:

- (i) Gênero – especificamente nominal (**o** amigo / **a** amiga)
- (ii) Pessoa – especificamente verbal (**nós** conversamos / **eles** conversam)
- (iii) Número – Nominal ou verbal.

Dentre os mecanismos citados, destacamos a concordância de número, que por sua vez pode ser nominal ou verbal. A verbal ocorre entre o sintagma nominal (doravante SN) sujeito e o verbo (ele estuda / **eles** estudam), a nominal ocorre entre os elementos flexionáveis do SN, como em (1), ou entre o SN sujeito e o predicativo, como em (2):

Os elementos constituintes do Sintagma Nominal:

(1) “**Meus pais** moram comigo.”¹⁰



O Sintagma Nominal sujeito e o predicativo do sujeito:

(2) “**Eles** tão se sentido **incomodados**.”



A partir dos exemplos supracitados (1) e (2), é possível afirmar que a concordância nominal de número ocorre a partir da inserção da marca de plural -s em todos os elementos flexionáveis do sintagma nominal. Portanto, a marca formal de plural nos nomes em português é representada pelo -s, considerando que a forma regular de se fazer este tipo de plural no PB é através da adição de um segmento fricativo não labial, cujas realizações podem ser [s], [z], [š], [ž] ou [h], estas realizações ocorrem em função de contextos linguísticos específicos ou em função de características regionais dos falantes do português do Brasil (SCHERRE, 1996, p. 87).

Mesmo diante dos muitos estudos realizados sobre o fenômeno da concordância nominal de número (cf. Figura 4) a tradição gramatical brasileira apresenta-a como uso obrigatório, atribuindo a não marcação de plural em todos os elementos flexionáveis do SN alguns estigmas, como veremos no subitem posterior.

Graças aos avanços dos estudos dialectológicos e sociolinguísticos é possível contrapor as imposições da tradição gramatical, uma vez que esses apresentam argumentos que comprovam que o fenômeno da concordância de número no SN parece não ser tão obrigatório, quando se trata da oralidade, especificamente quando se trata dos falantes do português brasileiro. Portanto, no PB contemporâneo, a concordância nominal de número dentro do sintagma nominal pode apresentar-se através de três possibilidades:

Inserção da marca de plural em todos os elementos do SN

(3) “As crianças também aprendem.”

¹⁰ Todas as frases apresentadas neste trabalho, a título de exemplificação, são retiradas dos *corpora* a serem analisados no Capítulo V.

Marcação de alguns elementos do SN

(4) “Gosto *das minhas coisa¹¹*.”

Marcação de plural em um único elemento do SN

(5) “Estava com *as perna^Ø inchada^Ø*.”

Isso implica dizer que, ao se tratar do português brasileiro, a concordância nominal de número não é tão invariável como afirmam algumas gramáticas, uma vez que pesquisas realizadas sobre os dialetos brasileiros vêm demonstrando que esse fenômeno não é uma regra categórica, mas sim variável conforme certos determinantes linguísticos e sociais (SCHERRE, 1994).

Assim, ao se tratar da variação na concordância de número no SN, podemos afirmar a existência de duas variantes interligadas ao fenômeno: a variante explícita (marcada), e a variante zero de plural (não marcada). Vejamos através dos exemplos (6) e (7):

Variante explícita:

(6) “Ele fez **as estações**.”

Variante zero de plural:

(7) “Eu ia com **alguns colega^Ø**.”

Questões como a concordância (nominal e verbal) têm sido discutidas sob a perspectiva da funcionalidade das marcas linguísticas, que trataremos melhor no subitem posterior. No entanto, a partir dos exemplos (6) e (7), percebemos que no português do Brasil é possível eliminar algumas marcas de concordância dentro do SN, respeitando o princípio da economia, preservando as marcas necessárias e eliminando as redundantes, como em (7); ou, pelo contrário, preservar as marcas redundantes, como em (6).

Diante do exposto podemos deduzir, ainda que de forma superficial, que o português vernacular do Brasil apresenta variações nos processos de concordância de número, o que o distancia da visão de homogeneidade e de rigidez pregadas pelas gramáticas tradicionais que

¹¹ A simbologia \emptyset utilizada neste trabalho, reflete a confluência formal entre o zero plural e a morfologia do singular, caracterizada pelo morfema zero, como define Scherre (2006, p. 107).

tomam como base o português padrão de Portugal. Essa visão tradicional tem favorecido a propagação da ideia que a variante explícita de plural é a variante de prestígio, enquanto a variante zero (não marcada) é imprópria, ou simplesmente errada e restrita a grupos que “não sabem falar o português”. Porém, diante das pesquisas realizadas no campo sociolinguístico, já é incontestável que a concordância de número realiza-se variavelmente tanto nos dialetos populares, quanto na chamada norma culta, no Brasil.

Desde os primeiros estudos sobre a variação na concordância nominal de número no português do Brasil (*cf.* subitem 3.5.1) tem-se observado que os falantes têm optado por marcar apenas o primeiro constituinte pluralizável do SN, o que segundo Gonçalves (2000) tem favorecido o enfraquecimento da morfologia nominal. No que tange ao português africano, especificamente o português de Cabo Verde, Jon-And (2010, p. 40) também afirma que a “[...] posição pré-nuclear adjacente é essencial e favorece fortemente a marcação de plural”.

Mesmo diante dos estudos que comprovam a existência da variação na concordância nominal no PB, muitos falantes ainda são temerosos ao assumirem esse modo peculiar de se expressar. Dificilmente, encontra-se um falante brasileiro que assuma a não marcação de plural em todos os elementos flexionáveis do SN, uma vez que tal estratégia foge do padrão elitizado/canônico. Nesse sentido, Perini (2003, p. 19) afirma que

sabemos que é freqüente, entre nós, fazer o plural marcando-o apenas no primeiro elemento do sintagma [...]. Essas construções, quando não são simplesmente ignoradas, são dadas como da linguagem das ‘pessoas incultas’, ou de ‘baixa classe’. Segundo essa opinião, não se trataria de um fato normal do português brasileiro, mas de um ‘erro’ cometido por aquelas pessoas (coitadas) que não tiveram a sorte de uma educação formal suficiente.

O autor segue afirmando que levantamentos mostram o contrário, pois construções em que se deixa de marcar alguns dos elementos do SN são amplamente utilizadas pelos falantes do Português Brasileiro, incluindo os “cultos”. Entende-se, então, que tal fenômeno é característico da oralidade sem exclusões de grupos minoritários e/ou majoritários.

Dessa forma, torna-se evidente que a ausência da marca de plural é comum em todas as comunidades brasileiras, condicionada por fatores linguísticos e extralinguísticos, mas cabe ressaltar que, assim como citado acima, alguns fatores extralinguísticos, como pouco grau de escolaridade, tendem a favorecer à marca zero de plural, o que não implica dizer que tal fenômeno se restrinja a tal variável social. Nesse sentido, “[...] sabe-se que fatores de cunho extralinguístico têm um peso considerável na maior ou menor frequência dos padrões

canônicos e não canônicos, como se depreende de estudos sobre o PB [...]” (BRANDÃO; VIEIRA, 2012, p. 9).

O que já foi exposto permite perceber que a variação no uso das regras de concordância é amplamente estudada, portanto é natural que muitas vezes torne-se tema de discussão quando se trata da gênese do português do Brasil, tanto por aqueles que veem as características do português do Brasil como o resultado de um processo secular de deriva apenas intensificado pelo contato do português com as línguas indígenas e africanas (NARO; SCHERRE, 2007), quanto por aqueles que relacionam os processos de variação que se observam hoje no português do Brasil, sobretudo em suas variedades populares, ao amplo e massivo contato entre línguas que marca a história sociolinguística do país (BAXTER; LUCCHESI, 1997; LUCCHESI, 2003, 2008).

Pensar na variação da concordância de número no SN é entender que o português falado no Brasil apresenta características peculiares ao ser comparado com o português europeu, características essas que incitam pesquisas, não apenas no âmbito do PB, mas pesquisas como esta, que trace paralelos entre o PB e outras variedades do português, neste caso o português falado em Angola. Só assim começaremos a trilhar o caminho que nos conduzirá a resultados que comprovem que o processo de concordância nominal é, fazendo alusão ao autor de nossa epígrafe, *uma metamorfose ambulante* – português vernacular – e não aquela velha opinião formada sobre tudo – gramática tradicional.

3.1 CONCORDÂNCIA NOMINAL: PERSPECTIVA TRADICIONAL E FORMAL

Recentes pesquisas vêm demonstrando uma forte tendência da não marcação de plural no SN, mas esse fenômeno ainda não é descrito pelos compêndios gramaticais. Para discorrer sobre tal processo, torna-se necessário elencar alguns conceitos encontrados nesses compêndios, no que tange à concordância de número no SN.

Em geral, a concordância é compreendida como uma relação de solidariedade entre os elementos de um dado enunciado, com a inserção da marca de plural -s em todos os elementos flexionáveis do sintagma. Para Dubois *et al.* (1973, p. 136), concordância é

o fenômeno sintático pelo qual um substantivo ou um pronome pode exercer pressão de alteração formal sobre os pronomes que o representam, os verbos de que ele é sujeito, e os adjetivos ou participios que a ele se referem.

Dentre os autores renomados, temos Evanildo Bechara, autor de várias das principais gramáticas da língua portuguesa, que tenta sintetizar a concordância de número no SN, como:

Diz-se *concordância nominal* a que se verifica em gênero e número entre o adjetivo e o pronome (adjetivo), o artigo, o numeral ou o particípio (palavras determinantes) e o substantivo ou pronome (palavras determinadas) a que se referem. (BECHARA, 2009, p. 543 - grifos do autor)

Após a definição essencialmente gramatical para *concordância nominal*, Bechara (2009, p. 544-554) aponta uma série de contextos de ordem formal e semântica em que vocábulos determinantes subordinam-se aos seus núcleos, mediante a ação da concordância. Os contextos são os seguintes:

A. CONCORDÂNCIA DE PALAVRA PARA PALAVRA

- (i) *Há uma só palavra determinada* – a palavra determinante irá para o gênero e número da palavra determinada;
- (ii) *Há mais de uma palavra determinada:*
 - Se as palavras determinadas forem do mesmo gênero, a palavra determinante para o plural e para o gênero comum, ou poderá concordar, principalmente se vier anteposta, em gênero e número com a mais próxima;
 - Se as palavras determinadas forem de gêneros diferentes, a palavra determinante irá para o plural masculino ou concordará em gênero e número com a mais próxima.
- (iii) *Há uma só palavra determinada e mais de uma determinante* – a palavra determinada irá para o plural ou ficará no singular, sendo, neste último caso, facultativa a repetição do artigo.

B. CONCORDÂNCIA DE PALAVRA PELO SENTIDO – a palavra determinante pode deixar de concordar em gênero e número com a *forma* da palavra determinada para levar em consideração, apenas, o *sentido* em que esta se aplica;

C. OUTROS CASOS DE CONCORDÂNCIA NOMINAL

Até o momento parece estar tudo muito bem estruturado aos olhos da gramática tradicional, mas cabem algumas contestações a respeito da eficácia e viabilidade dessas conceituações/regras apresentadas, por muitos gramáticos, em especial pelo autor da *Moderna Gramática Portuguesa*, de que nos servimos para uma reflexão. Vamos pensar juntos:

- (i) Qual o real sentido de se estabelecer regras, aparentemente fixas, e em seguida de forma bastante tímida apresentar as observações, ou as tradicionais e constantes *exceções*?
- (ii) Por que afirmar ao leitor a existência de três contextos referentes à *concordância nominal*, sendo que um deles (*Outros casos de concordância*) aglomera 21 (vinte e um) contextos específicos?

- (iii) Será que os gramáticos vão conseguir categorizar e “gramaticalizar” um fenômeno tão variável quanto a *concordância nominal de número* ou caminharão para o “mundo sem fim” das exceções do português do Brasil?

A primeira reflexão surge das muitas exceções apresentadas pelas gramáticas tradicionais. Na *Moderna Gramática Portuguesa* podemos observar que após conceituar a *concordância de palavra para palavra*, em que há uma só palavra *determinada*, o autor traz a seguinte informação:

OBSERVAÇÃO: Os nomes femininos como *sentinela*, *guarda*, *guia* e assemelhados, quando aplicados a pessoas do sexo masculino, mantêm o gênero feminino, e levam para este gênero os determinantes a eles referidos: a *sentinela avançada*. (BECHARA, 2009, p. 545 – grifos do autor)

Quando se trata da *concordância de palavra para palavra*, em que há *mais de uma palavra determinada* o autor apresenta mais três “observações”, uma delas diz: “Se as palavras determinadas se referirem a uma só pessoa ou coisa, impõe-se o singular do determinante” (BECHARA, 2009, p. 545). Essas “observações” vão se fazendo presentes em boa parte dos contextos de ordem formal e semântica apresentados por Bechara (2009), porém cabe ressaltar que o fenômeno das exceções gramaticais não se restringem a essa gramática específica, mas sim a toda e qualquer gramática que tente tratar o português como homogêneo.

A segunda reflexão nos faz entender que a intenção de aglomerar 21 (vinte e um) contextos específicos em apenas um, faz com que se evite defini-las como EXEÇÕES, uma vez que iria transparecer ao leitor que, no caso da concordância nominal, há mais exceções do que regras.

A terceira e última reflexão incorre no grande debate existente entre os tradicionalistas e os sociolinguistas, sobre a variabilidade inerente a língua falada, particularmente, ao se tratar do português falado no Brasil.

No que tange a visão formalista temos Câmara Jr. (2000), linguista que desenvolveu muitos trabalhos e inseriu-se na corrente da linguística estruturalista. Nessa perspectiva dita formal, a língua é analisada fora de seu contexto de uso, considerando-se apenas aspectos internos. No que tange a concordância acredita que essa ocorre mediante uma “coerção formal”, entendendo que esse é um fenômeno exclusivamente estrutural. Câmara Jr. (2000, p. 31), acredita que ao se tratar da concordância entende-se que “[...] na flexão há obrigatoriedade e sistematização coerente. Ela é imposta pela própria natureza da frase”.

A variação no uso do plural é abordada pelo autor, mas apenas para alertar sobre a realização do arquifonema /S/, o qual se atualiza, na oralidade, através das fricativas não-labiais ou sibilantes, /s/ - /z/, ou chiantes /s'/ - /z'/, atualizações essas não representadas na grafia do fonema indicador de plural (-s) pontuando que tais variantes estão relacionadas à região do falante e são, sobretudo, alomorfas fonologicamente condicionadas.

Diante de uma superficial amostragem das concepções formalistas, a noção de concordância tem a ver com a concepção de língua como objeto invariável. Atribui-se, então, à estrutura da língua toda responsabilidade pelo seu funcionamento.

Enquanto os estudos formalistas baseiam-se na estrutura da língua, a gramática normativa baseia-se no registro escrito – o literário, cuja função é o exercício estético da linguagem – e, portanto não diz respeito a padrões reais de uso. Desse modo, tais formas de se explicar o fenômeno da concordância nominal de número são pouco elucidativas no que se refere ao seu emprego enquanto prática social.

3.2 CONCORDÂNCIA NOMINAL DE NÚMERO VARIÁVEL NO PORTUGUÊS DO BRASIL E DE ANGOLA

3.2.1 Concordância nominal de número variável no português do Brasil

A partir da década de 70 do século XX, o linguísta norte-americano Anthony Julius Naro inicia o que designou de *cruzada sociolinguística*. Naro não só implementou a Sociolinguística Quantitativa no Brasil, como se fez e se faz presente nos mais diversos estudos desenvolvidos nessa perspectiva, e não seria diferente com esse trabalho a ser desenvolvido no intuito de traçar paralelos no que tange a variação na concordância nominal de número nas variedades do português do Brasil e de Angola.

Desde os primeiros estudos sobre concordância nominal desenvolvidos no Brasil foi possível confirmar o que Teyssier (1997, p. 98) já apontava quando afirmou que as divisões dialetais no Brasil são menos geográficas que socioculturais.

No arcabouço da Teoria da Variação e Mudança linguística, linha laboviana, vários fenômenos que se enquadram nas mais diversas esferas, a citar: fonética/fonologia, morfossintaxe, sintaxe. Dentre essa é na morfossintaxe que se enquadra o fenômeno que nos propomos a analisar, a variação na concordância nominal de número. Diante dos muitos trabalhos que tratam desse fenômeno, nas mais diversas regiões do país, como poderão

observar na *Figura 4*, pesquisadores afirmam ser este um aspecto “bem-resolvido” no conjunto de fenômenos linguísticos variáveis no português do Brasil.

Scherre (1994) afirma que as análises que envolvem a variação da concordância de número no sintagma nominal têm início com o estudo desenvolvido por Jeroslow, em 1974, que abarca uma pesquisa na área rural cearense, e faz uma descrição estruturalista dos diversos subsistemas linguísticos do falar dessa área e, dentre os diversos aspectos descritos, registra-se a concordância de número verbal e nominal, mas esses estudos ainda não tinham o aporte teórico metodológico laboviano. Sendo assim, destacam-se como pioneiros os estudos realizados por Braga e Scherre, em 1976, que analisaram dados de sete falantes do Rio de Janeiro, de classe social e procedência geográfica distintas; o de Braga (1977) que desenvolve sua dissertação com dados de sete falantes da classe média e baixa do Triângulo Mineiro, com três níveis de escolaridade; o de Lemle e Naro (1977) que utilizando os princípios labovianos da teoria da variação, inicia a pesquisa sobre concordância verbo/sujeito com 20 falantes semi-escolarizados da zona urbana Rio de Janeiro; o de Scherre (1978) que analisou, em sua dissertação de mestrado, dez falantes da zona urbana Rio de Janeiro, com diferentes níveis de escolaridade.

Esses e outros estudos, posteriormente desenvolvidos, evidenciam

[...] que o fenômeno da variação na concordância de número no português falado do Brasil, longe de ser restrito a uma região ou classe social específica, é característico de toda a comunidade de fala brasileira, apresentando diferenças mais de grau do que de princípio, ou seja, as diferenças são mais relativas à quantidade de marcas de plural e não aos contextos linguísticos nos quais a variação ocorre. (SCHERRE; NARO, 1993, p. 38)

Em um breve panorama dos estudos sobre a variação na concordância nominal desenvolvido no Brasil, temos:

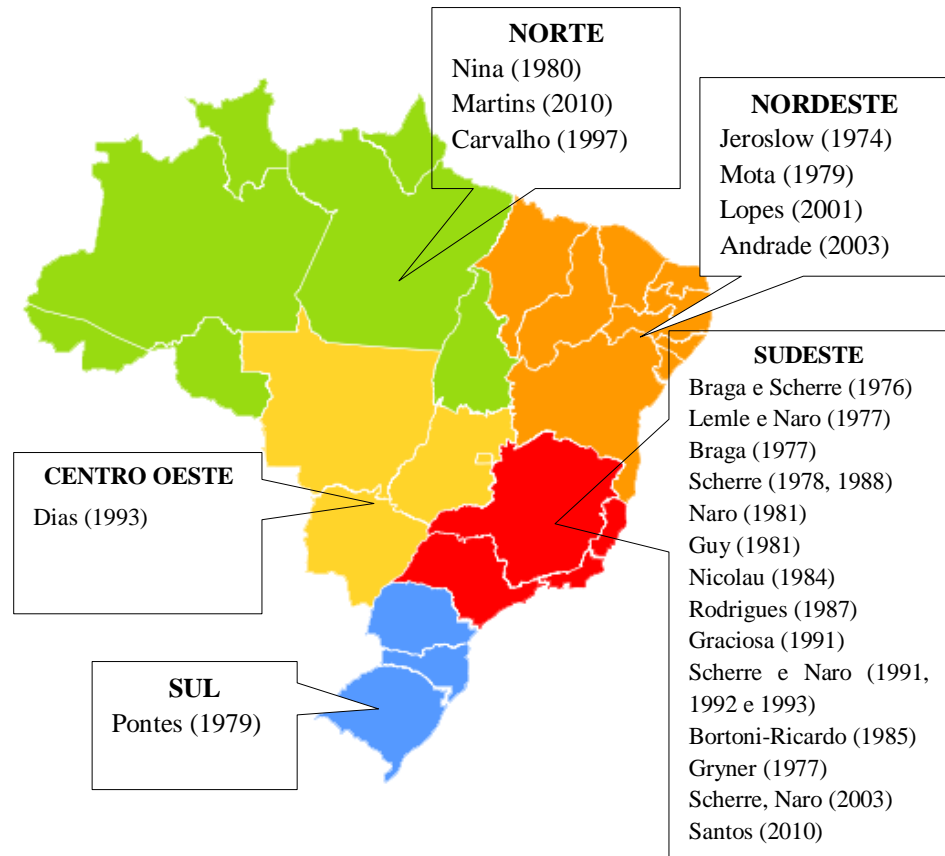


Figura 4: Panorama dos estudos sobre variação na concordância nominal desenvolvidos no Brasil

Os trabalhos registrados na *Figura 4* não representam todos os trabalhos já desenvolvidos no Brasil sobre variação na concordância nominal, contudo, são uma evidência de que esse é um fenômeno recorrente no português falado do Brasil, não se restringindo a uma região ou classe social específica, como ressaltamos no início deste capítulo, mas sendo característico de toda a comunidade de fala brasileira.

Entre os trabalhos realizados por sociolinguistas no Brasil, merecem destaque a tese de doutorado de Scherre (1988)¹², intitulada *Reanálise da concordância nominal em português*, que hoje é considerada como uma obra de referência ao se tratar da variação na concordância nominal de número, já que através desse trabalho a autora analisa todos os mecanismos que regem o fenômeno da variação da concordância no SN no português falado no Rio de Janeiro, analisando-o de forma criteriosa na duas abordagens de análise, a atomística ou mórfica e a sintagmática ou não-mórfica, termos de que trataremos no capítulo sobre a metodologia

¹² Analisa a concordância nominal no PB com dados extraídos do *corpus* Censo do Programa de Estudos sobre o Uso da Língua (PEUL), grupo de pesquisa vinculado ao Departamento de Linguística e Filologia da Faculdade de Letras da Universidade Federal do Rio de Janeiro – UFRJ.

adotada no presente trabalho. Em suas considerações finais Scherre (1988, p. 277-278) apresenta os fatores que condicionam ou não a variação da concordância no SN, nas perspectivas de análise supracitadas¹³.

Essa tese de doutorado serviu e serve como base e/ou apoio bibliográfico para pesquisas posteriores que trataram e/ou tratam da concordância nominal. Nessa sessão, daremos destaque a outros trabalhos, primeiramente à tese de Lopes (2001)¹⁴, intitulada *Concordância nominal, contexto linguístico e sociedade*, que analisa a variação da concordância de número no SN nas falas popular e culta na cidade de Salvador-Ba, em quatro diferentes faixas etárias, três graus de escolarização, nos dois gêneros, e em dois diferentes grupos étnicos, identificando fatores linguísticos e sociais que mais condicionam a sua variação. Lopes (2001) também trabalhou com amostras de fala do português popular, a partir do banco de dados do Programa de Estudo do Português Popular de Salvador (PEPP).

Diferentemente de Scherre (1988), Lopes (2001) trabalhou apenas na perspectiva atomística, apresentando também os fatores que condicionam o uso padrão, marcado, para a concordância de número no SN. A autora chega às seguintes conclusões:

- (i) São mais alvo de concordância elementos com mais material fônico na oposição singular/plural;
- (ii) Elementos à esquerda do núcleo recebem mais marca de plural que elementos nucleares;
- (iii) Elementos à direita do núcleo têm menos probabilidade de receberem marca de plural;
- (iv) Contextos antecedentes com numeral promovem mais concordância subsequente em segunda posição que contextos com marca formal antecedente;
- (v) Elementos em 3a, 4a ou 5a posição tendem a fazer mais concordância quando são antecedidos por mais de uma marca do que quando antecedidos por zero.
- (vi) Elementos são mais alvo de plural quando finalizam a sentença;
- (vii) Elementos com contexto subsequente com vogal sonora têm mais probabilidade de não fazerem concordância que contexto com vogal surda. (LOPES, 2001, p. 368-369)

Das conclusões expostas por Lopes (2001), ressaltamos, sobretudo, a de que a posição à esquerda do núcleo, adjacente a ele, favorece mais a presença de marca do que a posição à esquerda, não adjacente ao núcleo. Por tanto seus dados dão respaldo suficiente para afirmar

¹³ A autora afirma que em seus dados mostrou exercer influência as seguintes variáveis: marcas precedentes e posição; saliência fônica: dimensões processos e tonicidade; relação dos elementos não nucleares em função do núcleo e posição dos elementos nucleares no SN; formalidade dos substantivos e adjetivos; grau dos substantivos e adjetivos; animacidade dos substantivos; contexto fonético/fonológico seguinte e função resumitiva. Na perspectiva não-atomística mostram exercer influência: pluralidade do contexto; configuração sintagmática do SN; saliência fônica; grau de formalidade do SN; pluralidade do SN; função textual e localização do SN.

¹⁴ A análise foi pautada em inquéritos com indivíduos usuários do que se convencionou chamar de Norma Urbana Culta – NURC –, ou seja, com sujeitos portadores de diploma de curso superior.

que a posição à esquerda adjacente ao núcleo dos elementos não nucleares é a mais favorecedora da concordância.

O segundo destaque será dado à dissertação de Andrade (2003), intitulada *Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialeto afro-brasileiro*, também considerada uma obra de referência por tratar do uso variável da concordância nominal de número, no dialeto rural afro-brasileiro da comunidade de Helvécia-Ba¹⁵, buscando explicações da variação na sócio-história do português do Brasil. Como Scherre (1988), Andrade (2003) analisa os dados nas duas perspectivas, atomística e sintagmática, chegando a conclusão que os fatores que mais exercem influência na concordância de número no SN são:

- (i) *Posição do constituinte com referência ao núcleo* - o efeito mais significativo na determinação do uso do morfema de plural, no dialeto de Helvécia, é exercido pelo fator *primeira posição*, pois é nesta que o elemento recebe mais frequentemente a marca de plural;
- (ii) *Saliência fônica* - outro importante condicionador da marcação de número plural. Embora, a autora afirme que não exista no dialeto de Helvécia uma escala muito bem definida, pois foi observado que os elementos tidos como mais perceptíveis fonologicamente são mais propensos a levar a marca do plural do que os menos salientes do ponto de vista morfológico, do tipo *casa/casas*;
- (iii) *Marcas precedentes ao elemento nominal analisado* - tem como principal fator determinante da marcação de número plural a presença de um numeral imediatamente anterior ao constituinte focalizado.

Esses três trabalhos servirão como suporte para o desenvolvimento da análise que nos propusemos realizar e desde já podemos inferir que as variáveis linguísticas e extralinguísticas utilizadas em uma pesquisa sociolinguística são, de forma geral, semelhantes dentro da abordagem de análise (atomística e sintagmática), porém os dados mediante as amostras analisadas nem sempre apresentarão os mesmos resultados.

¹⁵ Em janeiro de 1994, diante do fato de não haver encontrado características mais radicais nas comunidades visitadas até então, Alan Baxter e Dante Lucchesi organizaram uma recolha na comunidade de Helvécia, no Extremo Sul do Estado da Bahia, contando com a valiosa colaboração do antropólogo Marcos Luciano Lopes Messeder, que já havia contribuído com a recolha nas comunidades de Rio de Contas, e com a participação da estudante de Letras da UFBA Nara Barreto. A comunidade de Helvécia tinha um significado especial, porque nela a dialetóloga Carlota Ferreira havia identificado, no início da década de 1960, o que chamou de “remanescentes de um falar crioulo”. (LUCCHESI; BAXTER; RIBEIRO, 2009, p. 16).

As pesquisas supracitadas trabalham com *corpora* orais, porém cabe uma breve ressalva para o recente trabalho intitulado *Concordância nominal: cenas da variação em palcos do século XIX*, publicado por Oliveira, Soledade e Santos (2009), que tratam do fenômeno em uma perspectiva sincrônica, como expõem os autores, de eras idas, passadas. Para tanto utilizam-se de atas escritas por africanos¹⁶, ao longo do século XIX, analisando o fenômeno através do aparato, na medida do possível, metodológico da Sociolinguística Quantitativa. No artigo chegam a seguinte conclusão:

os textos dos africanos, escritos ao longo de duas décadas do século XIX, se mostraram uma preciosa fonte a servir de base empírica para mostrar que, nos oitocentos, a variação no âmbito da concordância nominal já acontecia e, pelo que tudo indica, de forma sistemática, já que, levando para o passado alguns fatores que regem a concordância no presente, houve consonância com o que se observa nos estudos da concordância nominal em tempos que correm. (OLIVEIRA; SOLEIDADE; SANTOS, 2009, p. 314)

Os autores (2009, p. 277) afirmam em seu artigo não ser este o primeiro trabalho que utiliza como *corpus* a linguagem escrita, já que em 2008, Maria Tereza Borges da Costa, escreve um texto em que procura esclarecer o fenômeno na escrita de alunos do ensino fundamental e médio residentes no município de Ribeira do Pombal – Ba¹⁷.

Os trabalhos de Oliveira, Soledade e Santos (2009) e o de Costa (2008 *apud* OLIVEIRA; SOLEIDADE; SANTOS, 2009, p. 279) têm demonstrado que a língua nas modalidades oral e escrita, mesmo sendo a segunda mais monitorada, apresentam resultados semelhantes no que se refere aos fatores que favorecem/condicionam a variação na concordância de número no sintagma nominal.

Os estudos apontados na *Figura 4* e outros – não apontados neste trabalho, porém de fundamental importância no campo dos estudos sociolinguístico já realizados sobre concordância nominal - nos permitem delinear os fatores estruturais que condicionam a marca canônica de plural nos sintagmas nominais no português falado no Brasil.

3.2.2 Concordância nominal de número variável no português de Angola

Como visto no subitem anterior, no âmbito do português do Brasil muitos estudos já foram realizados sobre a concordância nominal, muitos no âmbito da língua oral e ainda

¹⁶ Documentos escritos por africanos na Salvador oitocentista.

¹⁷ O *corpus* constituiu-se de 40 produções escritas da rede particular e 20 por alunos da 3ª série do ensino médio da rede pública.

poucos na modalidade escrita, porém este cenário não é tão vantajoso ao se tratar de outras variedades do português europeu, em especial do português falado em Angola, o qual trataremos em nossas análises.

Diante dos poucos estudos realizados sobre o português falado em Angola, diferentemente do PB, traçaremos um breve panorama dos estudos realizado sobre a variação na concordância nominal desenvolvido no continente africano:

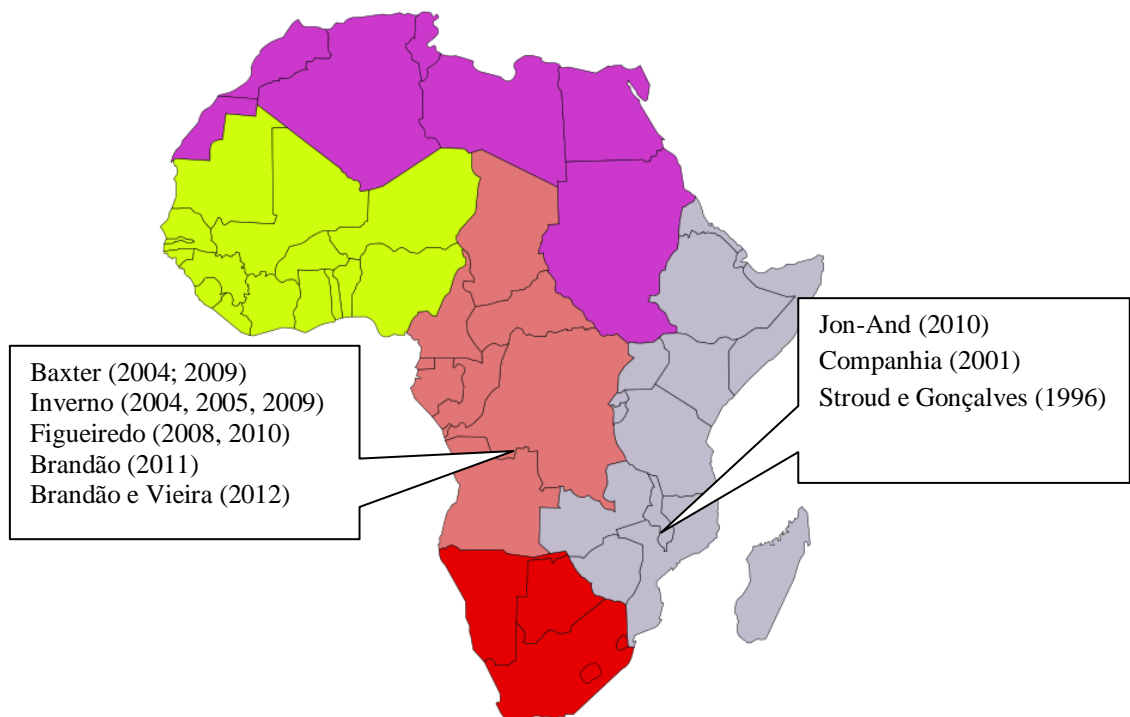


Figura 5: Panorama dos estudos sobre variação na concordância nominal desenvolvido na África

Como podemos observar na *Figura 5*, poucos e recentes são os estudos sobre o fenômeno concordância nominal no português africano, fato que incentiva a realização dessa pesquisa. De forma ainda preliminar, destacam-se os estudos quantitativos desenvolvidos na variedade de São Tomé (BAXTER, 2004; FIGUEIREDO, 2008), na variedade de Cabo Verde (JON-AND, 2009) e os estudos qualitativos nas variedades de Moçambique (COMPANHIA, 2001; STROUD e GONÇALVES 1997) e de Angola (INVERNO 2004, 2005, 2009; MARQUES, 1985). Mas uma vez ressaltamos que os trabalhos registrados na *Figura 5*, não dão conta de todos os trabalhos já desenvolvidos na África sobre variação na concordância

nominal, mas representam uma evidência de que esse é um fenômeno que merece destaque ao se tratar do português africano/angolano.

Dentre os trabalhos mencionados, destacaremos alguns que serão base bibliográfica para análise da nossa pesquisa. Primeiro o de Inverno (2004), intitulado *Português vernáculo do Brasil e português vernáculo de Angola: reestruturação parcial vs. mudança linguística*, que, como afirma a autora, constituiu um estudo piloto para uma comparação mais detalhada dos fatores sociais e linguísticos na origem das duas variedades vernáculas do português, a do Brasil e a de Angola, com o objetivo de procurar entender até que ponto os processos sociais e linguísticos que levaram à formação de cada um dos vernáculos terão sido semelhantes ou diferentes.

Inverno (2004, p. 08), em suas análises qualitativas, chega a conclusão que os fenômenos de concordância variável analisados, embora presentes nas duas variedades devido à influência do substrato banto, apresentam, contudo, uma menor estabilidade/cristalização no português vernacular angolano, parecendo, assim, apontar para uma formação mais recente desta variedade.

Temos também o importante trabalho de Inverno (2009)¹⁸, intitulado *A transição de Angola para o português vernáculo: estudo morfossintático do sintagma nominal*, que apresenta uma análise dos fatores sociolinguísticos e históricos relevantes para a compreensão da gênese e desenvolvimento do português vernacular de Angola, seguida de uma análise morfossintática do sintagma nominal. A autora formula a hipótese segundo a qual

[...] apenas na segunda metade do século XX um segmento suficientemente significativo da população angolana viria a ter acesso frequente à língua portuguesa e motivação suficiente para aprendê-la, iniciando assim um processo de reestruturação do português, ainda em curso, especialmente no interior entre falantes bilíngues do português e línguas banto, via mudança de língua. Esta hipótese é corroborada não só pelo evoluir da situação sociolinguística em Angola nos últimos cinco séculos, mas também pelas características morfo-sintáticas do SN no PVA, as quais denotam um índice de introdução de características de substrato típico de um processo de mudança de língua [...]. (INVERNO, 2009, p. 11)

Tratando de outras variedades do português, que não seja a angolana, destacamos dois trabalhos de Brandão (2011), que focaliza a concordância de número no sintagma nominal nas variedades urbanas do Português falado no Brasil e em São Tomé, com o objetivo de determinar os fatores que concorrem para a coexistência dos padrões variantes nelas

¹⁸ Os dados linguísticos analisados provêm de um *corpus* oral semi-espontâneo recolhido em 2004, no Dundo, província da Luanda Norte. O *corpus* foi recolhido através de trinta entrevistas informais, falantes que têm o português como segunda língua. Desses informantes, dezesseis apresentam idades compreendidas entre os 5 e os 17 anos e um grau médio de escolarização, 11 falantes apresentam idades compreendidas entre os 20 e os 56 anos e têm pouca ou nenhuma escolarização.

encontrados. Os resultados, advindos de análise fundamentada nos pressupostos da Sociolinguística Variacionista, demonstram que há significativas convergências entre as duas variedades, entre elas as fortes restrições de ordem estrutural e social que norteiam o cancelamento do morfema de número. Damos destaque também ao trabalho de Jon-And (2010), em que são analisadas, de forma quantitativa, algumas variáveis sociais e linguísticas que regem a concordância variável de número no sintagma nominal no português L2 falado em Maputo, Moçambique.

Portanto, de forma geral, esses são os estudos já realizados sobre concordância nominal, em especial, no português angolano. A escassez de estudos nessa perspectiva, e nessas variedades do português, deve ser vista como incentivo a novas pesquisas, uma vez que os resultados certamente contribuirão, e muito, para delinear as convergências e divergências entre as diferentes variedades do português.

3.3 A VARIAÇÃO NA CONCORDÂNCIA NOMINAL E A POLÊMICA SOBRE A FORMAÇÃO DO PORTUGUÊS DO BRASIL

O fenômeno da variação na concordância nominal tem sido o foco de muitas discussões no campo linguístico, mais especificamente no campo da sociolinguística, com os importantes trabalhos de Scherre e Naro, que hoje são considerados referências ao se tratar do assunto. As pesquisas atuais têm focado suas análises na busca de motivações linguísticas e extralinguísticas que colaborem para sistematização de tal variação e no favorecimento de evidências em torno da polêmica da formação do português brasileiro.

Ao se tratar da formação do PB, já vimos Capítulo 2, que duas correntes dividem os centros das discussões: a linha contatista e a linha derivista, as quais se encontram no âmbito das discussões sobre a origem do PB, desde o século XIX.

A primeira linha, derivista, acredita que as variações do português brasileiro sejam reflexos do português europeu, e que estas já estariam previstas no sistema linguístico. Scherre e Naro (1998), mesmo com seus poucos dados, o que torna suas conclusões contestáveis, afirmam ter encontrado no português europeu formas variáveis na concordância verbal e nominal, o que faz com que defendam que a variação não é uma marca peculiar ao português do Brasil, mas que além de ocorrer em Portugal, sempre ocorreu em toda a história da língua portuguesa. Estudos nessa vertente colaboram para a hipótese de um processo de

deriva natural da língua e descarta a hipótese de ocorrência de processos de crioulização na história da língua portuguesa falada no Brasil.

Lucchesi (1998, 2001, 2008, 2009), Baxter e Lucchesi (1997), acreditam que no Brasil não se formou um crioulo em termos gerais, mas um crioulo leve, mas que a língua portuguesa que hoje se fala aqui é herança de um massivo contato linguístico, que veio a gerar uma transmissão linguística irregular (*cf.* subitem 2.2.2).

No que tange à hipótese do contato linguístico, entende-se que as variedades orais do português popular do baiano oferecem campos de trabalho bastante promissores para o conhecimento da configuração sócio-histórica do PB, na grande área do semiárido baiano, dado o fato de esse ter sido gestado sob o efeito de amplo contato linguístico, durante o período Colonial e ter sido alvo também de um contato mais recente com variedades padrão e não-padrão do PB, propiciado por migrações, a partir da década de 40 do século XX.

Considerando o estudo sob uma perspectiva sócio-histórica, seria de supor que as características atuais da fala popular brasileira resultam de mudanças produzidas no contexto de contato do português com línguas de origem indígena e africana, hipótese assumida por diversos pesquisadores (LUCCHESI, 2009; MATTOS E SILVA, 2001; entre outros). Há ainda que se considerar outros fatos, como por exemplo, as mudanças ocorridas no português lusitano, alvo de mudança durante o período em que o PB foi gestado, o que torna a situação mais complexa, abrindo possibilidades comparativas com variedades atuais do português europeu (PE) contemporâneo.

No que concerne às variedades africanas do português, a questão da variação na concordância nominal de número é geralmente explicada através do contato linguístico, com possíveis influências das línguas bantu (BAXTER 2004, 2009; FIGUEIREDO, 2008; MARQUES, 1985; INVERNO, 2005), mas Jon-And (2010, p. 45) afirma que “por enquanto pode-se constatar que não podemos confiar unicamente na estrutura de línguas bantu como explicação para o favorecimento da marcação de plural da posição adjacente ao núcleo em variedades do português”.

Assim, defenderemos que as variedades orais do português do interior podem ser descendentes do chamado *português geral*, que, segundo Mattos e Silva (2001, p. 298-299) teria sido adquirido na oralidade e em situações de aquisição imperfeita, e que, na atualidade, há um processo de convergência da língua popular e da língua padrão/culta, conforme assumido por Lucchesi (2009) para o PB, motivado por contato em grandes centros urbanos e por vetores de standardização linguística, entre os quais a escolarização pode ser a causa principal (CARNEIRO; ALMEIDA, 2006, 2007).

É fato, como vimos no Capítulo I, que a formação histórica e social do Brasil é extremamente marcada pela integração de vários povos: os autóctones, africanos de várias partes, europeus, asiáticos. Agora resta-nos buscar indícios de que até que ponto esse cenário multilinguístico interferiu, ou melhor, favoreceu a formação do português falado no Brasil e de forma mais específica, para atingir o objetivo dessa dissertação, na variação da concordância nominal de número.

4 PRESSUPOSTOS TEÓRICOS METODOLÓGICOS

“Toda a teoria deve ser feita para poder ser posta em prática e toda a prática deve obedecer a uma teoria. Só os espíritos superficiais desligam a teoria da prática, não olhando a que a teoria não é senão uma teoria da prática, e a prática não é senão a prática de uma teoria. Na vida superior a teoria e a prática completam-se. Foram feitas uma para a outra.”

Fernando Pessoa

Para posterior análise da concordância gramatical de número plural entre os elementos flexionáveis do sintagma nominal no português do Brasil e no de Angola, adota-se uma perspectiva teórica que une aspectos da Teoria da Variação Linguística Laboviana ou Sociolinguística Quantitativa e aspectos da Teoria Funcionalista. Este capítulo também tece considerações acerca da Teoria dos 4M, de Myers-Scotton e Jake (2000 *apud* Lopes, 2001), que se refere a uma tipologia de morfemas, que neste trabalho é considerada como de grande interesse para o entendimento do fenômeno da concordância no sintagma nominal no português. Também apresentamos, neste capítulo, a abordagem metodológica utilizada no tratamento e análise dos dados. Para tanto delinearemos a área em que a pesquisa se insere e modelo quantitativo que a sustenta.

Mesmo que já tenhamos alguns manuais de “introdução aos estudos sociolinguísticos”, a citar o de Alkmin (2001), Mollica e Braga (2008), e muitas outras monografias, dissertações e teses que tratam dos pressupostos teórico-metodológicos da Sociolinguística, realizando descrições e análises da realidade linguística brasileira, ainda acredito na relevância de trazer alguns conceitos fundamentais, oferecendo, ainda que sumariamente, um pequeno arcabouço teórico-metodológico da Teoria da Variação, sem a qual não é possível compreender e trabalhar com os seus instrumentos de análises.

Para iniciar as discussões/devagações sobre a Teoria da Variação é necessário compreender, antes de tudo, que as línguas são organismos vivos multisseculares e que ao longo de seu tempo histórico, elas sofreram/sofrem algumas alterações, que nós, sociolinguistas, denominamos de variações e mudanças. Há mudanças fônicas, mórficas, sintáticas e léxico-semânticas, dentre essas mudanças, que constituem objetos de estudos da ciência da linguagem, a mórfissintática é a que nos interessa neste trabalho.

4.1 CAMINHOS DA SOCIOLINGUÍSTICA: TEORIA E MÉTODO

Ainda hoje podemos considerar que a Linguística é uma ciência relativamente recente, pois inicia-se, em bases teóricas, a partir das reflexões de Ferdinand de Saussure (2002 [1916]). A Linguística, ao se constituir como ciência autônoma, tem como objeto de estudo a *langue*, estrutura homogênea. Sua preocupação básica passou a ser a análise das relações internas entre os elementos linguísticos, estabelecendo, assim, a chamada “linguística estrutural”, sem se preocupar com as relações entre a linguagem e a sociedade. Hoje já se discute ser esta uma opção de caráter metodológico, pois Saussure reconhece a língua como um fato social.

Nesse sentido, Calvet (2002, p. 13-14), no seu livro intitulado *Sociolinguística: uma introdução crítica*, apresenta criticamente os progressos da Linguística. Inicia afirmando que foi o linguísta francês Antoine Meillet (1866-1936), contemporâneo de Saussure, quem insistiu em numerosos textos, no caráter social da língua, definindo-a como um fato social. Porém, com a publicação póstuma do *Curso de linguística geral*, Meillet afirma, em resenha que faz do livro, que ao separar a variação linguística das condições externas de que ela depende, Ferdinand de Saussure a priva da realidade; ele a reduz a uma abstração que é necessariamente inexplicável. Calvet (2002, p. 15) afirma que

A afirmação do caráter social da língua que se verifica em toda a obra de Meillet implica ao mesmo tempo a convergência de uma abordagem *interna* e de uma abordagem *externa* dos fatos da língua e de uma abordagem *sincrônica* e *diacrônica* desses mesmos fatos. Quando Saussure opõe linguística interna e linguística externa, Meillet as associa; quando Saussure distingue abordagem sincrônica de abordagem diacrônica, Meillet busca explicar a estrutura pela história [...]. Enquanto Saussure busca elaborar um modelo abstrato da língua, Meillet se vê em conflito entre o *fato social* e o *sistema que tudo contém* [...].

O autor observa ainda que próprio Labov (2008, p. 217) retrata o fato de que

Meillet, contemporâneo de Saussure, acreditava que o século XX assistiria o desenvolvimento da explicação histórica baseada no exame da mudança linguística encaixada na mudança social (1921). Mas discípulos de Saussure, como Martinet (1964), repudiaram ativamente essa opinião e empreenderam esforços para que a explanação linguística ficasse confinada às inter-relações de fatores internos, estruturais. Assim procedendo, eles certamente seguiam o espírito da doutrina de Saussure, pois o estudo mais atento de seus escritos sugere que, para ele, ‘social’ não significava muito mais do que ‘multi-individual’, sem nenhuma sugestão das implicações mais amplas da interação social.

Nessa perspectiva, passa-se a entender a linguística como uma ciência social, que busca dar conta dos processos de variação linguística e mudança social, ponto em que William

Labov se debruçará ao constituir o modelo teórico-metodológico da Sociolinguística Quantitativa.

Calvet (2002, p. 32) assinala que Meillet já afirmava que é preciso buscar a explicação da irregularidade das variações linguísticas nas flutuações da composição social da comunidade linguística e isso é retomado por Labov ao publicar, em 1966, seu estudo sobre a estratificação social do /r/ nas grandes lojas de Nova York. Porém, Calvet (2002, p. 33) ressalva que

Meillet, comparatista de alto nível, trabalhou sobretudo com línguas mortas, enquanto Labov trabalha continuamente com situações contemporâneas concretas, enfrenta problemas de metodologia da pesquisa, em suma, constrói um instrumento de descrição que tenta ultrapassar, integrando-os, os métodos heurísticos da linguística estrutural.

Portanto, é a partir das pesquisas desenvolvidas pelo americano William Labov que nascerá a corrente linguística denominada linguística variacionista, a qual propõe uma reação à ausência do componente social no modelo gerativo e que veementemente insiste na relação entre língua e sociedade e na possibilidade de se sistematizar a variação existente na própria língua falada.

Dentre os trabalhos publicados pelo autor, destaca-se o publicado em 1963¹⁹, um dos célebres trabalhos de Labov, em que trata sobre o inglês falado na ilha de Martha's Vineyard, no estado de Massachussets - Estados Unidos, destacando o papel decisivo dos fatores sociais na explicação da variação linguística, isto é, da diversidade linguística observada. Em 1964, Labov realiza um estudo sobre a estratificação social do inglês em Nova York, a partir do qual fixa um modelo de descrição e interpretação do fenômeno linguístico no contexto social de comunidades urbanas – conhecido como Sociolinguística Variacionista ou Teoria da Variação, de grande impacto na linguística contemporânea.

Com base em estudos empíricos, os linguistas Uriel Weinreich, William Labov e Marvin Herzog lançaram os fundamentos da teoria da mudança linguística com o texto *Empirical Foundations for a Theory of Language Change* (1968), no qual demonstram como se desenvolver estudos sobre a gramática de uma comunidade de fala, buscando explicar o funcionamento dessa gramática a partir de seus aspectos internos e externos. Através deste livro, de estudos de bases empíricas, foi possível compreender melhor a heterogeneidade

¹⁹ Nessa pesquisa, Labov relacionou fatores como idade, sexo, ocupação, origem étnica e atitude em relação ao comportamento linguístico dos nativos da ilha, no que se referia à pronúncia de determinados fones do inglês (as vogais dos ditongos [ay] e [au]), ao qual se seguiram estudos sobre a estratificação social do inglês falado na cidade de Nova York.

linguística do português do Brasil, instaurando uma nova perspectiva de linguagem, uma vez que ao enfatizar o lugar central da variação e da mudança nas línguas naturais e explicitar os diversos aspectos envolvidos nesses processos, os autores abrem espaço, como afirmam as autoras Duarte e Paiva no posfácio do livro, para o desenvolvimento de uma modelização orientada por questões precisas.

Na apresentação da versão brasileira desse livro, Carlos Alberto Faraco comenta que o texto do livro foi apresentado no Simpósio “Direções para a Linguística Histórica”, na Universidade do Texas, e publicado 1966. Momento em que se buscava renovar as pesquisas no âmbito dos estudos linguísticos históricos, no intuito de torná-la mais visível aos olhos da linguística. Faraco ainda afirma que os dados dialectológicos coletados na década de 1950 e o surgimento da Sociolinguística no início da década de 1960 já apontavam para uma renovação teórica e metodológica da linguística histórica.

Nesse contexto, Weinreich, Labov, Herzog, em sua introdução afirmam que *Empirical Foundations for a Theory of Language Change* (1968)

[...] se baseia na observação de que as teorias estruturais da língua, tão frutíferas na investigação sincrônica, sobrecarregaram a linguística histórica com um fardo de paradoxos que não foram totalmente superados. Ferdinand Saussure, ao assentar as fundações do estudo sincrônico, estava ciente da correspondente intratabilidade da mudança linguística, e claramente resignado com isso. Mas entre a maioria dos linguistas depois de Saussure, a escolha entre estudar ou a estrutura ou a história das línguas não ficou bem assentada. (WEINREICH; LABOV; HERZOG, 1968, traduzido por BAGNO, 2008, p. 33)

No que tange a aproximação entre a Sociolinguística e a Linguística Histórica, Oliveira (2006) comenta a existência de uma relação estreita entre essas áreas, uma vez que, estudando a variação da língua dentro da estrutura social da comunidade em que ocorre, identifica e prevê processos de mudança em curso, que, obviamente, são decorrentes dessas variações. A autora ainda afirma que “sendo a língua um meio de interação entre o indivíduo e a sociedade em que ele atua, está claro que há forças sociais que agem sobre essa interação.” (OLIVEIRA, 2006, p. 45).

Dessa forma, os estudos que hoje buscam o relacionamento casual entre língua e sociedade, são frutos da proposta de análise contida em *Empirical Foundations for a Theory of Language Change* (1968), em que se propõem estudos sobre a gramática de uma comunidade de fala, buscando explicar o funcionamento dessa gramática a partir de seus aspectos internos e externos. Hoje esse texto tornou-se um importante material, uma vez que

Nele – depois de uma revisão crítica do pensamento neogramático e do estruturalismo saussuriano, em que apontam os paradoxos introduzidos pelos

imanentistas na reflexão histórica com sua concepção homogeneizante de língua – aqueles autores, revendo trabalhos empíricos de dialectologia e sociolinguística com implicações para a compreensão da mudança, delineiam uma estratégia geral que se apóia justamente nesses fundamentos empíricos para estudar a mudança linguística. (FARACO, 2005, p. 189)

Faraco (2008) - na apresentação da versão brasileira dos *Fundamentos Empíricos* – afirma que os textos de Weinreich, Labov, Herzog apresentam complexibilidade, ao trazer desafios para os estudos da mudança linguística e ao atentar para o fato de que os fatores que desencadeiam e restringem as mudanças são certamente mais diversificados em comunidades linguísticas contemporâneas do que em sociedades pré-urbanas, que até então era foco de parte das investigações históricas tradicionais em linguística.

Dessa forma, a heterogeneidade linguística sai do plano hipotético dos estudos saussureanos e ganham espaços nas investigações sociolinguísticas, que passa a entender a língua dentro do seu contexto de variabilidade e mudança. Nesse novo contexto teórico, cinco são os problemas apontados por Weinreich, Labov, Herzog (1968, pp. 121-125) quando se trata de qualquer teoria da mudança linguística. São eles:

- (i) **O problema dos fatores condicionantes (*the constraint problem*)** - A teoria precisa determinar o conjunto de mudanças possíveis e as condições possíveis para a mudança por meio de um estudo minucioso de mudanças em progresso;
- (ii) **O problema da transição (*the transition problem*)** - As mudanças têm uma distribuição contínua através de sucessivas faixas etárias da população;
- (iii) **O problema do encaixamento (*the embedding problem*)** - As mudanças linguísticas devem ser vistas como encaixadas no sistema linguístico como um todo. Há dois tipos de encaixamento:
 - *Encaixamento na estrutura linguística* – Apresenta um modelo de língua em que se tem: a) estratos discretos, coexistentes, definidos pela co-ocorrência estrita, que são funcionalmente diferenciados e conjuntamente disponíveis a uma comunidade de fala; b) variáveis intrínsecas, definidas por co-variação com elementos linguísticos e extralinguísticos, o que permitiu afirmar que um conjunto limitado de variáveis num sistema altera seus valores modais gradualmente de um pólo a outro. Portanto, as variantes das variáveis podem ser contínuas ou discretas;

- *Encaixamento na estrutura social* - A estrutura linguística está encaixada no contexto mais amplo da comunidade de fala. Portanto, variações sociais e geográficas são elementos intrínsecos da estrutura, uma vez que os fatores sociais pesam sobre o sistema como um todo;
- (iv) **O problema da avaliação (*the evaluation problem*)** - O linguista deve estabelecer empiricamente os correlatos subjetivos da avaliação dos diversos estratos e variáveis numa estrutura heterogênea. Além disso, o nível de consciência social é uma propriedade importante da mudança linguística que tem de ser determinada diretamente;
- (v) **O problema da implementação (*the actuation problem*)** - O processo global da mudança linguística pode envolver estímulos e restrições tanto da sociedade quanto da estrutura da língua. O problema em torno da implementação advém do grande número de fatores que influenciam a mudança;

Diante do legado deixado por Weinreich, Labov, Herzog (1968), podemos entender que a sociolinguística é um espaço de investigação interdisciplinar que estuda e correlaciona aspectos dos sistemas linguísticos e dos sistemas sociais, através dos empregos concretos da língua, dentro das comunidades de fala.

Nesse sentido, Oliveira (2006) afirma que a concepção e o alcance do modelo sociolinguístico são, ao mesmo tempo, sincrônicos e diacrônicos, já que toda mudança implica uma variação. Nesse novo modelo, que surge na década de 1960, a estrutura de uma língua somente será totalmente entendida à medida que se compreendam efetivamente os processos históricos de sua configuração.

4.1.1 A Sociolinguística quantitativa

A autêntica pesquisa sociolinguística circunscreve-se na relação entre língua e sociedade e na possibilidade de investigar e descrever sistematicamente a variação existente em contextos específicos de fala, no uso real da língua. Essa base teórica deve-se ao americano William Labov, a quem se deve as primeiras investigações com base nos postulados da sociolinguística, que também é conhecida como “sociolinguística quantitativa”,

pois em se tratando de variação, torna-se imprescindível se operar com números e tratamento estatístico dos dados coletados.

No que tange a análise quantitativa, Guy e Zilles (2007) dedicam-se intensamente à tarefa de oferecer ao leitor subsídios para uma análise quantitativa adequada aos princípios teórico-metodológicos da sociolinguística quantitativa variacionista.

Para o desenvolvimento de uma análise sociolinguística, em que seja possível abarcar fatores linguísticos e extralinguístico, é necessário um grande número de dados, para que seus resultados sejam representativos e possam, realmente, indicar os caminhos da variação e/ou mudança de certas variantes linguísticas. Para tanto, se faz necessário todo um cuidado metodológico já abordado por Tarallo (1997), Guy e Zilles (2007), Mollica e Braga (2008) e outros pesquisadores, que vai desde a coleta de dados a interpretação dos resultados. Como já afirma Guy (2007):

Para desvelar tanto a estrutura linguística quanto a estrutura social, devemos, necessariamente, coletar grande quantidade de dados de muitos indivíduos; conseqüentemente, devemos enfrentar problemas ligados a controle de qualidade e confiabilidade, a manuseio e apresentação de dados, e a interpretação e inferência. (GUY, 2007, p. 19)

O autor segue comentando que parece sensato afirmar que toda pesquisa dialetal, seja ela de ordem geográfica ou social é essencialmente quantitativa, uma vez que estudos dialetais anteriores podem comprovar que dificilmente se fala em termos categóricos sobre as propriedades dos dialetos. O autor ainda explica que isso acontece diante da heterogeneidade da língua falada, já que as isoglossas não formam grupos perfeitos, o que implica dizer que todo dialeto está propício a variação linguística (GUY, 2007, p. 20). Portanto, cabe a sociolinguística fazer a abordagem linguística e social dos mais diversos dialetos, tomando como suporte metodológico a análise quantitativa dos dados, podendo assim controlar o grau de variabilidade do fenômeno em análise, e apresentá-los através dos mais diferentes recursos: quadros, tabelas, gráficos.

Numa análise quantitativa, o pesquisador busca sempre os números, as probabilidades, as estatísticas, para que seja possível apreender a sistematicidade da variação linguística, em análise, uma vez que esses resultados refletirão o uso da língua em um contexto social e linguisticamente heterogêneo. Dessa forma, entendemos que na abordagem da sociolinguística quantitativa

os padrões linguísticos são descritos, e explicados (na medida em que objetos dessa natureza podem ser explicados), em termos de uma gramática de regras variáveis que operam com probabilidades associadas a fatores sociais e restrições hierarquizadas da estrutura linguística. (LUCCHESI, 2004, p. 196)

De qualquer modo, Guy (2007) define três fases numa análise quantitativa. São elas:

- (i) **COLETA DE DADOS** – momento em que se organiza uma amostra de dados, a qual deve apresentar altos índices de confiabilidade e representatividade, para que os dados não sejam tendenciosos;
- (ii) **REDUÇÃO E APRESENTAÇÃO DE DADOS** – abordagem que busca formas que minimize detalhes sem importância e apresente o que realmente é relevante à análise, sem que isso torne os dados tendenciosos ou distorcidos, ou até mesmo para não incorrer no risco de obscurecer/omitir dados/fatos importantes;
- (iii) **INTERPRETAÇÃO E EXPLICAÇÃO DE DADOS** – representa o objetivo final de qualquer estudo quantitativo dialetal, uma vez que os números/estatísticas apenas direcionam as análises e permitem interpretações e explicações dos fenômenos linguísticos.

Todo esse trabalho, da coleta a interpretação, ganhou um grande auxílio das ciências exatas e computacionais, uma vez que, a fim de facilitar a trabalho de quantificação de dados, estudiosos elaboraram programas computacionais que oferecerem um tratamento estatístico aos dados sociolinguísticos. Portanto, transforma todo o amontoado de ocorrências em probabilidades estatísticas, o que veio a facilitar a vida dos pesquisadores sociolinguísticos.

No que tange aos programas computacionais, o mais usados é o Variable Rule Analysis²⁰ (VARBRUL), que é um tipo de análise multivariada amplamente empregada em estudos de variação linguística. Este programa busca separar, quantificar e testar a significância dos efeitos dos fatores (sociais ou linguísticos) em uma variável linguística (GUY, 2007, p. 33).

Desde a criação do VARBRUL, muitas foram as alterações realizadas no programa para que esse se torne, a cada dia, mais prático, confiável e com mais ferramentas operacionais, o que tem colaborado na aplicação de testes, cruzamentos de dados. Todos esses recursos oferecidos pelo mundo moderno têm vindo a somar na realização pesquisas mais consistentes.

Em meios aos avanços e modernização da sociolinguística quantitativa, podemos afirmar que

²⁰ Análises de regras variáveis.

antes do advento da metodologia de quantificação, a variação linguística era considerada secundária, aleatória ou mesmo impossível de ser cientificamente apreendida. O uso de métodos estatísticos, contudo, 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)

Assim entende-se que a sociolinguística tem em muito contribuído para a análise de uma língua real, não idealizada. A partir do seu aparato teórico-metodológico, é possível “[...] enfrentar o desafio de tentar processar, analisar e sistematizar o universo aparentemente caótico da língua falada” (TARALLO, 1997, p. 5), contribuindo para discussões de aspectos que por muito tempo foram excluídos dos estudos tradicionais da língua, uma vez que agora se busca uma abordagem que trata da diversidade linguística, implicando num “[...] esforço que vem sendo feito há séculos pelos linguistas em sua difícil tarefa de descrever um fenômeno que nos governa e ultrapassa em tão grande medida” (PERINI, 2003, p.10).

Tudo isso tem contribuído para a configuração de uma ciência mais social, mais presente, mais humana e através dela, certamente, muitos conceitos poderão ser revistos ou até mesmo repensados.

Diante do amplo campo de atuação da Sociolinguística, muitos estudos têm sido realizados no Brasil, dentre eles destacamos o de Bortoni-Ricardo (2011), em seu livro *Do campo para a cidade: estudos sociolinguístico de migração e redes sociais*, em que discute as características da comunidade nacional de fala, abordando a amplitude da variação do português brasileiro na perspectiva de um contínuo dialetal que se estende das variedades rurais estigmatizadas à variedade urbana padronizada. Certamente a proposta lançada pela autora harmoniza-se, ou ao menos subsidiará o enquadramento dos nossos *corpora*, a serem descritos no subitem que trata da metodologia.

A autora foca sua atenção na Brazlândia, cidade-satélite de Brasília, e produz, como afirma Ataliba de Castilho no prefácio do livro, um rigoroso retrato sociolinguístico do local, operando com o paradigma das redes sociais²¹. A partir deste trabalho é que a sociolinguística urbana abre novos horizontes, podendo agora conhecer e reconhecer a perspectiva rurbana, zona de interseção que problematiza a chegada de contingentes rurais ao meio urbano.

As novas perspectivas assumidas pela sociolinguística vêm colaborando para o desafio de Weinreich, Labov e Herzog (2006 [1975]) de buscar um modelo formal capaz de acomodar a heterogeneidade ordenada como uma realidade inerente às línguas e não como um fundamento marginal.

²¹ “Conjunto de vínculos de todos os tipos entre os indivíduos em um grupo.”

4.2 TEORIA FUNCIONALISTA

A hipótese funcionalista que se discute aqui se baseia na “condição de distintividade” apresentada por Kiparsky (1972, p. 195, *apud* SCHERRE, 1997, p. 182), que acredita que “há uma tendência para a informação semanticamente relevante ser mantida na estrutura superficial”. Numa releitura de Kiparsky (1972 *apud* SCHERRE, 1997, p. 182), podemos afirmar que em um contexto comunicativo, um segmento que seja necessário à transmissão eficiente da mensagem tenderá a não desaparecer.

No quadro da corrente funcionalista, Scherre (1988, p. 45) comenta:

- (i) A língua é considerada, antes de mais nada, como um instrumento de comunicação inserido no seio de uma comunidade. Como tal, não pode ser estudada no abstrato sem que se leve em conta o contexto linguístico e, inclusive, todo o processo comunicativo (NICHOLS, 1984, p.97-98 *apud* SCHERRE, 1988, p. 45);
- (ii) No que tange ao objeto do estudo linguístico interessa, a esta corrente funcionalista, descrever os fenômenos linguísticos substantivos e não constructos teóricos e analíticos, buscando, conseqüentemente, universais de natureza substantiva (NICHOLS; WOODBURY, 1985, p.2 *apud* SCHERRE, 1988, p. 45). Os dados provenientes de intuição não têm, portanto, lugar nesta perspectiva funcionalista.

Voltando o olhar para o fenômeno da variação na concordância de número no SN, entenderemos, através de Kiparsky (1972, p. 195 *apud* SCHERRE, 1997, p. 182), que há uma série de regularidades na fonologia e na morfologia que são baseadas em condições de natureza funcional e que para explicá-las é necessário “[...] adicionar à teoria condições substantivas que pertencem não à forma da gramática (isto é, ao sistema de regras que gera a língua) mas à sua saída (output)”. Portanto, um destes aspectos regulares diz respeito a regras opcionais com frequências variáveis, conforme Kiparsky (1972, p.197 *apud* SCHERRE, 1988, p. 42).

Assim podemos inferir que o fenômeno da concordância de número plural, no português culto do Brasil, é redundante por repetir a mesma marca (-s) que contém as mesmas informações em toda cadeia sintagmática. A visão tradicional (*cf.* subitem 3.1) prevê muitos desses contextos, mas pesquisas (*cf.* subitem 3.5.1) vêm demonstrando que a concordância nominal é um fenômeno de natureza variável, que alterna entre a preservação das marcas

redundantes e a perda dessas marcas em algum momento da cadeia sintagmática e essa alternância está relacionada a fatores linguísticos e extralinguísticos. Nesse sentido, Scherre (1997, p. 183) afirma que na literatura da linguística variacionista, a colocação de Kiparsky tem sido denominada de hipótese funcionalista, e uma série de estudos tem sido realizada sob essa óptica, ora buscando, ora apresentando evidências que a fortaleçam (POPLACK, 1980, 1981; Guy, 1981; PAREDES, 1993; CAMERON, 1993).

4.3 TEORIA DOS 4M

O que será exposto sobre a teoria dos 4M, ou teoria dos 4 tipos de morfemas, baseia-se na releitura de Lopes (2001), a qual acredita que a variação da concordância nominal de número entre os elementos do SN, no português do Brasil, tem a ver com o tipo de aquisição por que passaram negros e índios, que pode ser respaldada na teoria dos 4M (quatro tipos de morfemas), de Myers-Scotton e Jake (2000).

Antes de discutirmos as questões implícitas a teoria dos 4M, precisamos entender que essa pressupõe que a “[...] a ordem em que os morfemas são adquiridos numa língua está relacionada a uma tipologia que, assim como explica a aquisição desses elementos, pode servir para entender o processo da variação dos mesmos” (LOPES, 2001, p. 34).

Myers-Scotton e Jake (2000, p.1054 *apud* LOPES, 2001, p. 93) relacionam a ordem de aquisição da linguagem aos diversos tipos de morfemas, encontrados na estrutura da língua. Os autores apresentam quatro tipos de morfemas:

- 1) *Os content morphemes*, ou morfemas de conteúdo (os substantivos, os adjetivos, os verbos são exemplos de *content morphemes*); e os *system morphemes*, ou morfemas sistêmicos, podendo ser de três tipos:
 - 2) Os *early system morphemes* (em inglês, *out* e *at*, em *to look out* e *to look at*);
 - 3) Os *bridge system morphemes* (o morfema de plural *-s*, em *two tables*);
 - 4) Os *outsider system morphemes* (o *-s* indicador de 3ª pessoa, em *he works*).
- (MYERS-SCOTTON; JAKE, 2000, p. 1054 *apud* LOPES, 2001, p. 92-93)

Através dessa proposta apresentam uma hierarquia em que os diversos fatos da estrutura linguística são adquiridos tanto em aprendizagem de primeira como de segunda língua. Dessa forma, os autores acreditam que

a produção lingüística inicia-se no nível conceptual (pré-lingüístico), onde são ativados conjuntos de traços semântico-pragmáticos lingüisticamente específicos. Lemas, intermediários entre as intenções e a produção de estruturas sintáticas, são selecionados por esses conjuntos. Os lemas diretamente escolhidos no nível conceptual são a base para os *content morphemes*, ou morfemas lexicais. Além desse tipo de morfemas, os morfemas de conteúdo, definidos no nível dos lemas, outros são utilizados para concretizar as diversas intenções dos falantes, pois esses primeiros não são suficientes para fazê-lo: são os *system morphemes*, ou morfemas gramaticais. Esses, apesar de não serem morfemas de conteúdo, podem ser selecionados também no nível dos lemas ou, como a maior parte deles, no nível gramatical, ou no *formulator*. Esse nível é o responsável pela definição das estruturas gramaticais superficiais, pelas relações puramente sistêmicas ou estruturais. (MYERS-SCOTTON, JAKE, 2000, p. 1055 *apud* LOPES, 2001, p. 94)

Andrade (2003, p. 78) também discute sobre a teoria dos 4M, sob a releitura de Lopes (2001), e infere que os autores acreditam que os *early system morphemes* seriam ativados concomitantemente aos *content morphemes* para atender necessidades funcionais imediatas, e a atuação destes explicaria, por exemplo, a concentração de marcas de plural no primeiro elemento do SN, no português popular do Brasil, em detrimento das demais posições. Os *bridge late system morphemes* atuariam, no caso da concordância nominal de número, nos demais itens pluralizáveis do SN, “pois eles são pluralizáveis apenas por cumprir a orientação gramatical, a da concordância” (LOPES, 2001, p. 96). Sendo assim:

a análise da concordância no sintagma nominal do português que aqui se faz considera que esse fenômeno estrutural envolve, principalmente, morfemas do tipo *bridge late system*, além, também, de *early system morphemes*, conforme se discute neste trabalho, na análise dos dados. São, pois, tratados como *early system morphemes* os morfemas de plural nos nomes quando são os primeiros ou os únicos elementos pluralizáveis do sintagma ou aqueles em elementos anteriores imediatamente ao nome. Consideram-se *bridge system morphemes* todos os outros morfemas de plural do sintagma, pois eles são pluralizáveis apenas para cumprir orientação gramatical, a da concordância.

Por os autores chamarem à atenção para o fato de que os mesmos traços não são necessariamente codificados pelos mesmos tipos de morfemas nas diversas línguas, podemos considerar, pois, que a morfologia referente à concordância dentro do sintagma nominal, no português, estaria ora entre os *early system morphemes*, ora estaria se comportando como os *late system morphemes* (LOPES, 2001, p. 98). Vejamos:

(8) **AS FESTAS** daqui é muito animada.

A partir de Lopes (2001, p. 98), conseguimos entender que num sintagma nominal “As festas”, da sentença (8), o morfema de plural de “as” é inserido antecipadamente, portanto deve ser considerado um *early system*. Já o morfema de plural de “festas” parece não ser

gerado no mesmo momento, pode ser considerado um tipo de *late system*. Essa pode, segundo Lopes (2001) e Andrade (2003), ser a explicação para o morfema de plural *-s* de *festas* ser registrado com menos frequência em processos de variação, e ser fixado mais tarde em processos de aquisição.

Ao tratar da importância dos contextos sócio-históricos

em que ocorre o deslocamento forçado de populações, e o aprendizado lingüístico se dá através de um modelo “o mais ‘empobrecido’ possível de pistas da língua alvo” (p. 98), a teoria prevê que é seguido o mesmo roteiro postulado para a aquisição lingüística de primeira língua: os morfemas de conteúdo e os *early system morphemes* formam-se cedo para contemplar a intenção comunicativa do falante, suprimindo as perdas que se observam em línguas crioulas. (ANDRADE, 2003, p. 78)

Dessa forma, Lopes (2001, p. 102-103) e Andrade (2003, p. 79) acreditam que no contexto de estruturação de línguas crioulas os *early system morphemes* do superstrato podem ser absorvidos por essas línguas e os *late system morphemes* não passam para as línguas crioulas. Portanto,

[...] os *early system morphemes* são apenas ‘satélites’ dos morfemas de conteúdo, só são utilizados quando os de conteúdo a que eles se associam são acessados, e são ativados, por isso, na sua fase de produção, conjuntamente com eles. Nos crioulos, apesar de não ser regular a existência desse tipo de morfema, eles ocorrem em algumas dessas línguas. Myers-Scotton (no prelo) diz que eles podem surgir na forma de morfemas aglutinados ao nome, como um morfema pré-nominal, a exemplo das partículas *du, de la* francesas, presentes nos crioulos de Reunião, de Ile de France, do Haiti e no Mauriciano, de base francesa. (LOPES, 2001, p. 103)

A partir do que foi discutido até momento e do que foi exposto por Lopes (2001) é possível entender, na língua portuguesa, que no processo de concordância no SN, mediante a inserção de marca de plural em todos os elementos flexionáveis do enunciado, “os morfemas de plural não são todos ativados no mesmo nível e, por isso, tenham implicações nessa língua” (p. 107). Portanto, como Lopes (2001) já demonstrou em suas análises que a teoria dos 4M torna-se uma alternativa, aparentemente viável, para explicarmos o processo de variação de concordância nominal de número e sua conseqüente baixa frequência em dialetos populares do Brasil e de Luanda.

4.4 ABORDAGEM METODOLÓGICA

4.4.1 Amostragem

O material analisado foi extraído dos *corpora* dos projetos *A língua portuguesa no semiárido baiano* (fases 1, 2 e 3²²) e do projeto *Em busca das raízes do português brasileiro*, respectivamente, amostras do Português popular do Brasil (PPB) e do Português popular de Angola (PPA), respectivamente. Quanto ao fato das amostras serem voltadas ao português popular, justifica-se pela proposta, assumida neste trabalho, de que, embora as variedades cultas e populares do PB e do PA se oponham ao PE, o português popular tem uma origem no processo de transmissão linguística irregular, diferente do que ocorreu, a princípio, com a variante culta, o que pode apresentar fatos linguísticos de maior distanciamento do PE.

As amostras orais serão consideradas como amostras de língua-E (exteriorizadas), portanto o ponto a ser tomado aqui será a assunção de que a mudança se dá durante o processo de aquisição de línguas e que o ambiente de contato linguístico assume um papel importante para o nascimento de uma nova língua, como é o caso do PB e PA, marcado por uma história de contato, de maior ou menor influência, sobretudo em sua vertente popular. Serão consideradas, como ponto de partida, causas exógenas, como base das alterações no curso do PB e do PA, opondo-os às variedades faladas em Portugal.

4.4.1.1 Amostras da comunidade de fala de Mato Grosso

O Projeto *A Língua Portuguesa no Semi-árido Baiano*, coordenado pelas professoras Zenaide Carneiro e Norma Lucia de Almeida, do Departamento de Letras e Artes da Universidade Estadual de Feira de Santana, tem por objetivo contribuir para o conhecimento da realidade linguística brasileira e, de forma específica, sobre a língua falada em áreas rurais do semiárido baiano.

A título de curiosidade, o semiárido baiano compreende a grande parte do território brasileiro e ocupa uma parcela relevante da Região Nordeste. Através dos dados fornecidos pelo site da *Sudene* (Superintendência do Desenvolvimento do Nordeste), podemos afirmar que o semiárido nordestino, caracteriza-se principalmente pelas frequentes secas que podem ser

²² *Corpus* do projeto *A língua portuguesa no semiárido baiano, fase 3*, desenvolvido no Núcleo de Estudos de Língua Portuguesa (www.uefs.br/nelp) da Universidade Estadual de Feira de Santana, inicialmente implementado por Norma Lucia Fernandes de Almeida, Zenaide de Oliveira Novais Carneiro e equipe e na atual fase a cargo de Silvana Araújo, associado ao projeto *Vertentes*, coordenado por Dante Lucchesi (www.vertentes.ufba.br).

caracterizadas pela ausência, escassez, alta variabilidade espacial e temporal das chuvas. As características do meio ambiente condicionam a sociedade regional, a sobreviver principalmente de atividades econômicas ligadas basicamente à agricultura e a pecuária. Estas se realizam sempre buscando o melhor aproveitamento possível das condições naturais desfavoráveis, ainda que apoiadas em base técnica frágil, utilizando na maior parte dos casos, tecnologias tradicionais.

A *Sudene* ainda afirma que apesar da urbanização ocorrida nos últimos anos, a ocupação principal de sua força de trabalho é a agropecuária. A estrutura fundiária é extremamente concentrada, embora seja grande o número de pequenos estabelecimentos ou unidades de produção familiar. O semiárido brasileiro é composto por 10 (dez) estados²³, dos quais destacamos, neste trabalho, a Bahia, com 417 municípios, dos quais 265 integram o semiárido baiano, correspondendo a 63,55% dos municípios baianos (SUDENE, 2005). A partir do mapa abaixo podemos visualizar melhor a extensão territorial do semiárido baiano.

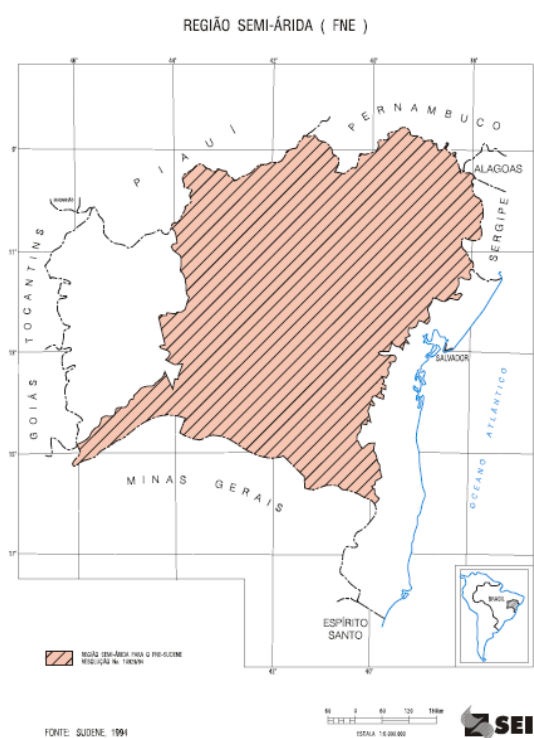


Figura 6: Mapa da região semi-árida da Bahia

²³ O Semiárido brasileiro é representado pelos estados do Piauí (56,95%), Ceará (81,52%), Rio Grande do Norte (88,02%), Paraíba (76,23%), Pernambuco (65,95%), Alagoas (37,25%), Sergipe (38,67%), Bahia (63,55%) e Minas Gerais (50,60%).

As amostras representativas do português popular do Brasil, escolhidas para desenvolvimento deste trabalho, se restringem a região semiárida do estado da Bahia. Tais amostras, organizadas em quatro volumes, foram coletadas no período de 1994 a 2002 e integraram a *Coleção Amostras da Língua Falada no Semiárido Baiano*, publicada em 2008, que visou difundir *corpora* linguísticos diferenciados do português vernacular do interior da Bahia, selecionando alguns municípios representativos do processo de constituição demográfica na região do semiárida. Vejamos quais cidades estão inclusas nos volumes da coleção:

VOLUME	CIDADES	AMOSTRAS
I	Anselino da Fonseca (Piemonte da Chapada)	Piabas
II	Rio de Contas (Chapada Diamantina)	Bananal / Barra dos negros Mato Grosso
III	Feira de Santana	Matinha
IV	Jeremoabo	Casinhas Tapera Lagoa do Inácio

Quadro 2: Descrição dos volumes da *Coleção Amostras da Língua Falada no Semiárido Baiano*

Almeida e Carneiro (2008) afirmam que a escolha das comunidades levou em conta dois importantes fatores do processo de (re)urbanização do interior do estado, a ‘agropecuária’ (séc. XVII a XIX) e a ‘mineração’ (séc. XVIII basicamente). Assim, objetivando constituir *corpora* linguísticos diferenciados do português vernacular do interior da Bahia foram selecionados alguns municípios que tinham em sua sócio-história o contato entre as pessoas envolvidas na ocupação e produção rural baiana, inicialmente, a partir de aldeamentos e da criação de vilas e, posteriormente, do contato entre escravos e senhores e, ainda, de negros alforriados, brancos pobres e mestiços (CARNEIRO; ALMEIDA, 2008). Portanto, comunidades que podem fornecer indícios sobre a forma de interação e contatos linguísticos distintos e sobre o avanço e consolidação da língua portuguesa na região.

Dentre as comunidades representativas a um estudo sociolinguístico, foi escolhida, para esta pesquisa, a comunidade de Mato Grosso, situada no município de Rio de Contas. Um dos

fatores que levou a escolha desta comunidade, em comparação a de Luanda, foi que embora com contexto sócio-histórico distintos, havia indícios de contato linguístico. Já que

É evidente que essa distinção ‘com contato versus sem contato’ não pode ser encarada em termos absolutos, nomeadamente com respeito à comunidade de Mato Grosso. Por mais que os habitantes dessa comunidade afirmem a sua ‘pureza étnica’ e proclamem uma descendência direta de emigrantes provenientes da Região de Mafra, em Portugal, não se pode seriamente pensar que a formação histórica dessa comunidade se deu em um vácuo no interior do universo colonial brasileiro; sobretudo, se pensarmos que sua origem remonta às atividades de garimpo implementadas na região a partir dos finais do século XVII. Qualquer comunidade que se tenha estabelecido no Brasil nesse período teve em sua constituição lingüística alguma influência do massivo contato entre línguas que estava em curso então. (LUCCHESI, 2000, p. 57 – grifos nossos)

Portanto, assim como Lucchesi (2000, p. 57), acreditamos que as diferenças que esperamos encontrar entre as duas comunidades, Mato Grosso e Luanda, se situam nas frequências de uso que podem revelar uma interferência mais direta do contato entre línguas ou “[...] *vis-à-vis* marcas diluídas que refletem residualmente uma influência indireta do contato entre línguas” (LUCCHESI, 2000, p.57).

Almeida e Carneiro (2008), em introdução a amostra de fala da comunidade de Mato Grosso, contam que

a primeira menção a esse povoado foi por nós encontrada em uma transcrição de um relatório feito pelo desembargador Luiz de Siqueira da Gama, no início do século, quando fez viagem pela Chapada para que, por ordem de D. João V, examinasse a maneira como trabalhavam naquelas minas e para que se escolhesse o lugar para ereção de uma nova vila, ‘bem como a melhor forma de arrecadação dos quintos do ouro’ (AMARAL, 1940:19). (ALMEIDA; CARNEIRO, 2008 s/p)

Em notas, as autoras afirmam, através das palavras do desembargador Luiz Siqueira, que em Mato Grosso, e nas áreas próximas, em que havia mineração, trabalhava uma população totalmente diversa:

Achavão-se a este tempo no dito riacho setecentos trabalhadores entre batêas e almocafres, além de outros que andavão em vários riachos, e alguns em novos descobrimentos, com que seguramente passavão de duas mil pessoas. Compunha-se este número de toda a variedade de gente, que para aquela parte tinha ocorrido, como Paulistas do Serro do frio e Minas-geraes, homens brancos de pequena esfera, que deste recôncavo, de muitas partes do sertão tinha ido, mulatos e negros, e entre todos havia vários criminosos (...) uns roubando, e outros matando, e logo que em algum ribeiro acertavão alguns com melhor pinta, caia aquella multidão... (28-29) (ALMEIDA; CARNEIRO, 2008, s/p sic)

O trecho supracitado do desembargador Luiz Siqueira, demonstra que havia muitos brancos pobres, vindos das mais diversas partes do sertão, mulatos e negros, o que nos indica que houve, contatos intra e interlingüísticos. Almeida e Carneiro (2008) informam

que nos relatórios o desembargador não informa se havia portugueses entre os brancos, fala apenas dos paulistas e mineiros. Mas

[...] sabe-se que portugueses foram para Mato Grosso explorar ouro. Com a guerra dos Marotos, muitos deles deixaram a vila de Rio de Contas às pressas. No entanto, como o povoado não é muito acessível, pode ser que alguns deles tenham se mantido lá, pois é fato que o povoado é constituído por pessoas de pele branca, como inclusive observou o antropólogo americano Harris na década de 50 do século XX quando estudou a região (cf. Harris, 1959). (ALMEIDA; CARNEIRO, 2008, s/p)

A população de Mato Grosso é quase exclusivamente formada por brancos descendentes de portugueses, que chegaram à Bahia no final do século XVII, e início do século XVIII, vindos da cidade portuguesa de Mafra, que ainda é umas das mais representativas famílias de Mato Grosso.

O site do IBGE (Instituto Brasileiro de Geografia e Estatística) demonstra que a ocupação da região de Rio de Contas se inicia na última década do século XVII, através do estabelecimento de uma rota de viagem entre Goiás e o Norte de Minas e a cidade de Salvador, capital da então Província da Bahia. Com o intuito de se estabelecer um "ponto de pouso", nesta rota de viagem, foi fundado o povoado de *Creoulos*, situado em um planalto da serra das Almas, na margem esquerda do Rio de Contas Pequeno, atual Rio Brumado. A descoberta de ouro, no leito do rio Brumado, atraiu à região grande número de garimpeiros que, subindo o rio, fundaram o povoado Mato Grosso, no bojo do grande afluxo de bandeirantes mineiros e paulistas para a região. Estes dados também são apresentados pelo projeto *Vertentes do português popular do Estado a Bahia*, vinculado ao Departamento de Letras Vernáculas da Universidade Federal da Bahia (UFBA), coordenado pelo Prof. Dr. Dante Lucchesi.

As evidências para defender a posição de que a variedade linguística formada no sertão baiano na área em que se situa a comunidade selecionada para este estudo é de natureza popular são ancoradas em pesquisas sócio-históricas voltadas para essa região e para a comunidade a ser estudada de forma específica.

É farta a historiografia sobre o tipo de processo de expansão do povoamento e a criação de vilas na Bahia. Carneiro e Almeida (2007) apoiam-se no padrão apresentado por Bandeira de Mello e Silva (1989, pp. 94-97), padrões notadamente reconhecidos como tendo envolvido colonos brasileiros, majoritariamente pobres, população autóctone e africanos:

- Um povoamento no nordeste da capitania com tendência à linearidade (devido às rotas de boiadas e das tropas);

- Uma grande dispersão com vilas localizadas em pontos de interseção das estradas que serviam de rotas para o gado, ouro, etc.; e
- Um grande número de estradas nas áreas de pecuária e mineração, ambas no sertão. Esse processo ocorreu, inicialmente, a partir das entradas baianas em busca de metais preciosos, da expansão pecuária exigida pela também expansão da economia açucareira e a expulsão ou extinção dos índios ou o seu aldeamento pelos jesuítas, franciscanos e capuchinhos. Tais aldeamentos transformaram-se posteriormente em lugares e vilas, dando, a partir daí, a uma das origens do processo de urbanização do interior da Bahia.

Portanto, a comunidade de Mato Grosso, situada no município de Rio de Contas representará a amostra do português popular do Brasil, nas análises a serem descritas e comentadas no capítulo posterior, diante dos critérios extralinguísticos descritos em tabela.

AMOSTRAS DO PORTUGUÊS POPULAR DO BRASIL				
	Nome	Sexo	Escolaridade	Faixa Etária*
1	J.A.M	M	Fund. incompleto	I
2	N.	M	Fund. incompleto	I
3	J.J.S.	M	Analfabeto	II
4	G. G. L.	M	Analfabeto	II
5	J. M.	M	Fund. incompleto	III
6	F. S.	M	Fund. incompleto	III
7	S. A. S.	F	Fund. incompleto	I
8	M. F. S.	F	Fund. incompleto	I
9	M.D.	F	Analfabeta	II
10	A. C. S.	F	Fund. incompleto	II
11	A.O. M	F	Analfabeta	III
12	I.S.L.	F	Analfabeta	III

*As faixas etárias I, II e III correspondem, respectivamente, aos informantes até 30 anos, 31 a 60 anos, a partir de 61 anos.

Quadro 3: Descrição da população amostra do PB

Lopes (2001, p. 114) acredita que numa pesquisa sociolinguística, a seleção dos informantes deve refletir as hipóteses iniciais e orientar a constituição de uma amostra representativa das variáveis sociais em estudo. Portanto, a comunidade escolhida, como já discutimos, é reflexo do contato entre línguas e no que tangem aos 12 (doze) informantes que compõem a amostra, todos se encontram na faixa etária de 18 a 76 anos e possuem, no máximo, fundamental incompleto (até a 4ª série).

4.4.1.2 Amostras da comunidade de fala de Luanda

O projeto *Em busca das raízes do português brasileiro*²⁴, coordenado pela professora Eliana Pitombo Teixeira, do Departamento de Letras e Artes da Universidade Estadual de Feira de Santana (UEFS), parte do pressuposto de que as similaridades observadas entre as variedades linguísticas das duas ex-colônias portuguesas, Brasil e Angola, devem-se a causas exógenas como o contato com línguas africanas e a transmissão linguística irregular. Hipótese essa assumida também nessa dissertação.

Atualmente o banco de dados do projeto é composto por amostras representativas do português angolano, dentre essas, amostras do português falado na capital - Luanda, que constitui um dos *corpora* de nossas análises.

A realidade linguística do país é, porém, bem diferente da brasileira. A população, majoritariamente concentrada nas maiores cidades costeiras, tem como língua materna uma ou mais das 40 línguas banto existentes no país. Sendo assim, língua portuguesa em Angola, na maioria das vezes, restringe-se a língua veicular, língua da administração, comunicação social e ensino, a identidade nacional é preenchida pelas línguas africanas (INVERNO, 2004, p. 2). Porém, cabe a ressalva de que somente a partir do início do século XIX, mais especificamente após a Independência do Brasil, que Angola se tornou uma colônia de povoamento extensivo, o que permite afirmar que o português angolano ainda está em formação.

De acordo com o site oficial de Angola, a estimativa da população angolana era cerca de 11.190.786 habitantes em 2005, distribuídos pelas 18 (dezoito) províncias administrativas do país, o que lhe confere uma densidade populacional de cerca de 09 (nove) habitantes por

²⁴ Projeto tem como objetivos centrais: a) formar banco de dados de falantes nativos e não nativos do português da comunidade de Luanda; b) descrever e analisar o material coletado nos níveis lexical, fonético-fonológico e morfosintático, comparando os resultados dessas descrições e análises com as do português brasileiro; c) produzir literatura teórica e aplicada para fundamentar e estimular futuras pesquisas no campo desta investigação.

km². As cidades mais populosas são Luanda, Huambo e Benguela. O site ressalta que a população da província de Luanda resultou do processo de deslocação das populações após a independência nacional e, por outro lado, as províncias do interior apresentam uma quebra de população.

Da grande extensão territorial de Angola (Cf. Figura 7), o *corpus* a ser trabalhado, nesta dissertação, se restringirá a amostras de falas do português falado em Luanda, capital de Angola.



Figura 7: Mapa territorial de Angola

Fonte: ANIP (Agência Nacional de Investimento Privado)

Luanda, capital de Angola, caracteriza-se como a maior cidade em extensão territorial e tem sido o grande centro urbano e costeiro do país, para onde convergiu a grande maioria da população, proveniente de todas as zonas do território nacional.

Teixeira (2013) afirma que na atual sincronia, os luandenses se veem defrontados a um “turbilhão linguístico”, como qualificam Pagotto e Duarte (2005) a situação experimentada pelos brasileiros no final do século XIX. A autora segue afirmando que

se pode referir o português moderno, introduzido pelos portugueses no século XIX, um português local que teria se desenvolvido na situação de contato e praticado pela “elite caluanda” ou “sociedade crioula” e, por outro lado, o português praticado pela maioria da população, a que veio se reunir uma grande massa de migrantes, falantes de línguas nacionais que aprenderam o português em circunstâncias adversas, em uma situação de transmissão linguística irregular. (p. 1)

No que tange ao contexto sócio-histórico de Angola, Fage (2010) acredita que a ocupação angolana passou por fases que vão desde a tentativa de apoderação das minas de

prata pelos colonizadores até a intensificação do tráfico negreiro, quando Angola se torna uma feitoria, um ponto de apoio para a guarda e embarque dos escravos. Cabe ressaltar, que esse momento, que se estende até a Independência do Brasil, no início do século XIX, foi marcado por acordos e lutas internas, e pela invasão dos holandeses.

Nestes encontros e desencontros de povos, os portugueses foram se tornando cada vez mais africanizados e não é de estranhar que tenham adotado dos africanos outros tipos de comportamento, incluindo o linguístico (TEIXEIRA, 2012). Assim como a autora, acreditamos na hipótese de uma influência linguística nesse processo de contato entre colonizadores e colonizados.

Para alcançar o objetivo desta pesquisa - identificar convergências/divergências entre as duas variedades do português, brasileiro e angolano,- trabalharemos com dados de língua falada de 12 (doze) amostras representativas do português popular de Angola, dentro dos critérios extralinguísticos descritos:

AMOSTRAS DO PORTUGUÊS POPULAR DE ANGOLA			
Nome	Sexo	Escolaridade	Origem
J. J.	M	Fund. incompleto	Luanda
S.	M	Fund. incompleto	Luanda
S.D.M.	M	Fund. incompleto	Luanda
D. C.	M	Analfabeto	Outra
T.	M	Fund. incompleto	Outra
F.A.S	M	Fund. incompleto	Outra
A.E	F	Fund. incompleto	Luanda
M.M.	F	Fund. incompleto	Luanda
J.	F	Analfabeta	Luanda
P. A. B.	F	Fund. incompleto	Outra
A.V.N	F	Fund. incompleto	Outra
T. M. N.	F	Fund. incompleto	Outra

Quadro 4: Descrição da população amostra do PA

Mais uma vez ressaltamos a importância na seleção dos informantes, uma vez que essa deve refletir as hipóteses iniciais e orientar a constituição de uma amostra representativa das variáveis sociais em estudo. Portanto, a comunidade escolhida, como já discutimos, é reflexo do contato entre línguas e no que tangem aos 12 (doze) informantes que compõem a amostra, todos se encontram na faixa etária de 18 a 76 anos e possuem, no máximo, fundamental incompleto (até a 4ª série), sendo que 02 (dois) deles são analfabetos.

4.4.2 Coleta e análise de dados

As gravações dos projetos supracitados foram realizadas através de conversas gravadas numa situação que Labov (1983, p. 266) chamou de “o paradoxo do observador”, pois “queremos observar a fala do falante quando ele não é observado”, se esperam registros de fala informal, descontraída, do vernáculo, mas o contexto oferecido ao informante é de entrevista, que se constitui numa situação formal.

As 24 (vinte e quatro) entrevistas que compõem as amostras descritas anteriormente, caracterizam-se por entrevistas do tipo DID (Diálogo entre Informante e Documentador), com duração média de 30 a 60 minutos, seguindo o modelo de entrevista sociolinguística semidirigida. O método utilizado implica em um diálogo de tema livre, no qual o entrevistador tentar neutralizar o ambiente de gravação no intuito de que o falante utilize o português em ambientes informais. A coleta do vernáculo é o objetivo maior das entrevistas e para tanto o projeto apóia-se nos pressupostos teórico-metodológicos propostos por William Labov.

Nas análises trabalharemos com dados orais de 24 (vinte e quatro) informantes, sendo que 12 (doze) amostras representam o português popular de Mato Grosso e 12 (doze) representam o português popular de Luanda, todos com baixo nível de escolaridade (até a quarta série) e/ou analfabetos. Poderemos visualizar melhor na tabela abaixo:

AMOSTRAS DOPORTUGUÊS POPULAR DO BRASIL					AMOSTRAS DO PORTUGUÊS POPULAR DE ANGOLA			
Nome	Sexo	Escolaridade	Faixa Etária*	Nome	Sexo	Escolaridade	Origem	
1	J.A.M	M	Fund. incompleto	I	J. J.	M	Fund. incompleto	Luanda
2	N.	M	Fund. incompleto	I	S.	M	Fund. incompleto	Luanda
3	J.J.S.	M	Analfabeto	II	S.D.M.	M	Fund. incompleto	Luanda
4	G. G. L.	M	Analfabeto	II	D. C.	M	Analfabeto	Outra
5	J. M.	M	Fund. incompleto	III	T.	M	Fund. incompleto	Outra
6	F. S.	M	Fund. incompleto	III	F.A.S	M	Fund. incompleto	Outra
7	S. A. S.	F	Fund. incompleto	I	A.E	F	Fund. incompleto	Luanda
8	M. F. S.	F	Fund. incompleto	I	M.M.	F	Fund. incompleto	Luanda
9	M.D.	F	Analfabeta	II	J.	F	Analfabeta	Luanda
10	A. C. S.	F	Fund. incompleto	II	P. A. B.	F	Fund. incompleto	Outra
11	A.O. M	F	Analfabeta	III	A.V.N	F	Fund. incompleto	Outra
12	I.S.L.	F	Analfabeta	III	T. M. N.	F	Fund. incompleto	Outra

*As faixas etárias I, II e III correspondem, respectivamente, aos informantes até 30 anos, 31 a 60 anos, a partir de 61 anos.

Quadro 5: Descrição da população amostra

É importante declarar que o fenômeno da concordância de número no sintagma nominal pode ser analisado sob duas perspectivas, atomística e não atomística.

Na perspectiva atomística considera-se como dado de análise cada um dos constituintes flexionáveis dos SNs plurais. Dessa forma, se o constituinte estiver formalmente marcado, considera-se presença de concordância, caso contrário, considera-se ausência. Porém, Scherre (1988, p. 64) chama atenção para o termo ‘concordância’, uma vez que não o considera apropriado a todas as situações, já que muitas vezes apenas um elemento do sintagma é formalmente marcado, como em (9), outras vezes tem-se SNs sem nenhuma marca formal de plural, como em (10). Por isso, a autora prefere tratar o fenômeno como indicação de pluralidade e não concordância, já que a “concordância gramatical implica harmonia formal

em pelo menos dois elementos de uma dada construção” (SCHERRE, 1988, p. 64), o que não ocorre em (9) e (10).

(9) **As** menina foi lá buscar e não pegou **as** brusa.

(10) Quando era **cinco hora**, a gente já ta no pouso de novo.

Nesta perspectiva de análise as diversas propriedades linguísticas relativas a cada um dos elementos nominais da amostra são analisadas cuidadosamente e o linguista pode trabalhar com um conjunto expressivo de variáveis²⁵, buscando as possíveis adequações ao seu objeto de estudo ou fenômeno em análise.

Já na perspectiva não atomística – conhecida também como sintagmática – analisa-se o SN como um todo, não havendo análises individuais dos elementos flexionáveis do SN, como propõe a perspectiva atomística. Sendo assim, leva-se em conta “[...] a presença de marca em todos os elementos flexionáveis do SN *versus* ausência de marca em pelo menos um dos elementos flexionáveis do SN” (SCHERRE, 1994, p. 39), considerando como aplicação da regra, a presença de plural formal em todos os elementos flexionáveis que constituem o SN, como em (11), e como não aplicação, a falta de pelo menos uma marca formal de plural, como em (12).

(11) Se eu pudesse eu tava cuidando **das** prantas.

(12) Ele é que sabe **as** históriaø.

As variáveis independentes²⁶ utilizadas por Scherre (1988), pioneira nos estudos da variação da concordância nominal no Brasil, nas quais Lopes (2001), Andrade (2003) e outras também se baseiam, foram também tomadas aqui como ponto de partida para a definição dos grupos de fatores trabalhados.

²⁵ No trabalho, na perspectiva atomística, sob a variação na concordância de número no SN, costuma-se utilizar as seguintes variáveis, sendo estas utilizadas por Scherre (1988), a citar: processos morfofonológicos de formação do plural; tonicidade dos itens lexicais singulares; número de sílabas dos itens lexicais singulares; posição linear do elemento no SN; classe gramatical do elemento nominal; marcas precedentes ao elemento nominal analisado; contexto fonético/fonológico seguinte ao elemento nominal sob análise; função sintática do SN (codificada em cada um de seus constituintes); animacidade dos substantivos; grau dos substantivos e dos adjetivos e formalidade dos substantivos e dos adjetivos.

²⁶ Chama-se de variáveis independentes os grupos de fatores que são considerados como possivelmente relacionados ao fenômeno em estudo (LOPES, 2001, p. 128).

Como já ressaltado em outros momentos, define-se, neste trabalho, como variável dependente²⁷ a presença ou ausência de marca de plural nos elementos do sintagma nominal. Portanto, observa-se nas amostras dois padrões básicos e opostos:

- (i) Uma variante redundante, em que se tem a marca de plural (morfema -s) em todos os constituintes flexionáveis do SN;
- (ii) Outra, simplificada, em que não se tem a marca de plural em todos constituintes do SN.

Dessa forma, toma-se como hipótese inicial que a presença de marca pode estar relacionada a 09 (nove) variáveis independentes, 03 (três) sociais e 06 (seis) linguísticas ou estruturais, as mais identificadas em estudos já realizados no âmbito do PB. São elas:

- (i) Configuração sintagmática do SN;
- (ii) Função sintática do SN;
- (iii) Número de constituintes absolutos do SN;
- (iv) Número de constituintes pluralizáveis do SN;
- (v) Posição do SN em relação ao verbo;
- (vi) Saliência fônica.

Quanto aos grupos de fatores externos ou variáveis sociais, serão considerados para a amostra do português brasileiro:

- (i) Sexo:
 - Feminino;
 - Masculino.
- (ii) Faixa etária:
 - Faixa I – até 30 anos;
 - Faixa II – de 31 a 60 anos;
 - Faixa III – a partir de 61 anos.

²⁷ Chama-se variável dependente o fenômeno que está sendo estudado. Considera-se dependente por se ter como pressuposto que a presença de marca não é aleatória, mas condicionada por (ou dependente de) outras variáveis, as independentes (LOPES, 2001, p. 129).

Para amostra do português angolano:

- (i) Sexo:
 - Feminino;
 - Masculino.

- (i) Origem:
 - Luanda;
 - Outra.

Nestes grupos de fatores, cabe uma ressalva para o fator *faixa etária* e *origem*, já que o primeiro restringe-se a amostra do PB e o segundo a amostra do PA. A escolha e/ou exclusão desses fatores justifica-se diante das prioridades desta pesquisa, compor amostra do português popular. Portanto, o banco de dados, ainda em construção, do projeto *Em busca das raízes do português*, não ofereceu informantes que pudessem preencher os requisitos preliminares para composição do *corpus*, adotando o grupo de fator faixa etária. Desta forma, substituímos o grupo de fatores faixa etária, por origem, não sendo esse menos ou mais importante que aquele.

Para o tratamento quantitativo dos dados foi utilizado um programa computacional apropriado, conhecido como GOLDVARB X, uma versão atualizada VARBRUL (Programa das Regras Variáveis). Esta atualização do programa permite sua utilização em ambientes Macintosh, Windows e Linux, operando de forma semelhante em todos ambientes, traz também de novidade a pesquisa de referência cruzada e suas múltiplas ferramentas, o que possibilita ao pesquisador encontrar facilmente o que deseja. Além disso, o programa apresenta resultados com uma casa decimal (SANKOFF; TAGLIAMONTE; SMITH, 2005), o que tem favorecido resultados mais precisos.

Portanto, os dados serão submetidos a análises quantitativas que visem estabelecer as similaridades e divergências entre o português falado em Mato Grosso e em Luanda, respectivamente, amostras do português do Brasil e de Angola, no intuito de reunir características relevantes à interpretação do fenômeno da concordância nominal e a caracterização das variedades em análise, momento em que, retomando nossa epígrafe, *toda a teoria deve ser posta em prática e toda a prática obedecerá a uma teoria*.

5 A CONCORDÂNCIA NOMINAL DE NÚMERO SOBRE O OLHAR DA SOCIOLINGUÍSTICA QUANTITATIVA

“A linguagem e a vida são uma coisa só. Quem não fizer do idioma o espelho de sua personalidade não vive; e como a vida é uma corrente contínua, a linguagem também deve evoluir constantemente.”

João Guimarães Rosa – Museu da Língua Portuguesa

A língua portuguesa em suas origens, suas especificidades, suas variedades e tendências de mudança entrecruza-se com o objetivo almejado, neste trabalho, de identificar convergências/divergências entre essas duas variedades do português popular, brasileira e angolana. Para tanto, trazemos como ponto de partida às reflexões a colocação colossal de João Guimarães Rosa – que tem o seu lugar mais que merecido numa das paredes do Museu da Língua Portuguesa²⁸, em São Paulo. Talvez sejam as palavras mais belas em língua portuguesa sobre a linguagem e, portanto, escolhida para iniciar o *fim* de um *começo* de análises sob a concordância de número no SN em variedades do PB e PA.

O capítulo em que cabem as análises das variáveis linguísticas e sociais, descritas no capítulo anterior, que julgamos atuar sobre a concordância de número entre os elementos do sintagma nominal, baseia-se na Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 2006), na teoria dos 4M (MYERS-SCOTTON; JAKE, 2000) e nos pressupostos da Transmissão Linguística Irregular (LUCCHESI, 2009, 2008, 2003), através dos quais acreditamos ser possível compreender o processo de variação na concordância nominal na variedade do português nos *corpora* representativos das duas ex-colônias de Portugal, Brasil e Angola.

Separamos as estruturas sintagmáticas nominais em que as marcas de plural não se realizam em todos os itens flexionáveis e aquelas em que há concordância em todos eles. Dessa forma, pretendemos apontar os fatores que podem motivar as marcas explícitas de plural nos sintagmas nominais, além de colaborar para a apreensão de convergências/divergências entre as duas variedades e, em última instância, para os debates sobre as origens do Português do Brasil (LUCCHESI, 2009, 2008, 2003, 1998; NARO, 2003; NARO e SCHERRE, 2007).

²⁸ <http://www.museulinguaportuguesa.org.br>

5.1 ANÁLISE DOS DADOS

Como já relatado no capítulo que trata da metodologia, esta pesquisa trata especificamente da concordância de número plural entre os elementos flexionáveis do sintagma nominal. Os resultados que constam nas análises são fornecidos pela ferramenta computacional GOLDVAB X, que oferece resultados estatísticos das variáveis, em termos de frequência, oferece também pesos relativos, que medem efeitos simultâneos de todas as variáveis incluídas na análise.

Os dados serão analisados através da abordagem sintagmática ou não atomística, no intuito de certificar-nos do nível de aplicação da regra de concordância de número entre os elementos do sintagma nominal. Portanto, fica estabelecido que o valor positivo é atribuído a SNs cujos constituintes são todos marcados, como em (13), e àqueles SNs em que nem todos os elementos recebem marcação de plural é atribuído o valor negativo, como em (14) e (15). Para tanto, nas amostras analisadas, foram identificados três contextos em que a marca explícita de plural pode ser encontrada:

(i) Todos os elementos flexionáveis do SN

(13) *Não aprendem muito devido às novelas.*

(ii) Em alguns elementos flexionáveis do SN

(14) *Eu sempre gosto de escutar umas boaØ músicaØ.*

(iii) Em apenas um dos elementos flexionáveis do SN

(15) *As amizadeØ ali foi ótima.*

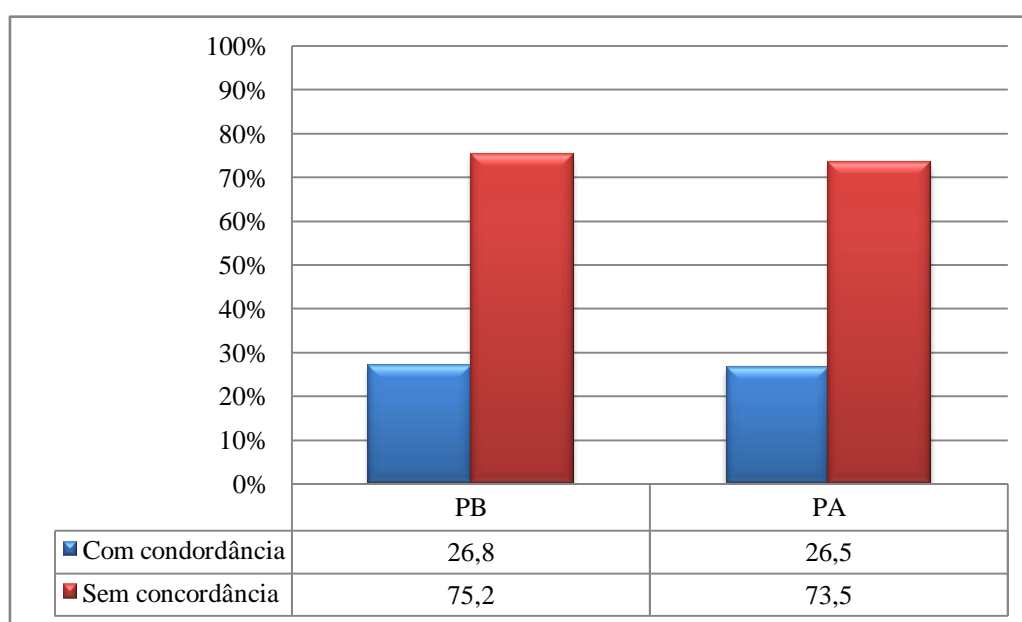
Assim, ressalta-se que para a análise, através da variável dependente binária, fica estabelecido que a aplicação da regra ocorre se houver presença de marcas em todos os itens pluralizáveis do SN, enquanto que a ausência dessa marca, seja em apenas um dos elementos do SN sob análise, constitui-se falta de concordância.

Foram analisados 1549 SNs, dos quais 933 SNs são representativos do português brasileiro e 620 SNs do português angolano. Através da *Tabela 4* podemos visualizar os resultados obtidos, no que tange a marcação e não marcação de plural, nas amostras.

Tabela 5: Taxas de uso da concordância no SN em amostras representativas do PB e PA

SINTAGMA NOMINAL	PB	%	PA	%
Com concordância	250/933	26,8	164/ 620	26,5
Sem concordância	683/933	73,2	456/ 620	73,5

Os resultados indicam que em ambas as amostras há uma tendência ao uso do valor negativo na indicação do plural nos SNs (73,2% para o PB, 73,5% para o PA) em relação ao valor positivo (26,8% para o PB, 26,5% para o PA), o que equivale a dizer que, ainda em uma análise preliminar, se observa que as duas variedades do português apresentam resultados similares no que tange à variação na marcação de plural em todos os elementos flexionáveis do SN. Tais resultados podem ser visualizados no gráfico que segue.

**Figura 8:** Frequência da variação na concordância no SN em amostras representativas do PB e PA

Antes de iniciarmos as análises quantitativas do fenômeno sob as variáveis – linguísticas e sociais -, faremos uma breve análise qualitativa dos dados, mesmo que esta não esteja inserida na proposta inicial. A opção de tal análise surge diante de alguns dados encontrados na fase de levantamento de dados.

A primeira ocorrência encontrada que merece destaque são os sintagmas em que apenas o primeiro elemento não recebe marca de plural, contrariando o que até então tínhamos visto em Shcerre (1988), que se baseia na “condição de distintividade”, em que a primeira posição favorece a presença de marca de plural em qualquer que seja a categoria.

Vejam os 12 (doze) casos encontrados nos *corpora*, sendo 10 (dez) do PA e apenas 02 (dois) do PB:

PB

[Indefinido + Nome]

(16) Só essa coisas que num vende aqui... aparelho de som.

[Determinante + Possessivo + Nome]

(17) Eu quase num participava das coisa não porque eu cuidava do meus avó que era velhinho.

PA

[Determinante + Nome]

(18) As crianças que tamo a ver aqui agora tão a aprender só coisas assim na ruas então, todas aquelas faz com que a crianças as veze as coisa da escola te passa na mente, mete as coisa de fora.

(19) Eu vejo que aquilo ainda vai conjugar com as minhas valore que eu quero pelo menos conseguir o dias ainda não posso dizer, não posso sonhar com aquilo que eu ainda não vi.

(20) Lá, quer dizer, quando você gostar duma miúda você primeiro tem que falar ainda os pais, falar a famílias.

(21) Luanda é só desenrascar, quer dizer, a vidas, desenrascar algumas coisa.

(22) Eu falo o Umbundu com quem, quer dizer, com meus amigos do próprios.

[Adjetivo + Nome]

(23) Logo vinha um, um... uma Toyota Corola ai de cima do Morro da Luz, com alta velocidades.

[...Nome...Adjetivo...]

(24) Eu acho que sim devem aprender a língua nacionais na escola.

(25) As pessoa que estão mais na cidade são pessoa mais desenvolvidas já.

[Determinante + Possessivo + Nome]

(26) E a vez eu fico assim a ver a revista, ai meu Deus! Que tem é que essa cegueira vai sair na minhas vistas.

(27) Vou na minhas tias e são mais velhas não falam português.

(28) Muito difícil e eu gosto de ver a minhas tias a falar principalmente o Kimbun... o Umbundu.

Como pode ser observado a não marcação no elemento de primeira posição acontece em diferentes estruturas sintagmáticas e a extensão do sintagma não parece ser um fator que impeça essa ocorrência, já que em (17), (24), (25), (26), (27) e (28) temos sintagmas compostos com mais de dois elementos e a inserção do morfema de plural apenas nos elementos de segunda e terceira posição. No que se refere ao fator função sintática, temos SN com função de sujeito, como em (16), objeto direto, como em (19), objeto indireto como em (17), complemento, como em (18). Com a leitura desses dados, retomamos o que foi discutido no capítulo anterior, referente aos pressupostos teóricos, e passamos a lançar reflexões sobre os casos explicitados em (16) a (28).

Sobre a hipótese funcionalista que se baseia na “condição de distintividade” apresentada por Kiparsky (1972, p. 195, *apud* SCHERRE, 1997, p. 182), acredita-se que “há uma tendência para a informação semanticamente relevante ser mantida na estrutura superficial”. Portanto, em um contexto comunicativo, um segmento que seja necessário à transmissão eficiente da mensagem tenderá a não desaparecer, o que leva Scherre (1988) a acreditar que há uma tendência à inserção da marca de plural, apenas no elemento de primeira posição. Mas como explicar os casos de (16) a (28), em que não temos marcação de plural no primeiro elemento do SN?

Na ótica da teoria dos 4M segundo a qual a ordem em que os morfemas são adquiridos numa língua obedece a uma tipologia que, assim como explica a aquisição desses elementos, pode servir para entender o processo da variação dos mesmos (LOPES, 2001, p. 34), também

nos questionamos diante dos casos de (16) a (18) em que não se adquire o morfema de plural no primeiro elemento.

Lopes (2001) acredita que os *early system morphemes* seriam ativados concomitantemente aos *content morphemes* para atender necessidades funcionais imediatas, e a atuação destes explicaria, por exemplo, a concentração de marcas de plural no primeiro elemento do SN, mas o que poderia explicar a não marcação neste primeiro elemento? Já que os morfemas de plural não são todos ativados no mesmo nível, em que momento seria ativado este morfema de plural no primeiro elemento do SN?

Certamente muitas das nossas respostas estariam presentes numa análise atomísticas, em que analisaríamos o efeito de cada variável em cada um desses elementos, o que poderia evidenciar as variáveis que contribuem para a não aleatoriedade da presença/ausência das marcas formais de concordância gramatical nos elementos flexionáveis do SN. No entanto, nossa intenção em trazer esses dados e questionamentos não é, neste momento, de apresentar hipóteses que possam explicar essas ocorrências em amostras do PB e do PA, mas demonstrar aos pesquisadores sociolinguistas que as variedades do português a cada momento nos surpreendem com sua variabilidade. Portanto, resta-nos buscar mais dados para que se confirme se há uma tendência ou se esses são apenas dados isolados do português e que fatores motivariam a ausência de marca de plural no primeiro elemento no SN.

Outra ocorrência encontrada nos *corpora* foi o sintagma nominal em que o segundo constituinte sofre um processo fonológico em que se acrescenta um segmento fonético em posição final de palavra. Encontramos 29 (vinte e nove) casos, em diferentes estruturas sintagmáticas, em que temos o acréscimo de um segmento fonético [i], 13 (treze) casos do PB e 16 (dezesesseis) do PA. Vejamos alguns deles:

PB

[Numeral + Nome]

(29) São dois irmãos homens, quatro mulhere.

[Indefinido + Nome]

(30) O café vende pra Conquista, Livramento, esses lugare aí.

[Determinante + Nome]

(31) As veze ele pudece, é fácil, assim mesmo já cai, ainda cai no chão.

(32) Os rapaze mesmo quando vão pra São Paulo

PA**[Determinante + Possessivo + Nome]**

(33) Eu vejo que aquilo ainda vai conjuar com as minhas valore que eu quero pelo menos conseguir o dias ainda não posso dizer, não posso sonhar com aquilo que eu ainda não vi.

[Indefinido + Nome]

(34) Então aquilo já é, me corta aquelas festa porque às vezes não vou conseguir aqueles valore que eles tão a pedir.

(35) No Huambo também eu tinha uns familiare.

(36) Alguns familiare que foram pra escola escrevem umbundo e português.

[Numeral + Nome]

(37) Uma semana você vai duas veze, três veze.

(38) Éramos dai duas menina e dois rapaze.

[Determinante + Nome]

(39) Os cobradore eles falam “eu vou fazer a via”.

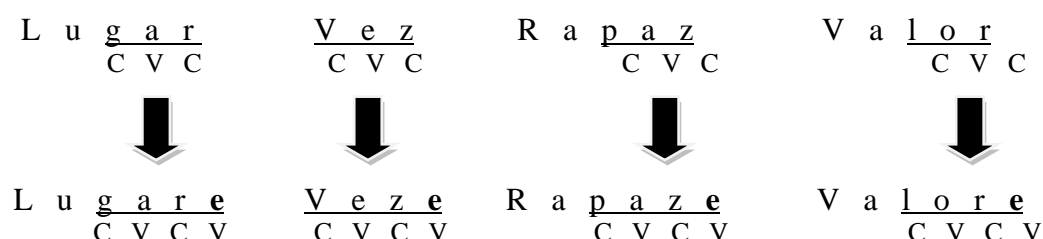
(40) Eu digo assim e uma coisa eu só podia saber só dos doutore fiz todas analise mas eles só me disseram que é infecção.

(41) Por exemplo quando vêm as mulhere dele vão pra praia.

[Indefinido + Possessivo + Nome]

(42) Fui falar com as primas com outros meu familiare eles me perguntaram assim: “vais pra onde?”

Inicialmente pensamos em defini-los como paragoge²⁹, já que de acordo com Xavier e Mateus (1990, p. 281), a paragoge pode ser definida como o “acrescentamento de um segmento fonético em posição final de palavra” e Massini-Cagliari (2005, p. 275) a entende como “um processo fonológico que acrescenta uma vogal neutra /e/ após sílabas terminadas por codas consonantais, a fim de transformar essas sílabas em estruturas canônicas do tipo CVCV”. Como pode ser observado nas palavras integrantes dos sintagmas de (29) a (42) há o acréscimo da vogal /e/ no final das sílabas, transformando as estruturas em CV:



Rodrigues (2010, p. 217) acredita que no contexto da reestruturação vocabular nos crioulos, está em evidência a emergência da sílaba CV e sua importância dentro da organização prosódica das línguas.

Esse aspecto estrutural ancora-se no contexto sócio-histórico de todo e qualquer crioulo: a situação de dominação faz com que os povos subjugados tenham que entender e reproduzir, num curto espaço de tempo, a fala de seus dominadores. No caso de Cabo Verde, os falantes de diferentes línguas africanas do grupo Mandé e línguas Oeste-Africanas, como Mandinga, Bambará, Balanta e línguas Afro-Asiáticas, como o Wolof, fizeram com que os vocábulos portugueses adquirissem uma forma silábica mais próxima às suas línguas de origem, que dão preferência ao tipo CV, na maioria dos casos.

Rodrigues (2010) trabalha com processos fonológicos de acréscimo – prótese, epêntese e paragoge – atuantes na formação da sílaba ótima no Caboverdiano e nas línguas crioulas em geral e acredita que para falarem o português, os escravos não se guiavam pelas regras da língua portuguesa, mas pelas de sua própria língua. O autor acredita que as restrições da língua materna dos africanos podem ter desencadeado processo de interferência capaz de gerar *outputs* dessemelhantes dos *inputs* recebidos por causa da re-hierarquização.

²⁹ A paragoge é um dos metaplasmos por adição de fonemas, sendo acrescentados um ou mais ao final da palavra. De acordo com Xavier e Mateus (1990, p. 281), a paragoge pode ser definida como o “acrescentamento de um segmento fonético em posição final de palavra”. Encontra-se definição semelhante em *The Concise Oxford Dictionary of Linguistics*: “Paragoge: The development of an additional sound or sounds at the end of a word [...]” (MATTHEWS, 1997).

Abrimos espaço para destacar o trabalho de E-Ching Ng (2013)³⁰, intitulado *Paragoge as an indicator of language contact*, em que trata foneticamente o fenômeno da paragoge relacionado ao contato linguístico, afirmando que

I have found that *paragoge* is in fact attested diachronically, but many cases coincide with well-known cases of language contact. [...] However, contact with African languages is certainly a plausible explanation for *paragoge* in early Brazilian Portuguese. (2013, p.1)

As leituras bibliográficas nos conduziam a acreditar que os casos de (29) a (42) sofriam processo de paragoge, mas nossas buscas, nos *corpora*, por casos similares no singular, como “a mulhere”, foram frustradas. A ausência de tais ocorrências nos induziram a assumir os casos supracitados com valor positivo na marcação de plural.

Sendo as ocorrências casos de paragoge ou simples apagamento da marca de plural, cabe a reflexão por se apresentarem similarmente em duas variedades do português, PB e PA. Teria os casos de (29) a (42) relação como o processo de contato linguístico ocorrido nas duas ex-colônias? Certamente, a comparação entre PB e PA, principalmente no campo fonético, constitui tema de trabalho que vale a pena prosseguir, partindo dos poucos dados aqui apresentados.

Ainda na perspectiva qualitativa, que nos propomos a fazer, destacamos os casos dos sintagmas em que temos um determinante e um nome no diminutivo³¹. Vejamos as ocorrências:

PB

(i) SNs não marcados

[Determinante + Nome]

(43) Vendeu *as terrinha* que nós tinha tudo.

(44) Esperar *os cardeirãozinho* de comida doz’outo, de sogro, de sogra em casa, é chato

[Indefinido + Nome]

(45) Briga aqui só quando uns, os cara bebe *uns golinho* a mais.

³⁰ Pôster apresentado no LSA 2013 (*Linguistic Society of America*), em Boston.

³¹ A formação de diminutivo é muito produtiva no Português Brasileiro, que exhibe duas maneiras de derivar as formas de diminutivo das palavras não verbais (nomes e adjetivos/advérbios): a sufixação -inho(a) e a sufixação -zinho(a). (LEE, 1999, p. 114)

- (46) A gente fazer uns... *uns pauzinho* aqui no chão, óh, parecendo com um quadrinho assim, óh, e marrava tudo com *umas cordinha*, marrava, fazia *uns boizinho* e botava lá dentro.
- (47) A gente ficava fazeno *uns carrinho*, fazia umas... umas brincadeira, umas... umas...
- (48) Briga aqui só quando uns, os cara bebe *uns golinho* a mais, eles faz confusão com outro.

[Demonstrativo + Nome]

- (49) *Aqueles pezinho*, com *aqueles pezinho* de café que você viu.
- (50) Eu mehmo só tenho *esses pezinho aqui*.

(ii) SNs marcados

[Determinante + Nome]

- (51) Todo mundo já leva *as esterinhas*, bota tudo ali cada um vai na carreira, vai botano, ói.

[Indefinido + Nome]

- (52) Hoje mehmo eu vim de Livramento, eu comprei *umas coisinhas* bem pa... paguei quatro reais.

PA

(i) SNs não marcados

[Indefinido + Nome]

- (53) Tem *uns bichinho* que eles fala que é minhoca.

[Demonstrativo + Nome]

- (54) *aqueles bichinho* como um micróbio então aquilo cavamos nos gimbo.

(ii) SNs marcados**[Determinante + Nome]**

(55) Então lhes metemos *nos azulinhos* da polícia.

Como é observado nos caso de (43) a (55), o levantamento de dados nos permitiu perceber que dos 17 (dezesete) casos em que o segundo elemento do SN é um substantivo (nome) que sofre o processo de sufixação (acréscimo do morfema –inho (a) e zinho), apenas 03 (três) apresentam o morfema de plural, permitindo que o SN seja flexionado em número, como em (51), (52) e (55). Os poucos dados também apontam uma maior frequência de SN, em que temos o acréscimo dos sufixos supracitados, no português brasileiro: das 17 (dezesete) ocorrências, 14 (quatorze) para o PB e 03 (três) para o PA. Resta-nos verificar até ponto a presença de uma palavra que sofre o processo gramatical de sufixação de diminutivo pode influenciar ou não a marcação de plural no SN, nas variedades do português. A resposta poderá estar presente numa análise atomística dos dados, a qual, por limitação de tempo, não nos propomos a realizar.

Encerramos aqui a apresentação das ocorrências que nos despertaram interesse, o que poderá vir a ser objeto futuro de pesquisa. Partiremos agora para a proposta metodológica de análise quantitativa.

Como já afirma Lopes (2001, p. 160) fazer análise de dados numa perspectiva sociolinguística é normalmente entendido como buscar razões, no contexto social, para os diversos usos linguísticos. É seguindo essa idéia que iniciamos nossas análises quantitativas, diante da variação na concordância nominal de número, na tentativa de fazer um exame minucioso das variáveis sociais, selecionadas pelo programa computacional GOLDVARB X. Para tanto, trabalharemos as variáveis sociais selecionadas para o PB, em seguida as selecionadas para o PA. Após a análise das variáveis sociais, são tratadas as variáveis linguísticas.

5.2 VARIÁVEIS SOCIAIS

Como afirma Scherre (1988, p. 430), a influência de variáveis sociais em fenômenos linguísticos variáveis, principalmente nos casos em que uma das variantes é estigmatizada, não constitui nenhuma novidade dentro da literatura linguística. Estudos, como os já descritos em capítulos anteriores, nos permitem entender que, em situação de variação que não envolve mudança, ou seja, na variação estável, normalmente se verifica uma relação proporcional entre as variáveis sociais e uso de formas de prestígio. Portanto, cabe-nos compreender que

não se pode mais seriamente defender que o linguista deve limitar suas explicações da mudança às influências mútuas dos elementos linguísticos, definidos pela função cognitiva. Nem se pode argumentar em qualquer sentido sério que o sistema linguístico em mudança é autônomo. [...] não é possível concluir uma análise das relações estruturais dentro de um sistema linguístico, sem considerar as relações externas. (LABOV, 1972 [1965], p. 181-182 *apud* LUCCHESI, 2004, p. 186)

Essa perspectiva, iniciada por Labov e adotada por outros tantos pesquisadores, possibilita a inserção do componente social na abordagem da mudança linguística e demonstra a importância da inclusão de condicionantes sociolinguísticos, já que:

todos esses estudos demonstraram que a mudança não é apenas uma função do sistema linguístico, mas uma função da interação da estruturação interna da língua com o processo social em que ela se realiza; sendo que, em muitos casos, os fatores funcionais e estruturais internos podem ser totalmente sobrepujados pela força das disposições sociais. (LUCCHESI, 2004, p. 185)

Considerando essas proposições, em que se coloca para a investigação da variação e da mudança da língua a necessidade de se correlacionar fatores linguísticos e extralinguísticos, buscamos verificar como o contato do falante com o universo exterior à comunidade pôde interferir no seu desempenho em relação ao uso da regra da concordância nominal no SN (ANDRADE, 2003, p. 123). Verificou-se que as variáveis sociais selecionadas como relevantes pelo programa estatístico GOLDVARB X foram gênero e faixa etária, para amostra do PB, e gênero e origem, para amostra do PA.

O fator gênero foi selecionado nas duas amostras. Trabalhos como o de Scherre (1988), Lopes (2001) e Andrade (2003) demonstram a importância de se controlar a diferença/variação entre os gêneros, já que esta é uma variável que tem muita importância quando se deseja estabelecer uma relação entre a variação e a mudança linguística (LOPES, 2001, p. 160). O fato de a variável gênero ter sido selecionada nas duas amostras de

variedades do português, já nos traz indícios de sua relevância ao tratar da variação na concordância de número nos SNs.

5.2.1 Português brasileiro: variáveis sociais

Para amostra do português brasileiro o programa computacional GOLDVARB X selecionou como relevantes, com significância 0.009, as variáveis gênero e faixa etária, cujos resultados apresentaremos em porcentagens e pesos relativos.

5.2.1.1 Gênero

A variável gênero tem sido focalizada pela sociolinguística, pois, do ponto de vista social, homens e mulheres têm papéis diferentes e estão, portanto, expostos a situações diversas, o que pode influenciar a escolha de uma ou outra forma linguística. O efeito da variável gênero sobre a presença de marcas de plural nos SNs pode ser observado, inicialmente, a partir da tabela e gráfico a seguir.

Tabela 6: Taxas de uso da concordância no SN em função do gênero no PB

GÊNERO	Frequência	%	Peso relativo
Masculino	47/421	11,2	.28
Feminino	203/512	39,6	.68

Na análise geral, essa variável foi selecionada e, como se vê, temos uma considerável diferença entre os pesos relativos (de .40), havendo uma tendência ao uso da concordância plural dos SNs entre as mulheres, o que ratifica a posição de Scherre (1988, p. 427) de que é normalmente aceito que os falantes do sexo feminino tendem a se aproximar da norma estabelecida mais do que os falantes do sexo masculino. Vejamos no gráfico os resultados através das porcentagens:

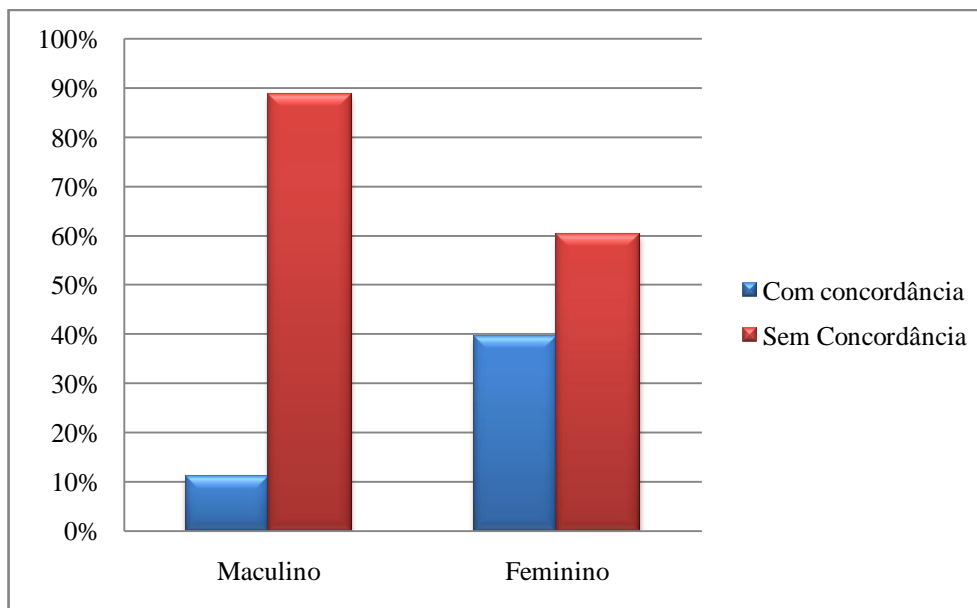


Figura 9: Efeito da variável gênero na variação na concordância de número nos SNs do PB

Entre os falantes do sexo masculino temos 11,2% para marcação de plural, contra 88,8% para ausência da marca de plural, entre os falantes do sexo feminino temos 39,6% contra 60,4%. Os resultados referentes à liderança das mulheres em direção ao padrão são reflexo, na comunidade de Mato Grosso, do acesso à educação, já que este direito, na época, se restringia às mulheres, independentemente de sua classe social (LUCCHESI, 2000).

Os resultados refletem o que geralmente se espera de uma comunidade, nos termos de Bortoni-Ricardo (2001), rurbana, já que representam um contínuo dialetal, que se estende das variedades rurais estigmatizadas à variedade urbana padronizada. Dessa forma,

em tal contínuo, adjacentes aos vernáculos altamente estigmatizados, encontram-se variedades não padrão que podem ser denominadas 'rurbanas'. Essas variedades são usadas por falantes de classes mais baixas, não alfabetizadas ou semialfabetizadas, que vivem na cidade, mas que, na maioria dos casos, têm antecedentes rurais, e pela população que vive em áreas rurais, onde já se vê introdução de tecnologia. (BORTONI-RICARDO, 2001, p. 21)

Em uma análise comparativa, que segue em tabela, vemos que nossos resultados contrariam os de Andrade (2003) que constata que os homens estão à frente na fixação da regra de concordância de gênero no dialeto descrioulizante de Helvécia - Ba. Vejamos tabela comparativa do estudo pioneiro de Scherre (1988), o de Lopes (2001) e o de Andrade (2003), que hoje são representativos ao tratar da variação na concordância de número no SN, no PB.

Tabela 7: Dados comparativos: uso da concordância no SN em função do gênero no PB

GÊNERO	SCHERRE (1988) <i>Rio de Janeiro</i>		LOPES (2001) <i>Salvador</i>		ANDRADE (2003) <i>Helvécia - Ba</i>		<i>Mato Grosso - Ba</i>	
	%	Peso relativo	%	Peso relativo	%	Peso relativo	%	Peso relativo
Masculino	65	.41	80%	.47	10	.63	11,2	.28
Feminino	77	.59	82%	.53	9	.39	39,6	.68

O quadro comparativo demonstra que nossos dados dialogam com os dados de Scherre (1988) e Lopes (2001) em que há uma tendência de que as mulheres realizem mais a concordância do que os homens, diferentemente do que demonstrou Andrade (2003) em suas análises, já que os homens em Helvécia - Ba foram expostos a diversos contextos sociais, viagens, enquanto que as mulheres dessa localidade se mantiveram mais dedicadas as tarefas do lar e à educação dos filhos.

5.2.1.2 Faixa etária

A variável faixa etária revela-se de fundamental importância no estudo dos fenômenos variáveis, uma vez que é essa variável que vai indicar se determinado fenômeno de variação linguística está estável ou está em progresso num estudo em tempo aparente. Sabe-se também que, nos eixos sociais, por exemplo, em geral, os falantes mais jovens são mais inovadores e os falantes mais velhos costumam preservar formas mais antigas. O efeito da variável faixa etária sobre a presença de marcas de plural nos SNs pode ser observado, inicialmente, a partir da tabela e gráfico que segue.

Tabela 8: Taxas de uso da concordância no SN em função da faixa etária no PB

FAIXA ETÁRIA*	Frequência	%	Peso relativo
Faixa I	177/410	43,2	.70
Faixa II	42/347	12,1	.31
Faixa III	31/176	17,6	.38

*As faixas etárias I, II e III correspondem, respectivamente, aos informantes até 30 anos, 31 a 60 anos, a partir de 61 anos.

Como podemos observar a faixa etária mais nova é a única que favorece o uso da concordância, com .70 de peso relativo, sendo a faixa que tende a marcar mais os elementos do sintagma nominal com o morfema de plural; a segunda faixa, falantes entre 31 e 60 anos, têm um peso relativo bem menor que os mais novos, uma diferença de .39, sendo essa diferença reduzida a .32 ao comparar faixa I e faixa III.

Nossos resultados contrariam os de Lopes (2001) e Scherre (1988) que demonstram que a faixa etária mais velha é favorecedora do uso da concordância, enquanto os mais jovens da variante inovadora. Por outro lado, os dados dialogam com os de Andrade (2003). Vejamos tabela comparativa:

Tabela 9: Dados comparativos: uso da concordância no SN em função da faixa etária no PB

FAIXA ETÁRIA	SCHERRE (1988) ³² <i>Rio de Janeiro</i>		LOPES (2001) ³³ <i>Salvador-Ba</i>		ANDRADE (2003) ³⁴ <i>Helvécia - Ba</i>		<i>Mato Grosso-Ba</i>	
	%	Peso relativo	%	Peso relativo	%	Peso relativo	%	Peso relativo
Faixa I	72	.49	81	.45	14	.74	43,2	.70
Faixa II	75	.54	80	.46	9	.48	12,1	.31
Faixa III	68	.47	85	.59	4	.23	17,6	.38

Assim como Andrade (2003, p. 125) – para comunidade de Helvécia – acreditamos que esses resultados confirmam que, em relação à concordância de número no SN, o dialeto de Mato Grosso está passando por um processo de mudança em curso, com a aquisição das regras de concordância nominal pelo grupo de falantes mais jovens. Conforme vimos em capítulos anteriores, o dialeto de Mato Grosso, passou pelo processo de contato linguístico, entre o português e as línguas africanas, e as heranças desse contato linguístico podem ser observadas ainda hoje, principalmente, no léxico e na morfologia, no nosso caso em especial,

³² Scherre (1988) trabalha com Faixa I correspondendo de 15 a 25 anos; Faixa II de 26 a 49 anos; Faixa III acima de 50 anos.

³³ Lopes (2001) trabalha com quatro faixas etárias, Faixa I correspondendo de 15 a 24 anos; Faixa II de 25 a 35 anos; Faixa III de 45 a 55 anos; Faixa IV acima de 65 anos. Portanto consideraremos nesta análise apenas as três últimas faixas etárias. Então onde lê Faixa I entende-se, nos parâmetros de Lopes (2001), por faixa II, onde lê Faixa II entende-se faixa III e onde lê faixa III entende-se faixa IV.

³⁴ Andrade (2003) trabalha com Faixa I correspondendo de 21 a 40 anos; Faixa II de 41 a 60 anos; Faixa III mais de 60 anos.

na expressiva redução da morfologia flexional nos nomes. No gráfico, poderemos observar a variação na concordância de número entre os nomes diante da variável faixa etária.

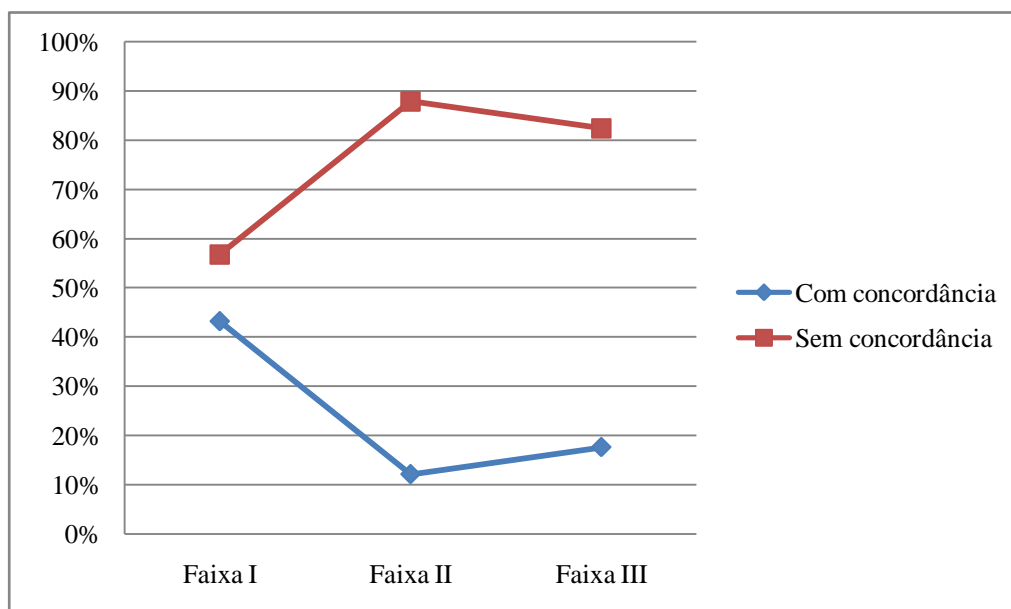


Figura 10: Efeito da variável faixa etária na variação na concordância de número nos SNs do PB

Como pode ser observado no gráfico acima a distribuição curvilínea das faixas etárias revela uma acentuada elevação no uso da forma de prestígio pelos jovens, sendo os altos índices da forma estigmatizadas representadas pelas faixas II e III, fazendo com que a realização da concordância de número plural aumente, conforme diminui a idade dos informantes analisados. Nesse sentido Andrade (2003, p. 126) acredita que

a baixa taxa de uso da regra de concordância pelos falantes mais velhos e a crescente utilização desta pelos mais jovens demonstra que o dialeto deixou-se ‘impregnar’ pelo contato com variedades outras do português, que se tornaram acessíveis pelo desenvolvimento social do Brasil [...].

Baxter (2001, p. 188-189) também, em estudo da concordância no sintagma nominal do português dos Tongas, identifica

a faixa etária dos mais velhos (ou primeira geração) como a mais desfavorecedora da concordância, diante de terem sido expostos a dados lingüísticos primários mais deficientes e incompletos [...]. Dessa forma, as primeiras gerações, diante dessa situação, têm propensão a demonstrarem mais relação estreita com essa história de aprendizagem. As gerações posteriores, com o contato com as diversas utilizações da linguagem e, ainda, com o convívio escolar com o padrão, tendem a incorporar, cada vez mais, traços da língua comum, perdendo vestígios da língua inicial.

Portanto, o fator faixa etária pode estar relacionado com outros aspectos sociais, tais como rede de relações sociais, mercado de trabalho e escolaridade, sendo assim, “[...] nem todo indício de mudança em curso apresentado pela distribuição dos resultados em função da faixa etária é reflexo somente da gradação etária dos falantes que constituem a amostra do estudo” (FREITAG, 2005, p.106).

5.2.1.2 Gênero *versus* Faixa etária

Para entendimento mais aprofundado das variáveis gênero e faixa etária, apresentamos na *Tabela 10* uma amostragem para observação da relação dessas duas variáveis.

Tabela 10: Taxas de uso da concordância no SN em função do gênero e da faixa etária no PB

GÊNERO	Faixa I		Faixa II		Faixa III	
	Frequência	%	Frequência	%	Frequência	%
Masculino	34/154	22	10/201	5	3/66	5
Feminino	113/256	56	32/146	22	28/110	25

Os resultados do cruzamento de dados ainda demonstram que as mulheres mais jovens tendem a seguir o padrão da marcação de plural nos SNs mais que os homens, como vimos no *subtitem 5.2.1.1*. Lopes (2001, p. 172) acredita, e comungamos da mesma idéia, que dados como esses talvez possam ser justificados “com o período em que há a relação mais forte com o ingresso no mercado trabalho, ou durante a sua fase mais produtiva no mercado de trabalho [...]. Essa pressão parece não mais atingi-las na última faixa etária”. Vejamos os dados expostos em gráfico:

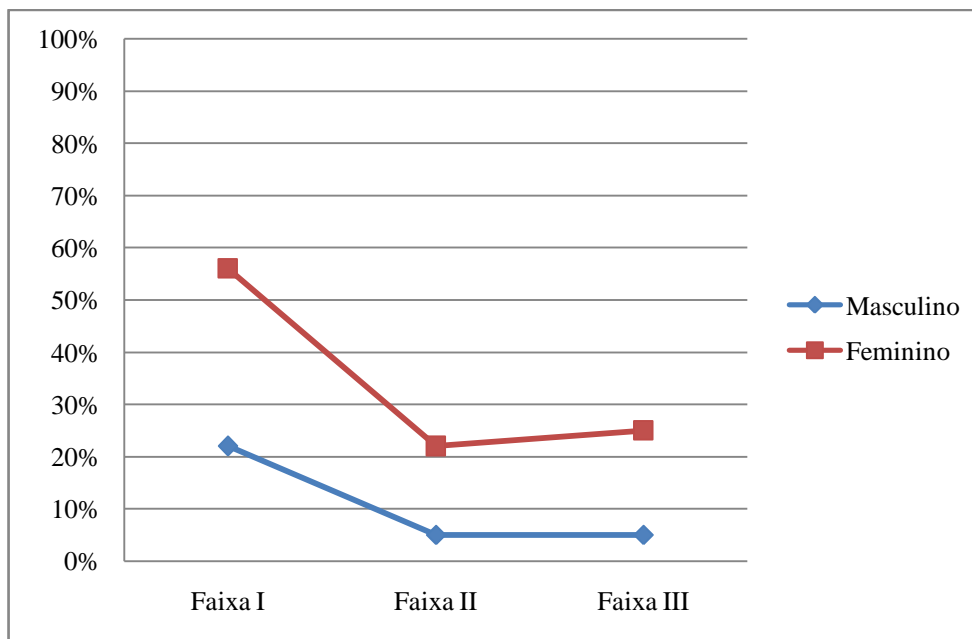


Figura 11: Efeito das variáveis Gênero e Faixa etária, associadas, na concordância

O gráfico demonstra que a distribuição etária, considerando todos os falantes, continua a se mostrar curvilínea, o que fica claro que independente de idade há uma escolha da não marcação do plural pelos falantes do sexo masculino e pela marcação de plural pelos falantes do sexo feminino.

5.2.2 Português angolano: variáveis sociais

Para a amostra do português angolano o programa computacional GOLDVARB X selecionou como relevantes, com significância 0.000, as variáveis gênero e origem, que nos propomos a apresentar, a seguir, seus respectivos resultados em porcentagens e pesos relativos.

5.2.2.1 Gênero

Como já discutido para o PB, do ponto de vista social, homens e mulheres têm papéis diferentes e estão expostos a situações diversas, o que pode influenciar a escolha de uma ou outra forma linguística. O efeito da variável gênero sobre a presença de marcas de plural nos SNs do PA pode ser observado, inicialmente, a partir da tabela e gráfico a seguir.

Tabela 11: Taxas de uso da concordância no SN em função do gênero no PA

GÊNERO	Frequência	%	Peso relativo
Masculino	95/274	34,7	.63
Feminino	69/346	19,9	.39

Os dados nos apresentam uma diferença entre os pesos relativos (de .24), havendo uma tendência no uso da concordância plural dos SNs entre os falantes do sexo masculino, peso relativo de .64, dialogando com os resultados de Helvécia (ANDRADE, 2003), em que há uma tendência de que os homens realizem mais a concordância do que a mulheres, mas contrariando os resultados obtidos através da nossa amostra do PB. Vejamos tabela comparativa:

Tabela 12: Dados comparativos: uso da concordância no SN em função do gênero

GÊNERO	ANDRADE (2003) <i>Helvécia - Ba</i>		<i>Mato Grosso</i>		<i>Luanda</i>	
	%	Peso relativo	%	Peso relativo	%	Peso relativo
Masculino	10	.63	11,2	.28	34,7	.63
Feminino	9	.39	39,6	.68	19,9	.39

Os resultados comparativos são surpreendentes, pois temos os falantes do sexo masculino liderando a marcação de plural nos SNs com pesos relativos idênticos aos de Helvécia: .63 para o dialeto de Luanda e .63 para o dialeto de Helvécia, diferentemente do que encontramos em Mato Grosso (peso relativo de .28). Com isso, podemos afirmar que no dialeto de Luanda, assim como no dialeto de Helvécia-Ba, os homens favorecem o uso da regra sintática da concordância de número, e as mulheres desfavorecem-na.

Abaixo, apresentamos em gráfico os resultados em percentual da variação em função do gênero no PA:

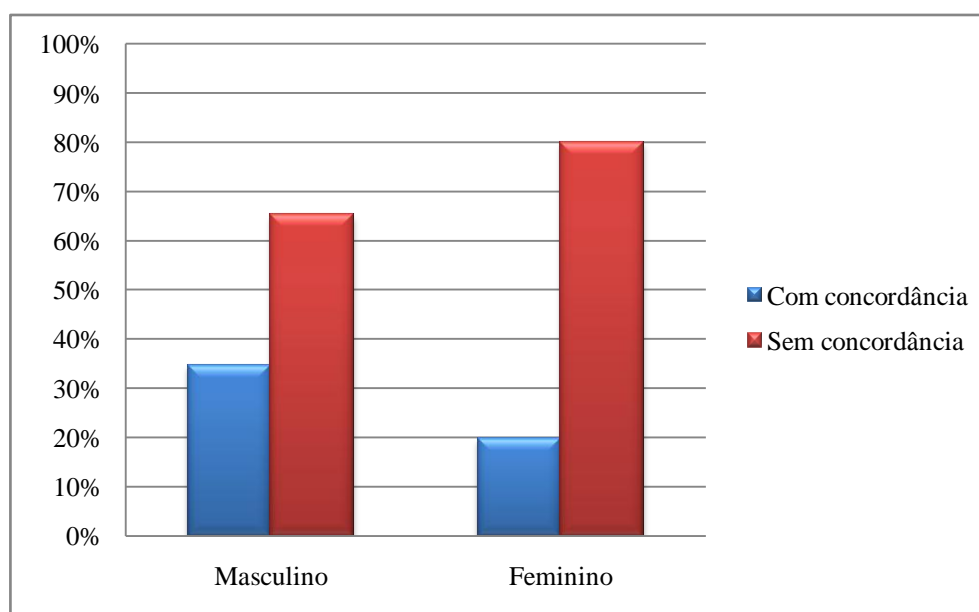


Figura 12: Efeito da variável gênero na variação na concordância de número nos SNs do PA

Na busca de novas respostas para o fato dos homens liderarem a variante prestigiada, encontramos uma explicação viável relacionada ao processo sócio-histórico de Angola,

relativamente ao nível de alfabetização e educação, a população angolana apresenta um índice muito elevado de analfabetismo (80,4%) e possui um nível de educação muito baixo [...]. As últimas estimativas disponíveis remontam a 1998, as quais apontam para apenas 42% da população com idade igual ou superior a 15 anos saber ler e escrever. *Esta taxa apresenta um grande diferencial entre ambos os sexos, em que cerca de 56% dos homens sabe ler e escrever e apenas 28% das mulheres detém essas mesmas características.* (BENTO, 2010, p. 05 – Grifo meu)

Mesmo nosso *corpus* sendo representativo do português popular de Luanda, em que os falantes são analfabetos ou estudaram até o 4º ano, é possível que esse acesso, ainda que mínimo, à escola seja um fator a ser considerado, principalmente, após o trabalho de Bento (2010) que afirma que 56% da população masculina têm e/ou tiveram acesso à leitura e à escrita e apenas 28% da população feminina tiveram acesso aos mesmos direitos. Portanto, o acesso ou não à escrita e à leitura, pode ser um fator a influenciar a aquisição das marcas de plural nos SNs.

Baxter (2009, p. 285), e outros pesquisadores (ANDRADE, 2003; LUCHESSI, 2000) acreditam que o fato dos homens liderarem a marcação de plural tem relação direta ao fator social, dessa forma

Os homens têm um maior grau de integração social, o que implica um maior contato com padrões linguísticos adventícios. Passam mais tempo fora da comunidade [...], enquanto as mulheres tendem a estar mais enraizadas na sua própria comunidade, dedicadas à roça e ao trabalho doméstico.

O fator escolaridade e o fator sócio-cultural são duas possíveis vertentes que conduzem a uma possível explicação do maior uso do morfema de plural pelos falantes do sexo masculino, no PA, já que ambos os fatores têm relação direta com o processo de contato com a língua alvo. Portanto, nossos resultados fornecem novos elementos para constataremos o importante papel dos condicionantes sociais nos rumos de um dialeto.

5.2.2.2 Origem

A confirmação de que uma maior integração social é um dos fortes condicionantes das mudanças/variações observadas no dialeto, segmentamos nosso *corpus* entre os falantes naturais da cidade de Luanda e os falantes vindos de outras regiões. Sendo este o segundo fator selecionado pelo programa computacional, com significância de 0.000. Vemos os resultados na tabela e gráfico que seguem.

Tabela 13: Taxas de uso da concordância no SN em função da origem no PA

ORIGEM	Frequência	%	Peso relativo
Luanda	107/236	45,3	.74
Outra	57/384	14,8	.34

Os dados revelaram que os falantes de origem luandense lideram a presença do morfema de plural nos SNs, com 45,3% e .74 de peso relativo, já os falantes de outras origens apresentaram apenas 14,8% dos seus SNs marcados, sendo atribuído, pelo programa, peso relativo de .34.

Os dados quantitativos deixaram em evidência um cenário que os falantes de origem luandense lideram a aquisição das marcas de plural, o acreditamos que advém do reflexo de uma cidade que é capital, considerada uma das mais desenvolvidas do país, em que a maioria dos habitantes são falantes nativos do português.

Como vimos em capítulos anteriores a língua portuguesa foi imposta aos angolanos, mas as muitas línguas africanas continuaram a servir como meio de comunicação em espaços e situações específicas, ao passo que a língua neolatina era utilizada de um modo

plurifuncional no país. Resta-nos compreender até que ponto as línguas africanas exercem ou podem exercer influência na marcação de plural nos SNs, uma variável a ser considerada.

Nesse sentido, Baxter (2009, p. 287), acredita que, considerando a presença histórica de grupos homogêneos de falantes de línguas africanas, parecem ser de interesse especial dois fatores: (i) português L2 de adultos, com transferências estruturais das L1 africanas; (ii) português L1 de afrodescendentes da mesma comunidade. Tal aspecto deve ser levado em consideração visto que

em sociedades multilíngues em que se tem desenvolvido uma segunda língua veicular e/ou oficial pós-colonial como L2 - por exemplo, [...] o português de Angola (INVERNO, 2005; MINGAS, 2000) e Moçambique (GONÇALVES, 1996, 2004) -, há amplas evidências da influência das línguas maternas originais dessas comunidades sobre as emergentes L2/L1. (BAXTER, 2009, p. 287)

5.3 VARIÁVEIS LINGUÍSTICAS

A fim de observar de que forma a variação no uso da regra de concordância se encaixa na estrutura linguística nos dialetos de Mato Grosso e Luanda, analisamos os seguintes grupos de fatores linguísticos, a citar: (i) *configuração sintagmática do SN*; (ii) *função sintática do SN*; (iii) *posição do SN em relação ao verbo*; (iv) *número absoluto de constituintes do SN*; (v) *número de constituintes flexionáveis do SN*; (vi) *saliência fônica dos elementos do SN*. Porém, desses, o programa selecionou como fator relevante ao uso da marca de plural no PB apenas (v) e (vi), com significância 0.009, e para o PA nenhum fator linguísticos foi selecionado.

Portanto será focalizado, a partir de então, as variáveis linguísticas que se tomam como possíveis grupos de fatores a interferir na variação da concordância no sintagma nominal.

5.3.1 Português brasileiro: variáveis linguísticas

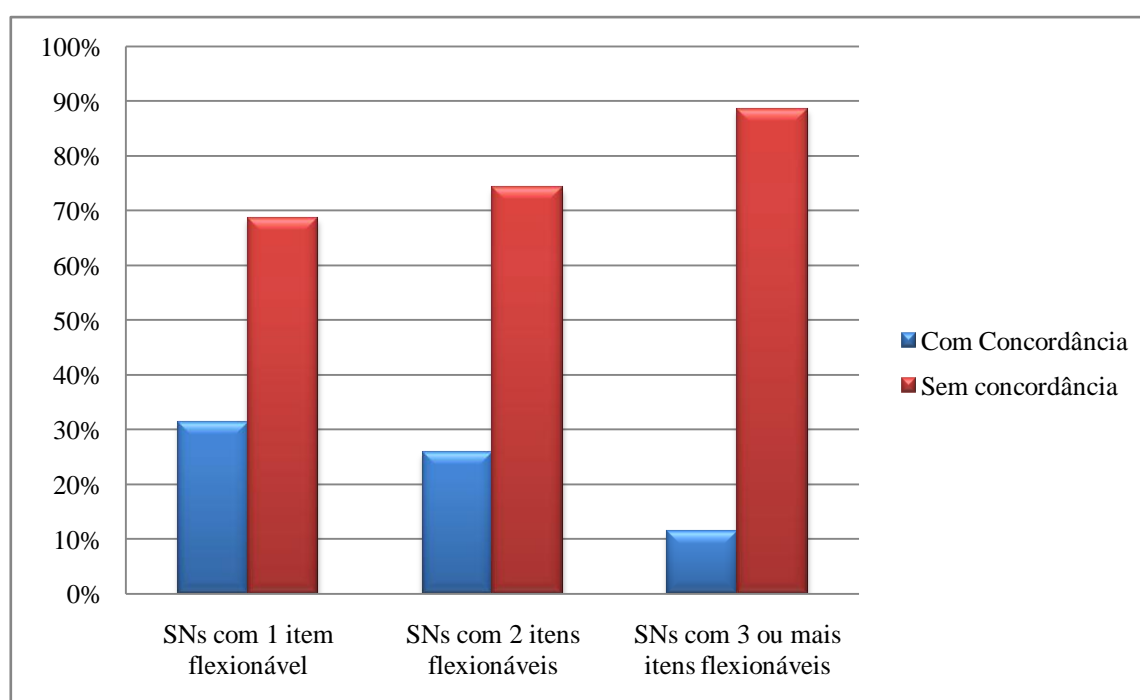
5.3.1.1 Número de constituintes flexionáveis do SN

Dentre as variáveis linguísticas o *número de constituintes flexionáveis do SN* foi selecionado em primeiro lugar, com significância 0,009. Vejamos os resultados através da tabela e gráfico.

Tabela 14: Taxas de uso da concordância no SN em função do número de constituintes flexionáveis no PB

Nº DE CONSTITUINTES FLEXIONÁVEIS DO SN	Frequência	%	Peso relativo
SN^s com 1 item flexionável <i>“Só quinze dias.”</i>	87/277	31,4	.61
SN^s com 2 itens flexionáveis <i>“Assisti muitos filmes.”</i>	158/612	25,8	.47
SN^s com 3 ou mais itens flexionáveis <i>“Esses lugares grandes, né?”</i>	5/44	11,4	.24

Os resultados indicam que os SNs que mais favorecem a aplicação da regra de concordância são aqueles constituídos por apenas um elemento flexionável, com uma frequência de 31% e peso relativo de .61. A frequência e o peso relativo do segundo fator, envolvendo SNs com dois constituintes pluralizáveis, sofre uma diminuição em termos percentuais, 25,8% e também há uma diminuição no peso relativo, que passa a ser de .47. O terceiro fator, SNs com três ou mais constituintes pluralizáveis, apresentou percentual de apenas 11,4% e peso relativos de .14. Vejamos as taxas de variação na concordância de número nos SNs diante desta variável.

**Figura 13:** Efeito da variável do número de constituintes flexionáveis na variação na concordância de número nos SNs no PB

Como é observado através do gráfico, número de elementos flexionáveis no SN é diretamente inversamente proporcional à presença de marcas de plural. Portanto, quanto maior o número de elementos flexionáveis no SN, a tendência é que o número de marcas de plural decresça. Como pode ser observado em tabela comparativa, nossos dados, mais uma vez, dialogam com os de Helvécia (ANDRADE, 2003).

Tabela 15: Dados comparativos: uso da concordância no SN em função do número de constituintes flexionáveis

Nº DE CONSTITUINTES FLEXIONÁVEIS DO SN	ANDRADE (2003) <i>Helvécia - Ba</i>		<i>Mato Grosso-Ba</i>	
	%	Peso relativo	%	Peso relativo
SN ^s com 1 item flexionável	15	.66	31,4	.61
SN ^s com 2 itens flexionáveis	7	.46	25,8	.47
SN ^s com 3 ou mais itens flexionáveis	5	.26	11,4	.24

Nossos dados comparados aos de Andrade (2003) confirmam que os sintagmas cuja demanda por flexão está em menor escala tende a ser mais marcados, já que em termos de processamento, são mais simples.

5.3.1.2 Saliência fônica

O fator saliência fônica foi introduzido nos estudos do português por Naro e Lemle (1976), inicialmente na concordância verbal, e posteriormente incorporada aos estudos de concordância verbal e nominal de Braga e Scherre (1976), de Braga (1977), de Scherre (1978), de Naro (1981), assim como de Guy (1981), de Scherre (1988) (*cf.* LOPES, 2001, p. 194) e mais recentemente destacamos os trabalhos de Lopes (2001) e Andrade (2003). Esses autores acreditam que a quantidade de material existente na oposição entre singular e plural interfere na possibilidade de se fazer a concordância, chegando à conclusão que os itens em que a oposição é mais saliente (mais material fônico perceptível na oposição singular/plural) devem ser os mais marcados com o morfema de plural (LOPES, 2001, p. 194).

A influência da saliência fônica na presença/ausência de plural nos elementos do SN é medida a partir de uma escala, considerando-se o nível de diferenciação material fônica na oposição singular/plural. Em língua portuguesa, essa oposição pode se dar

mediante o acréscimo do morfema de plural em palavras terminadas por vogal oral (*casa/casas*) ou vogal nasal (*algum/alguns; maçã/maçãs*); através da substituição da consoante final por um conjunto formado de vogal mais consoante (*papel/papéis*); pelo acréscimo da terminação *-es* (*mar/mares; mês/meses*); havendo ditongo nasal a oposição pode ser feita apenas com o acréscimo do morfema de plural (*cidadão/cidadãos; homem/homens*) ou com alterações no grupo vocálico (*pão/pães*); e ainda com mudanças metafônicas (*olho/ólhos*), o denominado plural duplo. (ANDRADE, 2003, p. 101)

Dessa forma, iniciamos com a hipótese que o item lexical não regular tem mais chances de ser marcado, portanto um SN que contenha pelo menos um constituinte não regular apresenta mais chances de conter todas as marcas formais de plural. Vejamos nossos resultados.

Tabela 16: Taxas de uso da concordância no SN em função da saliência fônica no PB

SALIÊNCIA FÔNICA	Frequência	%	Peso relativo
SN^s apenas com itens regulares “Assisti <u>muitos</u> filmes.”	166/679	24,4	.46
SN^s contendo itens que apresentam diferenciação fônica quando pluralizados “Ela tinha <u>oito meses</u> .”	84/254	33,1	.58

Nos dados, apareceram 679 SNs com plural regular, em que se insere apenas o morfema *-s*, dos quais 24,2% apresentaram concordância de plural, obtendo com peso relativo de .46. Ao observar os itens que se constituem como salientes foneticamente, temos 254 SNs em que 33,1% apresentam todas as marcas de plural, com peso relativo de .58. Assim, podemos afirmar que no *corpus* os itens mais salientes favorecem mais a presença de marca explícita nos elementos nominais dos SNs e os menos salientes, os regulares, favorecem menos a presença de marcas explícitas (SCHERRE; NARO, 1998, p.7).

Considerando as *Tabelas 16 e 17*, notamos que o quadro de variação na concordância de número no SN no dialeto de Mato Grosso está em conformidade com o que se tem postulado sobre a aquisição da regra a partir de contextos linguísticos mais salientes, demonstrando que formas nominais que apresentam maior diferenciação morfofonológica na oposição singular-plural, a frequência de aplicação da regra é maior do que com formas nominais que não apresentam essa diferenciação morfofonológica para marcar o plural.

Assim como as variáveis sociais, a variável linguística saliência fônica, encontra precedentes nos resultados de Helvécia (ANDRADE, 2003). Vejamos tabela comparativa:

Tabela 17: Dados comparativos: uso da concordância no SN em função da saliência fônica no PB

SALIÊNCIA FÔNICA	ANDRADE (2003) <i>Helvécia - Ba</i>		<i>Mato Grosso-Ba</i>	
	%	Peso relativo	%	Peso relativo
SN ^s apenas com itens regulares	8	.46	24,4	.46
SN ^s contendo itens que apresentam diferenciação fônica quando pluralizados	15	.66	33,1	.58

É possível observamos que encontramos o mesmo peso relativo, de .46, para os itens regulares, enquanto que para os itens salientes temos .58 para amostra de Mato Grosso e .66 para amostra de Helvécia. Ambos *corpora* indicam, como esperado, que o princípio da saliência fônica é relevante na aplicação da regra de concordância. Como afirma Andrade (2003, p. 114), outras pesquisas, que analisaram essa variável, apontam que os valores estatísticos sempre apontam para a força condicionadora das formas que se apresentam mais perceptíveis ao falante.

Ao tratar das questões implícitas à saliência fônica Lopes (2001) retoma alguns pressupostos de Guy (1981, p.296), o qual

ao estudar a saliência fônica na concordância na fala carioca, estabeleceu uma relação entre a quantidade de material fônico e a percepção do ouvinte, no momento da aquisição da regra (numa situação de suposta descrioulização), daí considerar os mais facilmente percebidos, ou mais salientes, como os inicialmente aprendidos ou incorporados. As formas em que a oposição singular/plural é mais saliente são, por isso, mais marcadas com o morfema.

Eis o ponto que entra em cena duas situações/opiniões contrárias: Naro (1981) acredita que é a partir dos itens com concordância menos saliente que a variação se inicia; e Guy (1981) acredita que é a partir dos itens com concordância mais saliente que se dá a percepção, na fase de aquisição (LOPES, p. 195). Nesse sentido Andrade (2003, p. 115) acredita que

há toda uma rede de elementos internos e externos envolvida nos processos de mudança, e um desses elementos é a força das formas lingüísticas de maior prestígio na sociedade, portanto, é plenamente razoável que o falante adquira formas da língua culta a partir de contextos lingüísticos mais salientes, por outro lado, não parece concebível a idéia de que o falante possa vir a substituir formas cultas por não cultas em contextos lingüísticos mais perceptíveis. Mas, ao entrever que as formas mais salientes são as que mais favorecem a descrioulização, Guy, de fato, não poderia concordar com a teoria de Naro e Lemle, pois, se há séculos a língua caminha no sentido de perder as marcas de concordância, como determina a sua deriva interna, não seria possível explicar porque os dialetos populares estão em processo de aquisição dessas marcas.

Portanto, nossos dados apresentam indícios estatísticos de que a variação na concordância de número no SN no dialeto de Mato Grosso dialoga com outros estudos sobre a aquisição da regra a partir de contextos linguísticos mais salientes, além de apresentar as mesmas tendências de outros dialetos populares, a citar o de Helvécia, no sentido de perder as antigas características provenientes de processos de transmissão linguística irregular.

CONSIDERAÇÕES FINAIS

Nesta dissertação, analisamos a concordância nominal no panorama da variedade do Português Brasileiro (PB) em comparação a variedade do Português Angolano (PA), buscando identificar convergências/divergências entre essas duas variedades do português, tendo como base de análise os padrões canônicos estabelecidos pelo Português Europeu (PE). Para tanto, buscamos explicações que pudessem colaborar com a hipótese de que variação existente no sistema de concordância nas normas populares do PB e PA seria decorrente da sócio-história, marcada pelo contato entre línguas, com processos de transmissão linguística irregular.

Certamente o processo de variação da concordância nominal de número é familiar no âmbito das discussões sobre as origens do português brasileiro, mas dessa vez propomos não apenas analisar o fenômeno no âmbito do PB, mas buscar suas similaridades ao comparar com o PA, também fruto do processo de contato entre linguístico.

Iniciamos pelas análises, fugindo do que havíamos proposto no capítulo metodológico, na perspectiva qualitativa, demonstrando casos específicos encontrados no levantamento de dados. Dentre eles destacamos: (i) os sintagmas em que apenas o primeiro elemento não recebe marca de plural, contrariando o que até então tínhamos visto em Scherre (1988); (ii) os sintagmas nominais em que há no constituinte núcleo o acréscimo do segmento fonético [i]; (iii) os sintagmas nominais em que temos um determinante e um nome no diminutivo, favorecendo a não marcação de plural. Os questionamentos suscitados diante dos dados representativos de (i), (ii) e (iii) não estão diretamente ligados a possíveis hipóteses, mas apresentam evidências de que as variedades do português, no nosso caso específico PB e PA, são fontes inesgotáveis de pesquisa. Por tanto, resta-nos, a partir de agora, buscar mais dados para confirmar se há uma tendência ou se esses são apenas dados isolados do português e que fatores motivariam a ausência de marca de plural nas ocorrências supracitada. Certamente, a comparação entre PB e PA constitui uma perspectiva de investigação que se desdobra em novas reflexões diante de fenômenos linguisticamente familiares, assim como a variação na concordância de número.

Os dados quantitativos, apontados diante das variáveis sociais e linguísticas, revelaram que o processo de variação e mudança no dialeto afro-brasileiro segue o mesmo padrão observado em Luanda, amostra do PA, observando-se apenas uma diferença no grau de intensidade das mudanças e variações em cada caso, como foi observado através das tabelas e gráficos estatísticos.

No que tange a variável social gênero, selecionada para o PB com significância 0.009 e para o PA com 0.000, é perceptível uma tendência ao uso da concordância plural dos SNs entre as mulheres para o PB, com .68, e entre os homens para o PA, com .64. Nos dados do PB temos uma tendência para os falantes do sexo feminino liderarem a norma gramaticalmente estabelecida; já os dados do PA justifica-se através do fato de que os homens têm um maior grau de integração social, o que implica um maior contato com padrões linguísticos normatizadores.

Ao tratar do fator social faixa etária, no PB, foi observado que a faixa mais nova é a única que favorece o uso da concordância, com .70 de peso relativo, sendo a faixa que tende a marcar mais os elementos do sintagma nominal com o morfema de plural, demonstrando evidência de um processo de mudança em curso, com a aquisição das regras de concordância nominal pelo grupo de falantes mais jovens. Cruzando os dois fatores sociais, gênero e faixa etária, ratificamos os resultados preliminares, demonstrando que independente de idade há uma escolha da não marcação do plural pelos falantes do sexo masculino e pela marcação de plural pelos falantes do sexo feminino, cabendo a faixa etária I (até 30 anos) a liderança desse processo de mudança.

Ainda para o PB foram selecionados duas variáveis linguística, *número de constituintes flexionáveis do SN* e *saliência fônica dos elementos do SN*. Através dos dados estatísticos da primeira variável linguística, podemos ratificar nossa hipótese inicial de que os SNs que mais favorecem a aplicação da regra de concordância são aqueles constituídos por apenas um elemento flexionável, em que obtivemos peso relativo de .61. Nesse sentido nossos dados afirmam que o número de elementos flexionáveis no SN está proporcional à presença de marcas de plural, portanto, quanto maior o número de elementos flexionáveis no SN, a tendência é que o número de marcas de plural decresça.

Em relação a saliência fônica, o segundo fator linguístico selecionado pelo programa GOLDVARB X, observamos e ratificamos nossa hipótese inicial de que os itens mais salientes favorecem mais a presença de marca explícita nos SNs, enquanto os menos salientes, os regulares, favorecem menos a presença de marcas explícitas.

Voltando aos dados do PA, ressaltamos que apenas duas variáveis sociais foram selecionadas pelo programa, gênero e origem, a primeira já discutida nestas considerações finais. A variável origem, com fatores que determinavam se os falantes eram nativos de Luanda ou vindo de outras regiões, revelou que os falantes de origem luandense lideram a presença do morfema de plural nos SNs, com 44,1% e .75 de peso relativo, já que essa é a

cidade mais desenvolvida do país, em que a maioria dos habitantes são falantes nativos do português.

O fato do programa estatístico selecionar como relevante duas variáveis sociais e duas linguística para o PB e apenas duas sociais para o PA, já nos traz indícios da importância dos fatores externos a análises de fenômenos variáveis, principalmente nos casos em que uma das variantes é estigmatizada, no caso ausência das marcas plural em alguns elementos do sintagma nominal.

Nos debates em função da existência da variação da concordância no sintagma nominal no PB, esta pesquisa procurou buscar meios de contribuir traçando análises comparativas com o PA. Entre as duas posições básicas assumidas pelos linguistas para as origens do português, esta dissertação colaborou, ainda que timidamente, para os debates em torno da tão conturbada polêmica, envolvendo a concordância, entre a deriva e o contato linguístico.

Sempre que possível, apresentamos as convergências/divergências entre as variedades do português, PB e PA, verificando em que medida se aproximam ou se distanciam no que tange ao processo de variação na concordância de número nos sintagmas nominais, entendendo tais mudanças/variações dentro do processo de formação sócio-histórica dos dois países, marcados pelo contato entre línguas.

REFERÊNCIAS

- ALBUQUERQUE, W. R. de; FRAGA FILHO, W. *Uma história do negro no Brasil*. Salvador: Centro de Estudos Afro-Orientais; Brasília: Fundação Cultural Palmares, 2006.
- ALKMIN, T. Sociolinguística. In: MUSSALIM, F.; BENTES, A. C. (orgs.). *Introdução à linguística: domínios e fronteiras*. São Paulo: Cortez, 2001. p. 21-47.
- ALMEIDA, N. L. F. de; CARNEIRO, Z. de O. N. *Coleção amostra da língua falada no semiárido baiano*. Feira de Santana: Universidade Estadual de Feira de Santana, 2008.
- ANDRADE, P. R. de. *Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialeto afro-brasileiro*. Dissertação de Mestrado em Língua Portuguesa: UFBA, Salvador, 2003.
- BAXTER, A. N. Concordância nominal. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). *O Português Afro-Brasileiro*. 1. ed., v. 1., Salvador: EDUFBA, 2009.
- _____. The development of variable NP plural agreement in a restructured african variety of portuguese. In: GENEVIÈVE, E.; ARMIN, S. (eds.). *Creoles, contact, and language change: linguistic and social implications*, vol. 27. Amsterdam: John Benjamins, 2004. p. 97-126
- _____. *The development of variable number agreement in a restructured African variety of Portuguese*. Texto apresentado no Colóquio sobre Línguas Crioulas. Coimbra: Universidade de Coimbra, 2001.
- _____; LUCCHESI, D. A relevância dos processos de pidginização e criolização na formação da língua portuguesa no Brasil, *Estudos Linguísticos e Literários*, 1997. p. 65-84.
- _____; _____. GUIMARÃES, M. “Gender agreement as a decreolizing feature of the Afro-Brazilian rural dialect of Helvécia”, *Journal of Pidgin and Creole Language*, 1997.
- BECHARA, E.. *Moderna gramática portuguesa*. 37. Ed. ver. e ampl. 16. Reimpr. Rio de Janeiro: Nova Fronteira, 2009.
- BENTO, C. T. J. S. F. *Contributos para uma caracterização linguística do Luandense*. Dissertação de Mestrado. Universidade de Aveiro. 2010.
- BRAGA, M. L. *A concordância de número no sintagma nominal no triângulo mineiro*. Dissertação de Mestrado. Pontifícia Universidade Católica. Rio de Janeiro, 1977.
- _____; SCHERRE, M. M. P.. *A concordância de número no SN na área urbana do Rio de Janeiro*. Anais do Encontro Nacional de Linguística. Rio de Janeiro: PUC, 1976. pp. 446-477
- BRANDÃO, S. F. Concordância nominal em duas variedades do português: convergências e divergências. *Revista em Estudos Linguísticos Veredas*. Atemática 1/2011. p. 164-178

_____; VIEIRA, S. R.. A concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística. *Revista PAPIA*. 22(1), 2012. p. 7-39.

BONVINI, E. Línguas africanas e português falado no Brasil. In: PETTER, M.; FIORIN, J. L. *África no Brasil: a formação da língua portuguesa*. 1. ed., 2ª reimpressão. São Paulo: Contexto, 2009.

_____; PETTER, M. M. T. *Portugais du Brésil et langues africaines*. *Langages*. Paris, n. 130, 1998. p. 68-83.

BORTONI-RICARDO, S. M. *Do campo para a cidade: estudo sociolinguístico de migração e redes sociais*. São Paulo : Parábola Editorial, 2011.

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

CÂMARA Jr., J. M. *Estrutura da Língua Portuguesa*. 13ª ed. Petrópolis: Editora Vozes, 2000.

_____. Ele como um acusativo no português do Brasil. In: _____. *Dispersos*. Rio de Janeiro: FGV, 1972.

CAMERON, R. Ambifuous agreement, functional compensation, and non epecifictu in Spanish of San Juan, Poerto Rico, and Madri, Spain. *Language variation and change*, v. 5., n. 3, 1993. p. 35-49.

CARNEIRO, Z. de O. N.; ALMEIDA, N. L. F. de. Elementos para uma sócio-história do semi-árido baiano. In: RAMOS, J.M.; ALKMIM, M. A.. (Org.). *Para a História do Português Brasileiro: estudos sobre mudança linguística e história social*. 1. ed. Belo Horizonte: Faculdade de Letras da UFMG, 2007, v. 1, p. 423-442.

_____. A criação de escolas a partir de critérios demográficos na Bahia do século XIX: uma viagem ao interior. In: Lobo, TÂNIA; Ribeiro, ILZA; CARNEIRO, Zenaide; ALMEIDA, Norma. (Org.). *Para a História do português brasileiro: novos dados, novas análises*. Salvador: Edufba, 2006. v. 2, p. 649-674.

CASTRO, I. *Curso de História da língua portuguesa*. Lisboa: Universidade Aberta, 1991.

CASTRO, Y. P. de. *Falares africanos na Bahia: um vocabulário afro-brasileiro*. 2. ed. Rio de Janeiro: ABL: Topbooks, 2005. 366 p.

_____. *Falares africanos na interação social do Brasil colônia*. Rio de Janeiro: Academia Brasileira de Letras/Topbook, 2001.

COELHO, F. A. (1880-1886). Os dialetos românicos ou neolatinos na África, Ásia e América. In: BARBOSA, J. M. (org.) *Estudos Linguísticos Crioulos*. Lisboa: Academia Internacional de Cultura Portuguesa, 1967.

COMPANHIA, Carlito. A. *Uso da morfologia flexional de número em sintagmas nominais plurais na interlíngua de crianças da 3ª classe*. Maputo: Universidade Eduardo Mondlane. Tese de doutorado, 2001.

CARVALHO, R. C.. *A concordância de número no sintagma nominal na fala urbana de Rio Branco*. 1997. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem - UNICAMP, Campinas, 1997.

DUBOIS, J. et. al. *Dicionário de linguística*. Trad.: Izidoro Blikstein et all. São Paulo: Cultrix, 1973.

Ethnologue: languages of the world. Disponível em: <<https://www.ethnologue.com/>>. Acessado em: 23 de março de 2014.

FAGE, J. D. *História da África*. Lisboa: Edições 70, 2010.

FARACO, C. A.. *Língua portuguesa: um breve olhar sobre sua história. Salto para o futuro*. Português: um nome, muitas línguas. Ano XVIII boletim 08 - Maio de 2008.

_____. *Linguística Histórica: uma introdução ao estudo da história das línguas*. São Paulo: Editora Parábola, 2005.

FAUSTO, B. *História do Brasil*. Edusp. 1996. Disponível em: <http://www.caccto.com.br/material/d00044/Material_6_E-MED_2A_195641.pdf>. Acessado em: 22 de março de 2013.

FIGUEIREDO, Carlos Filipe Guimarães. A concordância variável no sintagma nominal plural do português reestruturado de Almojarife. *PAPIA* 18. 23-43, 2008.

FIORIN, J. L.; PETTER, M. M. T. *África no Brasil: a formação da língua portuguesa*. 1. ed. 2. reimpressão. São Paulo: Contexto, 2009.

FREITAG, R. M. K. Idade: uma variável sociolingüística complexa. *Revista Línguas e Letras*. v. 6, n. 2, p.105-121, jan./jun., 2005. Disponível em: <<http://e-revista.unioeste.br/index.php/linguaseletras/article/viewArticle/875>>. Acesso em: 31 jan. 2010.

GONÇALVES, P. (*Dados para a história da língua portuguesa em Moçambique*). 2000. Disponível em: <<http://cvc.institutocamoes.pt/hlp/geografia/portuguesmocambique.pdf>>. Acesso em: 05 set. 2012.

_____. *Português de Moçambique: Uma Variedade em Formação*. Maputo: Livraria Universitária, 1997.

GUY, G. R.. Introdução à análise quantitativa da variação linguística. In: _____. *Sociolingüística quantitativa: instrumental de análise*. São Paulo: Parábola, 2007. 239 p.

_____. A questão da criouliização no português do Brasil. In: ZILLES, A. M. S. (Org.). *Estudos de variação linguística no Brasil e no Cone Sul*. Porto Alegre: Editora da UFRGS, 2005.

_____. *Linguistic Variation in Brazilian Portuguese: Aspects of Phonology, Syntax and Language History*. University of Pennsylvania, PhD Dissertation. Ann Arbor: University Microfilms International, 1981.

_____; ZILLES, A. M. S. *Sociolingüística quantitativa: instrumental de análise*. São Paulo: Parábola, 2007.

HOLM, J. Popular Brazilian Portuguese: a semi-creole. In: D'ANDRADE, E.; KIHN, A. (Org.). *Actas do colóquio sobre crioulos de base lexical portuguesa*. Lisboa: Edições Colibri, 1992. pp. 37-66.

_____. Creole influence on popular Brazilian Portuguese. In: GILBERT, G. (Org.). *Pidgin and creole languages: essays in memory of John E. Reineck*. Honolulu: University of Hawaii Press, 1987. pp 406-429.

HOUAISS, A.. *O português no Brasil*. Rio de Janeiro: Unibrade, 1985.

Instituto Brasileiro De Geografia E Estatística. Disponível em:
<<http://www.ibge.gov.br/home/>>. Acessado em 12 de dezembro de 2013.

INVERNO, L. A transição de Angola para o português vernáculo: estudo morfossintático do sintagma nominal. In: CARVALHO, Ana Maria. *Português em Contacto*. Madrid-Frankfurt: Vervuer, 2009.

_____. *Angola's transition to vernacular portuguese*. Coimbra: Universidade de Coimbra, Faculdade de Letras, Dissertação de Mestrado, 2005.

_____. Português vernáculo do Brasil e português vernáculo de Angola: Reestruturação parcial vs. mudança linguística. In: FERNÁNDEZ-FERREIRO, M.; VEIGA, N. V. *Los criollos de base ibérica: ACBLPE*. Madrid: Iberoamericana, 2004.

JEROSLOW, E. H. M. *Rural Cearense Portuguese: a study of one variety of nonstandard Brazilian speech*. (Dissertation). Cornell University, 1974.

JON-AND, A. Concordância variável de número no SN do português L2 de Moçambique: algumas explicações sociais e linguísticas. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola* 2. 2010. p. 28-50.

_____. Concordância de número no SN do português L2 falado em Cabo Verde. *Congresso da Associação de Crioulos de Base Portuguesa e Espanhola (ACBLPE)*. Colônia, Alemanha, 11-15 de agosto de 2009.

KIPARSKY, P. Explanation in Phonology. In: PETERS, S. *Goals of linguistic theory*. New Jersey: Prentice Hall, 1972. p.189-225.

LABOV, W. *Padrões sociolinguísticos*. Tradução Marcos Bagno, Maria Marta P. Scherre, Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008.

_____. *Modelos sociolinguísticos*. Madri: Cátedra, 1983.

LEMLE, M.; NARO, A. J. *Competências básicas do português*. Relatório final de pesquisa apresentado às instituições patrocinadoras da fundação Movimento Brasileiro de Alfabetização (MOBRAL) e Fundação Ford. Rio de Janeiro: Fundação Movimento Brasileiro de Alfabetização, 1977.

LOPES, N. da S. *Concordância nominal, contexto linguístico e sociedade*. Tese de Doutorado. Universidade Federal da Bahia - Instituto de Letras, Salvador. 2001.

LOURENÇO, Eduardo. *A nau de Ícaro e imagem e miragem da lusofonia*. São Paulo: Companhia das Letras, 2001.

LUCCHESI, D. A deriva secular na formação do português brasileiro: uma visão crítica. In: LOBO, T. et all. (org.). *ROSAE: Linguística histórica, história das línguas e outras histórias*. Salvador: EDUFBA, 2012.

_____. História do contato entre línguas no Brasil. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I.(Orgs.). *O Português Afro-Brasileiro*. 1. ed., v. 1., Salvador: EDUFBA, 2009. 576 p.

_____. Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão de conjunto. In: RONCARATI, C.; ABRAÇADO, J. (org.) *Português Brasileiro II: contato linguístico, heterogeneidade e história*, 366-390. Niterói: EDUFF, 2008.

_____. *Sistema, mudança e linguagem: um percurso na história da linguística moderna*. São Paulo: Parábola Editorial, 2004.

_____. 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, 2003. p. 273 – 284.

_____. As duas grandes vertentes da história sociolingüística do Brasil (1500-2000). *Revista D.E.L.T.A.* 17:1, 2001.

_____. *A variação na concordância de gênero em uma comunidade de fala afro-brasileira: novos elementos sobre a formação do português popular do Brasil*. Rio de Janeiro: UFRJ, Faculdade de Letras. Tese de Doutorado em Linguística: 2000.

_____. A constituição histórica do português brasileiro como um processo bipolarizador: tendências atuais de mudança nas normas culta e popular. In: GROBE, S.; ZIMMERMANN, K. (org.). *'Substandard' e mudança no português do Brasil*. Frankfurt: TFM, 1998.

_____; BAXTER, A. A transmissão linguística irregular. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). *O Português Afro-Brasileiro*. 1. ed., v. 1., Salvador: EDUFBA, 2009. 576 p.

_____; _____; RIBEIRO, I.. *O Português Afro-Brasileiro*. Salvador: UFBA, 2009.

MACHADO FILHO, Américo V. L. *A questão da constituição histórica do português brasileiro: revendo razões*. In: Cláudia Roncarati, Jussara Abraçado. (Org.). *Português brasileiro II*. Niterói: EdUFF, 2008, v. , p. 347-365.

MARQUES, I. G. Algumas considerações sobre a problemática linguística em Angola. In: Instituto de Cultura e Língua Portuguesas (ed.). *Actas do congresso sobre a situação actual da língua portuguesa no mundo*. Lisboa: Instituto de Cultura e Língua Portuguesas, 1985.

MASSINI-CAGLIARI, G. *A música da fala dos trovadores: estudos de prosódia do português arcaico, a partir das cantigas profanas e religiosas*. Tese (Livre Docência em Fonologia) – Departamento de Linguística/UNESP, Araraquara, 2005.

MATTHEWS, P.H. *The Concise Oxford Dictionary of Linguistics*. New York: Oxford University Press, 1997. p. 264.

MATTOS E SILVA, R. V. *Ensaio para uma sócio-história do Português Brasileiro*. São Paulo: Parábola Editorial, 2004.

_____. Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa. In: ALKMIM, T. M. *Para a história do português brasileiro: novos estudos*. São Paulo: Humanitas/FFCHL/USP: FAPESP, v. 2, 2001a.

_____. (org.). *Para a história do português brasileiro*. Vol. II: primeiros estudos, tomo II. São Paulo: Humanista/FFLCH/USP: FAPESP, 2001b.

MELLO, H. R. Português padrão, português não-padrão e a hipótese do contato linguístico. In: ALKMIM, T. M. (Org.). *Para a história do português brasileiro: Novos estudos*. v. III. São Paulo: Humanitas, 2002. p. 341-358.

MELLO E SILVA, S. C. B. de. *Urbanização e metropolização no estado da Bahia: evolução e dinâmica*. Salvador: Centro Editorial e Didático da UFBA, 1989.

MENDONÇA, R. *A influência africana no português do Brasil*. Brasília: FUNAG, 2012.

MINGAS, A. A. *Interferência do Kimbundo no Português falado em Luanda*. Luanda: Caxinde, 2000.

MOLLICA, C.a; BRAGA, M. L.. *Introdução à sociolinguística: o tratamento da variação*. 3. ed. São Paulo: Contexto, 2008.

MUSSA, A. *O papel das línguas africanas na história do português no Brasil*. Dissertação de mestrado. Rio de Janeiro: UFRJ, 1991, mimeo.

MYERS-SCOTTON, C.; JAKE, J. L. Testing the 4M model: an introduction. In: TRAVIS, C. *International journal of bilingualism*, 2000, v. 4, p. 1-8.

NARO, A. J. O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In: RONCARATI, C.; ABRAÇADO, J. *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7Letras, 2003. p. 285-302.

_____.; SCHERRE, M. M. P.. *Origens do português brasileiro*. São Paulo: Parábola Editorial, 2007.

NEGRÃO, E. V.; VIOTTI, E. Em busca de uma história linguística. *Rev. Est. Ling. Belo Horizonte*, v. 20, n. 2, jul./dez. 2012. p. 309-342.

NOLL, V. Os primeiros empréstimos Tupis no português do Brasil. In: NOLL, V.; DIETRICH, W. *O português e o tupi no Brasil*. São Paulo: Contexto, 2010.

OLIVEIRA, J. M. *O futuro da língua portuguesa ontem e hoje: variação e mudança*. 2006. 254f. Tese de doutorado. Faculdade de Letras, Universidade do Rio de Janeiro, Rio de Janeiro.

OLIVEIRA, K.; SOLEDADE, J.; SANTOS, V. de S. Concordância nominal. In: *África à vista: dez estudos sobre o português escrito por africanos no Brasil do século XIX*. Salvador: Edufba, 2009. pp. 255-316.

PAGOTTO, E.; DUARTE, M. E. L. Gênero e norma: avós e netos, classes e clíticos no final do século XIX. In: LOPES, Célia Regina dos Santos (org.). *A norma brasileira em construção*. Rio de Janeiro: UFRJ/ FAPERJ, 2005, p. 67-81.

PAREDES, S. V. L. Subject omission and functional compensation: evidence from written Brazillian Portuguese. *Language variation and change*, v. 5., n. 1, 1993. p. 35-49.

PARKVALL, Mikael. *Da África para o Atlântico*. Tradução Rodolfo Ilari. Rio de Janeiro: Editora Unicamp, 2012.

PERINI, M, A.. *Sofrendo a gramática: ensaios sobre a linguagem*. 3. ed. São Paulo: Ática, 2003.

PETTER, M. M. T. Aspectos morfossintáticos comuns ao português angolano, brasileiro e moçambicano. *PAPIA* 19, p. 201-220, 2009.

_____. A presença de línguas africanas no português do Brasil. *Estudos Linguísticos*. São Paulo: São José do Rio Preto, v. XXVII, 1998. p. 777-783.

POPLACK, S. The notion of the plural in puerto rican spanish: competing constraints on (s) deletion. In: LABOV, W. (edit). *Locating language in time and space*. New York: Academic Press, 1980.

_____. Mortal phonemes as plural morphemes. In: SANKOFF, D.; CEDERGREN, H. (eds.). *Variation Omnibus*. Canada: Linguistic Research Inc., 1981. p.59-71.

RODRIGUES, A. *Línguas brasileiras*. São Paulo: Layola, 1985.

RODRIGUES, J. H. A vitória da língua portuguesa no Brasil colonial. *Humanidades*. vol 1, n.4. Brasília, 1983.

RODRIGUES, U. R. de S. Em busca da sílaba ótima: processos de acréscimo no caboverdiano. *Linguagem: Estudos e Pesquisas*, Catalão, vol. 14, n. 2. 2010.

SANKOFF, D.; TAGLIAMONTE S.; SMITH E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

SAPIR, E. *Language: an introduction to the study of speech*. San Diego: HJB Books, [1921] 1949.

SCHERRE, M. M. P. *Concordância nominal e funcionalismo*. Alfa, São Paulo, 41(n.esp), 1997. pp.181-206.

_____. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: SILVA, G. M. de O.; SCHERRE, M. M. P. (Orgs.). *Padrões sociolinguísticos - análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Editora da UFRJ, 1996. p.41-62

_____. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa (RILP) - Norma e Variação do Português*. Associação das Universidades de Língua Portuguesa. 12: 37-49. dez. de 1994.

_____. *Reanálise da concordância nominal em português*. Tese de Doutorado em Linguística. Rio de Janeiro, Faculdade de Letras, UFRJ, 1988.

_____. *A regra de concordância de número no sintagma nominal em português*. Dissertação de Mestrado, Pontifícia Universidade Católica-Rio, 1978.

_____; NARO, A. J.. Duas dimensões do paralelismo verbal no português popular do Brasil. *DELTA* 9(1):1-14, 1993.

_____. Sobre a concordância de número no português falado do Brasil. In RUFFINO, G. (org.) *Dialettologia, geolinguística, sociolinguística*. (Atti del XXI Congresso Internazionale di Linguística e Filologia Romanza) Centro di Studi Filologici e Linguistici Siciliani, Università di Palermo. Tübingen: Max Niemeyer Verlag, 5:509- 523, 1998.

_____. Mudança sem mudança: a concordância de número no português brasileiro. *Scripta*, Belo Horizonte, v. 9, n. 18, p. 107-129, 1º sem. 2006.

SILVA NETO, S. da. *Introdução ao estudo da língua portuguesa no Brasil*. 3. ed. Rio de Janeiro: Presença, 1986 [1950].

STROUD, C.; PERPÉTTUA, G.. (eds.). *Panorama do português oral de Maputo*. Maputo: Instituto Nacional do Desenvolvimento da Educação, 1997.

Superintendência do Desenvolvimento do Nordeste (SUDENE). Disponível em: <<http://www.sudene.gov.br/>>. Acessado em: 12 de dezembro de 2013.

SUSSURE, F. *Curso de Linguística Geral*. São Paulo: Cultrix, 2002.

TARALLO, F. *A pesquisa sociolinguística*. São Paulo, Ática, 1997.

TARALLO, Fernando. Sobre a alegada origem crioula do português brasileiro: mudança sintática aleatória. In: ROBERTS Ian e KATO, Mary Aizawa (orgs). *Português brasileiro*. Campinas: UNICAMP, 1996.

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

TEIXEIRA, E. P. Aspectos da pronominalização no português vernacular de Luanda: uma comparação com o português do Brasil. In: LOPES, N.; BULHÕES, L.; CARVALHO, C. (Orgs.). *Sociolinguística: estudos da variação, da mudança e da sóciohistória do português brasileiro*. Feira de Santana: Editora UEFS, 2013.

_____. A representação do sujeito pronominal no português popular angolano. *PAPIA: Revista brasileira de estudos crioulos e similares*. n. 22(1). 2012. pp. 141-160.

TEYSSIER, P. *História da língua portuguesa*. Tradução Celso Cunha. 3 ed. São Paulo: Martins Fontes, 1997.

VENÂNCIO, R. P. Por uma história social do português no Brasil. In: RAMOS, J. M.; ALKMIM, M. A. (orgs.). *Para a história do português brasileiro: estudos sobre a mudança linguística e história social*. vol. 5. Belo Horizonte: FALÉ/ UFMG, 2007.

WEINREICH, W.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução de Marcos Bagno. São Paulo: Parábola Editorial, 2006.

XAVIER, M. F.; MATEUS, M. H. M. (Org.). *Dicionário de termos lingüísticos*. Lisboa: Cosmos, 1990.