

CONCORDÂNCIA NOMINAL EM TRÊS VARIEDADES DO PORTUGUÊS: RESULTADOS GERAIS, NOVAS INDAGAÇÕES

NOMINAL AGREEMENT IN THREE VARIETIES OF PORTUGUESE: MAIN RESULTS, NEW QUESTIONS

SILVIA FIGUEIREDO BRANDÃO

Universidade Federal do Rio de Janeiro/CNPq/FAPERJ

silvia.brandao@terra.com.br

Expõem-se, de forma sintética, os principais resultados de estudos de base sociolinguística sobre a concordância nominal de número em três variedades urbanas do Português desenvolvidos no âmbito do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*. Apresentam-se, ainda, algumas reflexões com base na análise da variável *Frequência de uso de um crioulo* no caso da variedade são-tomense, de modo a tratar de questões relacionadas ao contato multilinguístico, e de estudos de cunho geo-sociolinguístico sobre o /S/ em coda silábica na fala do Rio de Janeiro, com o intuito de discutir a possível interrelação entre a tendência à simplificação da estrutura da sílaba e o cancelamento da marca nominal de número.

Palavras chave: Concordância nominal, variedades do Português, contato multilinguístico, /S/ em coda.

We expose, in a synthetic way, the main results of sociolinguistic studies on nominal plural agreement in three urban varieties of Portuguese, developed under the Project *Comparative Study of agreement patterns in African, Brazilian and European varieties of Portuguese*. We also present some insights based on the analysis of the variable *Frequency of use of a creole*, in the case of Sao Tome variety, in order to deal with questions relating to multilingual contact, and on geo-sociolinguistic studies on /S/ in syllabic coda in the speech of Rio de Janeiro, with the aim of discussing the possible interrelationship between the tendency for simplifying the syllabic structure and the deletion of the nominal number mark.

Keywords: Noun agreement, varieties of Portuguese, multilinguistic contact, /S/ in coda.

0. INTRODUÇÃO

A importância do estudo da concordância – quer a verbal, quer a nominal – não só para a discussão das origens do Português do Brasil (PB), mas também para a compreensão do processo que levou à situação de polarização sociolinguística que o caracteriza suscitou o interesse pela análise do tema em outras variedades, como a Europeia (PE) e a de São Tomé (PST).

O foco no PE objetivava, de um lado, testar a hipótese de que, nele, a concordância teria caráter categórico ou semicategórico, nos termos de Labov (2003), e, de outro, verificar se procedia a hipótese de Naro & Scherre (2007) de que, nessa variedade, padrões variáveis de concordância poderiam ser detectados na fala de indivíduos de baixa ou nula escolaridade, o que, inclusive, levou esses pesquisadores a buscarem tais padrões em estudos de cunho dialetal e em antigos estágios do português, em especial em oito textos pré-clássicos.

O foco no PST, por sua vez, objetivava verificar se, em variedades africanas, se encontraria um quadro semelhante ao do PB por força de alguns aspectos relacionados ao seu povoamento, assemelhados (mas não idênticos) aos que ocorreram no Brasil e que teriam redundado também num quadro de polarização sociolinguística. De certo modo, se estaria testando, indiretamente, outra hipótese, a de Lucchesi (2003), de que o PB popular decorreria de um processo de aprendizagem linguística irregular. A escolha do PST, dentre as variedades africanas, deveu-se ao fato de São Tomé e Príncipe ser, nas palavras de Hagemeyer (2009), “o único país da África de língua portuguesa onde a maioria da população tem actualmente o Português como primeira língua, havendo, assim, condições para a emergência de uma nova variedade” (p.19-20). De acordo com o censo de 2001, o Português é ali falado por 98,9% da população, apesar de com ele coexistirem outras línguas, entre as quais o forro ou santomé, crioulo de base portuguesa, falado por 72,4% do contingente populacional com mais de cinco anos. Além disso, dado o caráter fortemente multilinguístico da sociedade santomense, o PST poderia fornecer indícios sobre o que teria ocorrido na formação do PB, variedade cuja emergência se deu em meio ao contato de várias línguas (gerais, indígenas e africanas, entre outras).

Com base em tais hipóteses, elas também polarizadas, deu-se início, em 2009, no âmbito do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias*, inicialmente apoiado pela CAPES e, desde 2011, desenvolvido como um projeto da ALFAL (o de número 21), a pesquisas sobre concordância nominal, que partiam da constatação de que, em Português, existem diferentes padrões de indicação de pluralidade, que implicam, num extremo, a presença do morfema /S/ em todos os constituintes do SN, como nos exemplos (1) e (2), a chamada concordância canônica, até, em outro extremo, a presença do morfema apenas em um de seus constituintes, como em (4), havendo, ainda, padrões intermediários, como em (3).

(1)...em conversa com [outras crianças mais/mais desprendidas] (Informante PE)

(2)...nunca me interessei [nos outros clubes] (Informante PST)

(3)...[todos os político()], não, o Lindinho (Informante PB)

(4)...peguei [as minha() ferramenta nova ()] (Informante PST)

Os *corpora* em que se apoiou a pesquisa foram organizados entre 2008 e 2010, de modo que, a partir de 2011, foram publicados artigos que iam dando conta, paulatinamente, dos resultados obtidos, alguns deles decorrentes de novas hipóteses por eles suscitadas. Entre tais artigos, encontram-se Brandão 2011a, 2013; Brandão & Vieira 2012a, 2012b; Vieira & Brandão (2014).

Neste trabalho, tem-se por objetivo relatar, muito brevemente, os principais resultados obtidos, até o momento, nas análises sobre a concordância nominal nas mencionadas variedades do Português e, sobretudo, ampliar algumas questões apresentadas por Brandão (2013: 90-93) na seção denominada “Some (few) insights” e que dizem respeito ao forte contraste entre “o

caráter aparentemente monolítico da regra de concordância nominal no PE, usada, ao que tudo indica, segundo o padrão canônico, em diferentes regiões (no continente, na Ilha da Madeira), por falantes de todos os níveis de escolaridade” e o caráter francamente variável da regra no PB e no PST.

Para tanto, na seção 1, apresentam-se os fundamentos teórico-metodológicos que nortearam as pesquisas, bem como se delineiam as áreas focalizadas nos estudos recém-citados. Na seção 2, indicam-se as diferentes análises realizadas e os resultados mais relevantes delas decorrentes em termos contrastivos. Na seção 3, sintetizam-se as mencionadas “reflexões” e apresentam-se, em caráter preliminar, elementos que possam contribuir para o debate sobre as origens do PB, focalizando, com base em dados sincrônicos, o cancelamento da marca de número (a) de um ponto de vista morfossintático, isto é, como um processo decorrente de tendência à redução de flexões, observado na formação de pidgins e crioulos e em processos de aprendizagem irregular derivada de contato linguístico (seção 3.1), e (b) de um ponto de vista mais propriamente fonético-fonológico, isto é, buscando indícios da vinculação desse processo à tendência à redução da sílaba à estrutura CV, fato considerado produtivo ao longo da história da língua e atribuído a uma deriva românica, embora também seja indicado como característico da constituição de pidgins e crioulos (seção 3.2). Na seção 4, tecem-se algumas considerações finais em forma de indagações.

1. ASPECTOS TEÓRICO-METODOLÓGICOS E PERFIL DAS COMUNIDADES

Todas as análises foram conduzidas segundo os pressupostos da Teoria da Variação e Mudança (Weinreich, Labov, Herzog 1968) e os desdobramentos da Sociolinguística Variacionista (Labov 1972, 1994, 2001, 2003). Organizaram-se as amostras com base em dados selecionados de entrevistas do tipo DID (Diálogo entre Informante e Documentador), realizadas, como já se observou, entre 2008 e 2010, com indivíduos distribuídos por sexo, três faixas etárias (18-35 anos; 36-55 anos; 56-75 anos) e três níveis de instrução (fundamental: 5 a 8 anos, médio: 9 a 11 anos; superior: 12 ou mais anos), todos naturais das localidades de perfil urbano definidas na Ilha da Madeira (Funchal), e nas regiões metropolitanas de Lisboa (Oeiras e Cacém), do Rio de Janeiro (Copacabana e Nova Iguaçu) e de São Tomé e Príncipe (São Tomé). Levaram-se em conta na análise do PB e do PST, além das três mencionadas variáveis sociais, oito variáveis de cunho estrutural: (1) saliência fônica; (2) tonicidade e (3) número de sílabas do item singular; (4) classe do vocábulo; (5) posição linear e relativa do constituinte no SN; (6) marcas precedentes; (7) contexto fonológico subsequente; (8) animacidade do núcleo do SN. Os dados foram submetidos ao Pacote GOLDVARB-X, de análise estatística, que fornece índices percentuais e pesos relativos concernentes aos fatores que motivam ou inibem a aplicação de regras variáveis. Na explanação dos resultados, tomou-se como valor de aplicação a marcação de plural. No caso do PE, consideraram-se os SNs como um todo, tendo em vista não haver condições, como se verá adiante, de realizar análises variacionistas com a amostra organizada.

Tendo em vista as dimensões territoriais dos países envolvidos e, conseqüentemente, o contingente populacional de cada um deles, bem como suas especificidades sócio-histórico-culturais, o contraste entre as variedades tem de ser relativizado. A Área Metropolitana de Lisboa ocupa 2.870km² e conta com 2.819. 433 habitantes, o que corresponde a aproximadamente ¼ da população portuguesa. A Área Metropolitana de Funchal tem uma população de cerca de 225 mil habitantes. Funchal, com 76,15km² e 111.892 habitantes, é a capital da Região

Autônoma da Madeira e a cidade mais populosa fora do território continental português. A Região Metropolitana do Rio de Janeiro tem 5.645km², onde vivem 12.434.611 indivíduos, ou seja, 76% da população do Estado. Na área urbana de São Tomé, encontra-se cerca de um terço da população de São Tomé e Príncipe (206.178 habitantes). A seleção das comunidades representativas das áreas metropolitanas do Rio de Janeiro e de Lisboa baseou-se nos critérios de prestígio e semelhanças sócio-culturais: respectivamente, Copacabana e Oeiras, as de altos índices de desenvolvimento humano, e as de Nova Iguaçu e Cacém, menos prestigiadas, em função de seus habitantes estarem em permanente interação com os moradores das respectivas cidades-sede, pois a maioria deles nelas desenvolvem atividades profissionais. Maiores detalhes sobre as comunidades podem ser obtidos nos estudos de Brandão e de Brandão & Vieira acima referidos.

2. AS ANÁLISES REALIZADAS: RESULTADOS MAIS RELEVANTES

Diversas análises foram realizadas, ora confrontando-se globalmente todas as áreas, ora centrando-se a atenção em um determinado grupo de falantes ou em uma variedade, conforme se discrimina na Tabela 1, devendo-se ressaltar que cada comunidade foi estudada individualmente.

Tipos de análise	Características da amostra	Número de dados
Contrastiva global	18 informantes por amostra do PE (Lisboa, Oeiras, Funchal), do PB (Copacabana e Nova Iguaçu) e do PST (São Tomé), considerando-se as três faixas etárias e os três níveis de escolaridade.	PE: 6.952 SNs PB: 7.493 constituintes do SN PST: 2.923 constituintes do SN
Restrita à variedade popular do PST	22 informantes de São Tomé distribuídos por sexo e faixa etária e dois níveis de instrução: fundamental (5 a 8 anos) e média (9-11 anos).	2.375 constituintes do SN
Restrita à fala de estudantes do PST	9 estudantes de diferentes níveis de instrução.	633 constituintes do SN

Tabela 1. Análises realizadas com diferentes amostras do PE, do PB e do PST

A análise realizada com dados do PE demonstrou que a regra de concordância nominal é implementada por todos os grupos de indivíduos considerados, tendo-se observado apenas oito casos de ausência de marca num dos constituintes nos 6.952 SNs considerados nas três comunidades, grande parte deles decorrente daquilo que Labov (2003)¹ indica como lapso quando uma regra é categórica. Observe-se, em duas das ocorrências que aqui servem de exemplo, que o indivíduo reorganiza o discurso imediata e automaticamente.

(5) a maioria d[as **pessoa**] das **mulheres** era tra/ era em casa era a bordar

(6) [outros **conhecimento**] outra **maneira** de ver as pessoas

Já no que se refere ao PB e ao PST, a regra de concordância é variável entre falantes de níveis fundamental e médio de escolaridade e semicategórica entre os falantes cultos, como se observa por meio dos índices percentuais que se expõem na Tabela 2 (*cf.* nota 1).

¹ Labov (2003) estabelece a existência de três tipos de regras linguísticas, com base nos índices de frequência com que operam: categóricas (100%), semicategóricas (95% a 99%) e variáveis (5% a 95%)

Constituintes flexionáveis do SN				
Amostra	Com marca de número		Sem marca de número	
	Nº de OCO	%	Nº de OCO	%
Copacabana	3432/3716	92,4	284/3716	7,6
Nova Iguaçu	3439/3777	91,1	338/3777	8,9
São Tomé	2524/2612	93,4	173/2612	6,6

Tabela 2. Distribuição de constituintes do SN com e sem marca de número em variedades urbanas do PB e do PST Fonte: Brandão (2013)

Análises multivariadas com os dados brasileiros e santomenses mostram que a ausência/presença da marca de número é motivada nas duas variedades por restrições de cunho estrutural (saliências posicional, semântica – traço de animacidade – e fônica – diferenciação entre as formas singular e plural) e social (nível de instrução, sexo, faixa etária). Dentre tais restrições, as que se efetivam, sistematicamente, como as mais salientes em todas as comunidades de fala são, de um lado, *Posição linear e relativa do constituinte no SN* e, de outro, *Nível de instrução*.

Com base nos resultados referentes à variável posicional, foi possível traçar o que se denominou de *continuum de marcação de plural no SN*, que se acredita válido não só para as variedades aqui representadas, mas também para as variedades africanas L1 ou L2.

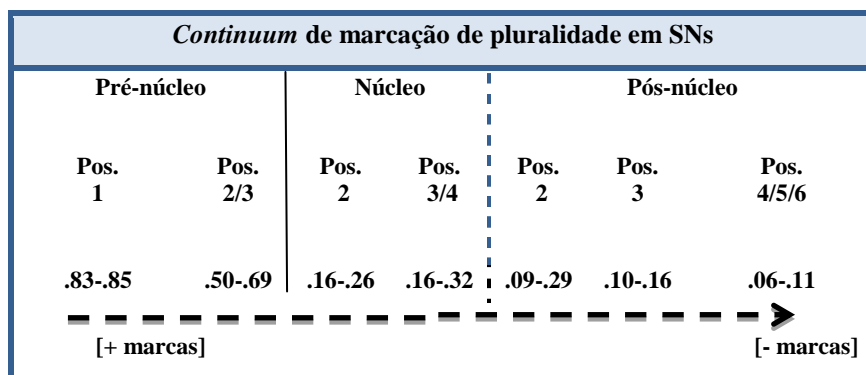


Figura 1. Continuum de marcação de plural em SNs prototípicos no PB e no PST, com base nos pesos relativos obtidos nas diferentes análises realizadas. Fonte: Brandão (2013)

Levando-se em conta o que aqui se denomina de SN prototípico, isto é, aquele composto por constituinte(s) pré-nuclear(es), núcleo e constituinte(s) pós-nuclear(es), verifica-se, por meio dos pesos relativos obtidos nas diferentes análises, que a primeira posição linear, o pré-núcleo, tende a ser mais marcada, o que sugere seja esse o *locus* por excelência da marca. A partir do núcleo em segunda posição, vai decrescendo gradativamente a presença da marca, quer se considerem os menores ou os maiores índices obtidos. No esquema proposto, a linha vertical contínua que separa o pré-núcleo do núcleo representaria o limite do espaço mais saliente para a marcação (o *locus* esquerdo), enquanto a linha vertical pontilhada entre o núcleo e o pós-núcleo indicaria que essas áreas não sofreriam, quanto à tendência à não marcação, solução de continuidade.

Do mesmo modo, os índices referentes à variável extralinguística permitem delinear outro *continuum*, o de nível de instrução, este também passível de ocorrer em outras variedades africanas do português. Observe-se que, à exceção de Nova Iguaçu, em que indivíduos de nível fundamental apresentam maior propensão à implementação da marca do que os de nível médio, em Copacabana e em São Tomé a marcação de plural no SN tem caráter nitidamente escalar, aumentando na fala dos menos aos mais escolarizados.

<i>Continuum de nível de instrução</i>			
	----->		
[- marcas]		[+marcas]	
Não escolarizados/ Escolaridade: 1 a 4 anos	Fundamental	Médio	Superior
	.9 - .14 - .38	.24 - .37 - .49	.76 - .76 - .82
	ST- COP-NIG	NIG-COP-ST	ST-NIG-COP

Tabela 3. *Continuum* de nível de instrução, com base em pesos relativos referentes à presença de marca de número, obtidos nas diferentes análises realizadas (ST: São Tomé; COP: Copacabana; NIG: Nova Iguaçu)

3. ALGUMAS (NOVAS) “REFLEXÕES”

Nesta seção, busca-se ampliar o debate iniciado em Brandão (2013), apresentando novos dados para reflexão. No mencionado estudo, a autora questiona o grande contraste que se verifica entre a variedade europeia - em que a regra de concordância nominal se mostrou categórica em todos os estratos sociais, seja no continente, seja na Ilha da Madeira - e as variedades não europeias, em que a regra é variável, embora entre os falantes de nível superior de instrução, sem dúvida por força da maior exposição aos bens culturais, se possa classificá-la como semicategórica. Dos aspectos por ela arrolados, caberia ressaltar as seguintes passagens:

- (a) “os dados de que se dispõe sobre o PB e, agora, sobre o PST, confrontados com os que vêm sendo divulgados sobre o Português L2 de Cabo Verde e de Moçambique (JON-AND 2010, 2011), sugerem que o massivo contato interlinguístico, que teria determinado uma aprendizagem irregular do Português, seria a motivação de base da regra variável que se acabou de descrever”.
- (b) “Nas gramáticas do século XVI e nas que as seguiram nos séculos XVII e XVIII, não há menção a padrões variáveis de concordância, embora diversas sejam as observações sobre aspectos fonético-fonológicos, em termos de correção e/ou de variação. Sabe-se que a variedade europeia passou por mudanças no século XVIII, mas estas parece não terem afetado o sistema flexional, e sim o sistema vocálico, dentre as quais a mais significativa foi a neutralização das vogais médias e altas em contexto pretônico, como se verifica em Teyssier (1997) e Castro (1991). Vale lembrar que Peres & Moia (1995), que indicam a concordância como “uma das áreas críticas” da língua portuguesa, não arrolam casos variáveis na estrutura do SN”.
- (c) “Dentre os fatores que poderiam ser indicados como motivadores do padrão canônico de concordância no PE, estaria a propensão à manutenção do padrão silábico (C)VC, na contramão da tendência universal à sílaba CV. No PB, em que o cancelamento e a vocalização de consoantes em coda são processos usuais, mesmo quando constituem marcas morfológicas (o /R/, do infinitivo, quase não mais produzido em alguns dialetos, o /S/ marcador de número nominal, que pode não ocorrer em todos os constituintes do SN), a marca de número que, na norma de alguns falantes, ocorre, em geral, apenas no

constituente pré-nuclear em primeira posição, parece espalhar-se por todo o SN: [os[gato preto]]SN +plural)”.
 (Brandão 2013: 90-91)²

Brandão ocupou-se, sobretudo, de tratar a questão indicada em (c), valendo-se de argumentos de ordem fonético-fonológica em consonância com trabalhos no âmbito da aquisição, como os de Mezzomo (2004), Freitas (*apud* Almeida 2007) e Cappelari & Zilles (2002). Com base neles, verificou que as crianças lusas e as brasileiras adquirem a coda silábica final antes da medial, que “trata a coda lexical e a coda morfológica de maneira singular desde o início das produções até sua estabilização” e que, “apesar de emergir após a coda lexical, a coda morfológica parece ajudar na estabilização precoce da coda final em relação à medial” (Mezzomo 2004, s/p). Por fim, concluiu que, apesar desses indícios, a aplicação ou não da marca de número no SN, como destacam Cappelari & Zilles (2002: 195), comentando um estudo de Lamprecht, depende do *input* a que estão expostas as crianças, pois algumas delas “parecem acreditar que não é necessária a marcação do plural quando este fica claro no determinante ou por meio de um numeral. Em resumo, as crianças conhecem o sistema mas parecem não saber que a marcação é desejável”. As autoras acrescentam, ainda, que o *input* teve influência evidente na pesquisa de Lamprecht, que “destaca o desempenho de duas crianças, designadas por ela como *marcadoras*, e constata que essas crianças provêm de famílias em que a linguagem usada é muito próxima da norma culta” (Cappelari & Zilles (2002: 197).

Pretende-se aqui retomar o que se assinala em especial em (a) com base nos resultados de uma das análises realizadas com dados do PST e de outras variedades africanas do português, e em (c) com apoio em estudos sobre a variável /S/ na fala do Rio de Janeiro.

Vale relembra, no que concerne às hipóteses sobre as origens do PB popular, que há, basicamente, duas linhas interpretativas: (a) a que vincula as características dessa variedade ao massivo contato multilinguístico e multiétnico que se instaurou no Brasil durante a fase de colonização – postulando-se, em sua versão mais forte, um processo de criouliização/descruiolização (Guy 1981, 1989; Holms 2003), e, em sua versão mais leve, um processo de transmissão linguística irregular (Lucchesi 2003, 2012) – e (b) a que atribui tais características à deriva da língua, em particular das línguas românicas (Naro & Scherre 2003).

Sabe-se que os portugueses, quando do início da colonização, já encontraram uma língua geral de base tupinambá adotada pelos índios da costa – pela qual optaram para fins de intercomunicação e catequese – e mais tarde denominada de língua geral paulista, cujo uso perduraria até o início do século XVIII. Também, no norte do país, havia a chamada língua geral amazônica ou nheengatu, ainda hoje falada em São Gabriel da Cachoeira, no Alto Rio Negro, Amazonas (Nobre 2011). No entanto, apesar do significativo contingente de africanos importados para servirem como escravos e que, em alguns momentos, ultrapassava em grande escala o número de homens brancos, não se tem registro cabal de pidgins ou crioulos de base portuguesa utilizados pelas diferentes etnias que aqui chegaram e que pudessem justificar as características encontradas sobretudo na variedade popular do Português do Brasil, em especial no que respeita à concordância nominal e verbal. Servem de apoio para justificar a possível existência de tal(is) língua(s), características encontradas em comunidades afrodescendentes isoladas, em geral remanescentes de antigos quilombos (como em Helvécia, por exemplo), caracterizadas por Lucchesi (2012: 81) como “sítios arqueológicos da história sociolinguística do Brasil”.

² O texto original está em inglês.

Partindo-se do princípio de que, no presente, se podem buscar indícios do que poderia ter ocorrido em outros estágios do desenvolvimento de uma língua, propõe-se, nesta seção, questionar, à guisa de exercício especulativo, possíveis interferências dos alegados contato multilinguístico e tendência à simplificação da estrutura silábica (CV) para o que se observa, no PB, no âmbito da flexão de número no SN. Para tratar do primeiro desses aspectos, recorre-se aos resultados concernentes à variável *frequência de uso de um crioulo*, controlada em uma análise sobre o Português de São Tomé (Brandão 2011b), bem como a índices de cancelamento da marca de número obtidos em pesquisas sobre outras variedades africanas faladas como L1 ou L2. Para abordar o segundo deles, utilizam-se percentuais referentes ao cancelamento do /S/ em contexto medial e final de vocábulo em algumas comunidades do Rio de Janeiro.

3.1. A variável *Frequência de uso de um crioulo*

Na amostra referente a São Tomé e organizada com base na fala de 18 informantes distribuídos por três níveis de instrução e três faixas etárias, observa-se alto índice de marcação de número no SN (93,4%), o que se atribui ao fato de a amostra ser representativa da fala da área urbana, onde não só é mais fácil o acesso à educação formal e informal, mas também há maiores oportunidades de trabalho que requerem proficiência em português, o que serve de incentivo ao domínio da norma de prestígio, no caso a do PE, veiculada nas escolas e nos meios de comunicação, e que, na visão de alguns dos informantes, é fundamental para a almejada ascensão social.

Não obstante esse quadro geral, uma análise mais detida demonstra que há, mesmo na zona urbana, um quadro de polarização sociolinguística bem semelhante ao que se detecta no PB, como se expõe na Figura 1. Levando-se em conta o *input* (.84) geral de aplicação da marca de número, a diferença entre os pesos relativos obtidos, na análise variacionista, para os níveis de escolaridade são bastante expressivos: .40, entre os menos e os mediantemente escolarizados e .26 entre estes e os mais escolarizados. Considerando-se os dois polos da escala, a diferença chega a .67.

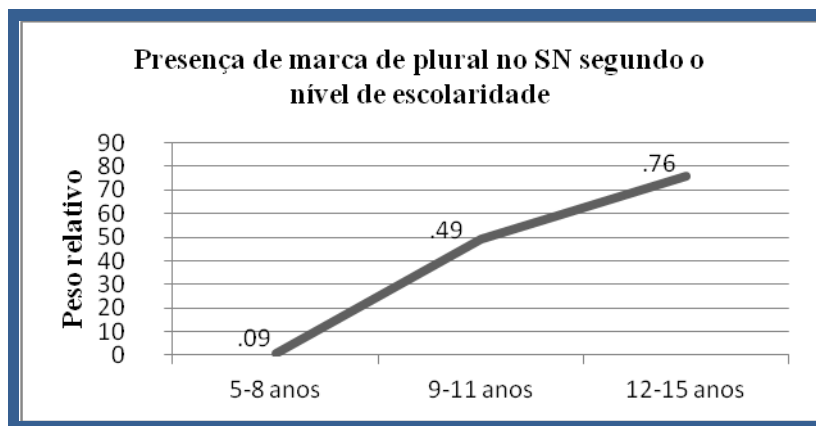


Figura 2. Pesos relativos referentes à presença de marca de número no SN em São Tomé segundo o nível de escolaridade

No sentido de melhor compreender esse quadro e de buscar indícios de que o contato multilinguístico continua a exercer, na sociedade santomense, influência na performance linguística dos falantes de português, decidiu-se focalizar a fala apenas dos indivíduos de níveis fundamental e médio, acrescentando mais dez informantes aos já pré-existentes, num total de

22, para testar a hipótese de que, apesar de todos eles se terem declarado falantes de português como L1, os indivíduos que, nas situações de intercomunicação diária, fazem mais uso de uma língua crioula do que do português, apresentariam maior probabilidade de cancelar a marca de número do que aqueles que se expressam fundamentalmente nesta última língua. Para tanto, formulou-se a variável *Frequência de uso de um crioulo*, tendo servido de referência para a determinação dos fatores que a compõem as informações prestadas pelos informantes nas entrevistas. Os fatores retratam situações prototípicas de São Tomé. Há indivíduos que (a) só usam o português e (i) não dominam ou (ii) dominam um crioulo; (b) usam o português e um ou mais crioulos (fundamentalmente o forro e/ou, ainda, o angolano, o lung'ie, o cabo-verdiano) (i) com menor ou (ii) maior frequência (em áreas públicas, como o mercado, ou em momentos de lazer, com os amigos). Assim, a variável ficou constituída pelos fatores: frequência do uso de um crioulo de (1) grau zero/baixo (GØ); (2) de grau médio (G1), de grau alto (G2). Na tabela a seguir, indica-se a distribuição de tais graus pelos informantes, verificando-se que, dos 22, nove se distribuem pelo GØ, dez pelo G1 e apenas três pelo G2.

Escolaridade	Nível fundamental (5 a 8 anos)		Nível médio	
	Masculino	Feminino	Masculino	Feminino
18-35 anos	ST-A-1-h-a (GØ)	ST-A-1m-a (G1)	ST-A-2-h-a (G1)	ST-A-2-m-a (GØ)
	ST-A-1-h-b (G1)	ST-A-1m-a-b (G2)	ST-A-2-h-b (GØ)	ST-A-2-m-b (GØ)
36-55 anos	ST-B-1-h-a (G1)	ST-B-1-m-a (GØ)	ST-B-2-h-a (G1)	ST-B-2-m-a (G1)
	ST-B-1-h-b (G1)	ST-B-1-m-b (G1)	ST-B-1-h-b (GØ)	ST-B-2-m-b (GØ-P)
56-75 anos	ST-C-1-h-a (G2)	ST-C-1-m-a (G1)	ST-C-2-h-a (GØ)	ST-C-2-m-a (GØ)
	-----	ST-C-1-m-b (G2)	ST-C-2-h-b (G1)	-----

Tabela 4. Distribuição dos informantes segundo as variáveis linguísticas e o grau de frequência de uso de um crioulo³.

Nessa análise, em que foram selecionadas a *posição linear e relativa do constituinte no SN*, o *nível de instrução*, a *saliência fônica* e a *animacidade do núcleo*, a variável *frequência de uso de um crioulo* foi a terceira a se mostrar mais relevante.

Independentemente da distribuição assimétrica, os indivíduos que se comunicam preferencialmente num crioulo são os que mais tendem a não implementar a marca de número (.25). Os três que se enquadram nesse caso têm nível fundamental de instrução: duas mulheres, uma da faixa mais jovem, outra da mais velha e um homem, também desta última faixa etária. Com esses indivíduos, contrastam os que nunca ou pouco se expressam em crioulo (p.r. .48) e os que o utilizam apenas eventualmente (p.r. 59).

³ Na tabela, aparecem as siglas referentes aos informantes, que contêm as seguintes informações: ST= São Tomé; faixa etária: A= 18-35 anos, B = 36-55 anos, C = 56-75 anos; nível de escolaridade: 1= fundamental, 2 = médio, 3= superior; sexo: m= mulher; h=homem; a (informantes da amostra anterior) e b (novos informantes).

Frequência de uso de um crioulo	P.R.
<i>Zero /Baixa (grau zero - Gϕ)</i>	.59
<i>Média (grau 1 - G1)</i>	.48
<i>Alta (grau 2 - G2)</i>	.25

Tabela 5. Atuação da variável *frequência de uso de um crioulo* para a implementação da marca de número no SN

Deve-se ressaltar que inclusive os que fazem pouco uso de um crioulo estão em contato regular com falantes multilíngues, sendo natural que, em determinadas situações de interação, sejam capazes de reconhecer e decodificar estruturas de outras línguas, o que implica, mesmo de forma indireta, algum grau de interferência em sua performance linguística, justificando-se, assim, o baixo índice (.59) sugerido para aplicação da marca de número para esse grupo de falantes.

A questão da interferência de outras línguas confirmada pelos resultados ora apresentados há muito vem sendo debatida. A tendência à marcação de plural no constituinte pré-nuclear adjacente ao núcleo tem sido apontada, de acordo com Jon-And (2010, 2011), como argumento para uma possível influência de línguas africanas (transferência ou influência de substrato) em variedades não europeias do português. Segundo a autora, nos casos do Português de Moçambique e de São Tomé (o grupo dos Tongas e a localidade de Almojarife), a hipótese se confirmaria, uma vez que essas variedades têm como substrato línguas banto, em que a marcação de plural é feita por meio de prefixo.

No entanto, ainda segundo Jon-And, a mesma explicação não se aplicaria a Cabo Verde, que não tem línguas banto como substrato. Nas línguas de substrato do crioulo cabo-verdiano, das famílias atlântico e mande, há uma variada forma de indicar a pluralidade, sendo a marcação à direita, por meio de sufixo, a mais usual. No entanto, no crioulo cabo-verdiano predomina a marcação no determinante em primeira posição, embora outros constituintes possam também receber marca. Da mesma forma, no português L2 de áreas urbanas de Moçambique e Cabo Verde, é essa a regra predominante, apesar de, nesta última variedade, os índices de aplicação da marca subirem a 82%, quando se consideram, em conjunto, analfabetos e escolarizados até o quarto ano.

Levando-se em conta o princípio do uniformitarismo, preconizado por Labov (1972) e que pressupõe a atuação, no estágio atual de uma língua, dos mesmos mecanismos que promoveram variação e mudanças ao longo do tempo e independentemente dos diversos argumentos que se apresentem, pode-se inferir que o quadro que se verifica, hoje, no PST serve de base para que se postule a interferência das diversas línguas em contato na gênese do Português do Brasil e nas variedades africanas, que têm no *continuum* retratado na Figura 1, acima, sua marca indelével.

3.2. A questão da simplificação da estrutura silábica

Em grande parte das línguas, o contexto pós-vocálico é altamente favorável à ocorrência de um reduzido número de segmentos consonantais, derivados, em geral, de processos de neutralização. Esse é o caso do português, que só admite, nessa posição, o rótico, a fricativa coronal, a lateral anterior e a nasal. Dentre eles, dois podem ter, eventualmente, valor mórfico, como o rótico (desinência de infinitivo ou de futuro do subjuntivo) e a fricativa (desinência

nominal de número ou desinência verbal número-pessoal). Também em grande parte das línguas, predomina a estrutura CV, considerada de caráter universal⁴.

Nos debates sobre as origens do PB, tem-se justificado a simplificação da estrutura silábica ora como um processo usual na constituição de pidgins e crioulos, ora como uma tendência inerente ao português, ou, em outros termos, como um fato da deriva que poderia ser atestado no decurso da história da língua e, inclusive, no da de outras línguas românicas. Nesse sentido, o PB e o PST atuais estariam seguindo essa deriva, contrariamente ao PE, que tende a manter e reforçar as consoantes em coda, quer o segmento tenha ou não valor mórfico.

Naro & Scherre (2005), no intuito de demonstrarem que uma série de processos fonético-fonológicos apontados como característicos do PB popular se encontram no PE, fazem, com base em quatro obras de cunho dialetal, o que denominam de um “garimpo fonológico” (cap. 5), em complementação aos capítulos 3 e 4, dedicados à coleta de ocorrências de cancelamento das marcas de número nominal e verbal em obras antigas e em estudos dialetais. Nele focalizam “cinco grandes grupos de processos fonológicos que asseguram a estrutura silábica CV” (p. 122), dentre os quais a queda de consoantes em coda silábica (/l/, /s/ e /r/). Dos dados referentes a /s/ sem valor mórfico, arrolam três casos – *reposta* (resposta), *mema* (mesma) e *refiado* (resfriado) – aparentemente como ocorrências únicas.

Assim, achou-se interessante observar o que sucede especificamente com o /s/ pós-vocálico em dois estudos de cunho dialetal (Almeida 2008 e Lima 2006), que focalizam a fala de indivíduos de nível fundamental de escolaridade e naturais de 16 localidades do Estado do Rio de Janeiro, entre elas Nova Iguaçu. Esses trabalhos constituem, respectivamente, uma tese de doutoramento (com 12 pontos de inquérito⁵) e uma dissertação de mestrado (com 4 pontos na Região Metropolitana do Rio de Janeiro⁶), que objetivaram organizar atlas fonéticos segundo os mesmos critérios de escolha de informantes⁷ e com base num mesmo questionário. Nas suas monografias, as autoras desenvolveram estudos sobre o /s/ com base nas respostas registradas nas cartas fonéticas, estudos esses que aqui se retomam para fins de exemplificação.

No questionário utilizado pelas duas autoras, há 58 cartas em que ocorre /s/ em coda: 31 em contexto interno e 27 em contexto externo: em 17 delas, o /s/ é marca de número nominal e, nas 10 restantes, não tem valor mórfico.

No estudo de Almeida (2002: 117) obtiveram-se, com base nos dados cartografados e relativos a 12 localidades, 4.062 ocorrências de /s/ em coda, assim distribuídas:

- 53% (2.144) em coda interna,
- 30% (1216), em coda externa com valor não-morfêmico, e
- 17% (702), em externa com valor morfêmico

Almeida estava particularmente interessada na concretização do segmento (como fricativa alveolar, pós-alveolar ou glotal), a fim de delimitar possíveis isoglossas. Por esse motivo, as tabelas em que apresenta seus resultados indicam o número de dados referente a cada variante por vocábulo. Como o que está em causa no presente trabalho é a concretização *versus* o cancelamento do segmento, as tabelas organizadas pela autora foram simplificadas.

⁴ Ao que tudo indica, o arrente, falado por um pequeno grupo de australianos, seria o único sistema que fugiria a esse princípio universal: nele as sílabas são do tipo VC(C), segundo BREEN, G.; PENSALFINI, R. Arrente: a language with no syllable onsets. *Linguistic Inquiry*, 30: 1-26, 1999. Informação obtida em <http://www.mitpressjournals.org/doi/abs/10.1162/002438999553940?journalCode=ling#.VC1Vqf4tDDc>. Acesso em 10 de agosto de 2014.

⁵ As 12 localidades são: São Francisco de Itabapoana, Quissamã, Cabo Frio, Itaguaí, Paraty, no litoral, e Porciúncula, Santa Maria Madalena, Cantagalo, Cachoeiras de Macacu, Valença, Três Rios e Resende, na área interiorana.

⁶ Nova Iguaçu, Duque de Caxias, Magé e Itaboraí.

⁷ Em cada localidade, inquiriram-se 6 informantes, distribuídos por sexo e três faixas etárias (18-35 anos; 36-55 anos; 56-75 anos).

Nº da carta	Vocábulo	Alveolar/Palatal/ Aspirada	Cancelamento	Total de ocorrências
19	Isca	70	0	70
54	Casca	72	0	72
55	Caspa	72	0	72
141	Cristo	71	0	71
81	Dentista	72	0	72
83	Desmaio	70	0	70
95	Escola	72	0	72
97	Escova	72	0	72
99	Escuro	72	0	72
101	Espinho	71	0	71
103	Esquecer	71	0	71
105	Estátua	69	1	70
107	Estômago	71	0	71
109	Experiência	65	0	65
110	Festa	71	0	71
119	Fósforo	49	4	53
134	Hospital	70	0	70
139	Intestino	65	2	67
149	Lesma	71	0	71
160	Máscara	72	0	72
161	Mastigar	71	0	71
167	Mesmo	67	2	69
170	Mosca	72	0	72
178	Nordestino	66	1	67
196	Pasta	72	0	72
205	Pescoço	72	0	72
208	Plástico	43	1	44
215	Poste	67	0	67
256	Transporte	70	1	71
292	Sexta	72	0	72
302	Agosto	72	0	72
Totais		2132 99,4%	12 0,6%	2144

Tabela 6. Ocorrências de /S/ em coda silábica interna: concretização *versus* cancelamento nos dados do MicroAFERJ. Fonte: Almeida (2008: 118)

Como se verifica, é ínfimo o índice de cancelamento da coda em contexto interno (12 em 2144 ocorrências: 0,6%), inclusive em vocábulos que, na fala corrente, tendem a perder o segmento, como é o caso de *mesmo*, com apenas 2 cancelamentos. Nas tabelas 7 e 8, discriminam-se as cartas referentes ao /S/ em coda externa, respectivamente, sem e com valor de morfema.

Na tabela 7, verifica-se que é baixo, também, o índice de cancelamento (3%) quando o /S/ não constitui marca de número. No entanto, deve-se atentar para o fato de que vocábulos terminados por /S/ sem valor mórfico são predominantemente oxítonos ou monossílabos tônicos, uma vez que, em português, sílabas fechadas (pesadas) tendem a atrair o acento. Pode-se perceber que tais vocábulos retêm o /S/ e podem, inclusive, reforçar a coda por meio de ditongação. Casos como *bônus*, *ônus*, *ânus*, *lâpis*, *íris*, que são exceções do ponto de vista acentual, terminam por sílaba átona, não protegida por acento, o que torna o segmento final mais suscetível ao cancelamento, como se verifica nos dois últimos vocábulos (que constam do questionário): é nessas sílabas que incide, no *corpus* de Almeida, o maior número de cancelamentos de /S/, respectivamente, 15 e 7. Os demais vocábulos inserem-se no conjunto cuja forma plural é mais marcada em relação à do singular, sendo natural a retenção do segmento. Lembre-se, ainda, que a adjunção de um morfema de número nominal (um elemento de natureza sufixal) não muda o padrão acentual do vocábulo, diferentemente do que pode ocorrer com sufixos (comparem-se *la'ranja* → *la'ranjas*; *la'ranja* → *laran'jal* → *laran'jada*).

Nº da carta	Vocábulo	Alveolar/Palatal/ Aspirada	Cancelamento	Total de ocorrências
21	Arco-íris	54	15	69
23	Arroz	70	1	71
80	Cruz	78	0	72
82	Depois	71	1	72
108	Feliz	70	2	72
130	Giz	71	0	71
141	Jesus	71	0	71
146	Lápis	65	7	72
154	Luz	72	0	72
174	Nariz	72	0	72
197	Paz	72	0	72
224	Pus	71	0	71
275	Voz	70	1	71
279	Dois	72	0	72
280	Três	72	0	72
283	Seis	72	6	72
287	Dez	71	1	72
Totais		1182 97%	34 3%	1216

Tabela 7. Ocorrências de /S/ em coda silábica externa sem valor de morfema: concretização *versus* cancelamento nos dados do MicroAFERJ. Fonte: Almeida (2008: 119)

Os dados que constam da tabela 8 foram obtidos por meio de questões que propiciam a formação de um sintagma composto por dois constituintes, o tipo de sintagma mais produtivo na modalidade oral. Na formulação da questão prevê-se, geralmente, um artigo ou um numeral (enunciado pelo documentador) e um substantivo (o núcleo, enunciado pelo informante), como, por exemplo, nas questões “*Os médicos nos hospitais cuidam dos _____ (doentes)*” - QFF 91 e “*Qual o móvel que fica na sala, é feito de espuma e que serve para duas ou três pessoas sentarem? E quando são dois, são dois ____ (sofás)*”- QFF 242 (Almeida, 2008: 121).

Mesmo havendo o estímulo do documentador, que já apresenta o primeiro constituinte no plural, grande parte dos informantes responde com o núcleo no singular, o que é uma prova cabal de que está seguindo a regra que prevê, no caso de indicação de pluralidade, o uso da marca preferencialmente no constituinte mais à esquerda. Está seguindo uma regra de motivação morfosintática, não propriamente fonético-fonológica, apesar de esses vocábulos se inserirem no conjunto em que a forma do plural em relação à do singular é a de mais baixa saliência. Não é de estranhar, portanto, tendo em vista uma outra variável fundamental (o nível de escolaridade dos informantes), ser alto o índice de cancelamento (50,15%), inclusive ligeiramente superior ao da concretização (49,85%).

Por outro lado, ainda com base no que se registra nas tabelas 7 e 8, se se considerar que, num universo de 72 informantes, por exemplo, 51 cancelaram a marca de número em *doentes*, 47 em *galhos*, 35 em *sofás* e apenas 7 cancelaram o /S/ em *nariz* e 15 em *íris*, pode-se deduzir que o falante que não aplica a marca de número em todos os constituintes do SN não cancela necessariamente a coda sem valor mórfico.

Nº da carta	Vocábulo	Alveolar/Palatal/Aspirada	Cancelamento	Total de ocorrências
67	Cócegas	4	60	64
90	Doentes	21	51	72
124	Galhos	6	47	72
140	Janelas	37	34	71
143	Lábios	53	18	71
182	Nuvens	16	56	72
185	Óculos ⁸	63	9	72
189	Olhos	66	5	71
243	Sofás	30	35	65
263	Unhas	35	37	72
Totais		350 49,85%	352 50,15%	702

Tabela 8. Ocorrências de /S/ em coda silábica externa com valor de morfema: concretização *versus* cancelamento nos dados do MicroAFERJ Fonte: Almeida (2008: 121)

O mesmo se verifica na Região Metropolitana do Rio de Janeiro, segundo os dados de Lima (2006). Como a autora utilizou o mesmo questionário que Almeida, apresentam-se, na tabela 9,

⁸ Como se trata de um caso de *pluralia tantum* e pelo fato de, na fala popular, ocorrer com frequência *meu óculo*, por exemplo, manteve-se o vocábulo nesta lista.

apenas os índices numéricos e percentuais para cada um dos contextos discriminados nas tabelas 6 a 8.

/S/ em coda interna				/S/ em coda externa sem valor mórfico				/S/ em coda externa com valor mórfico			
Concretiz.		Cancel.		Concretiz.		Cancel.		Concretiz.		Cancel.	
Nº	%	Nº	%	Nº	%	Nº	%	Nº	%	Nº	%
633	100	0	0	353	93	16	7	111	52,8	99	47,2

Tabela 9. Ocorrências de /S/ em coda silábica interna e externa sem e com valor de morfema: concretização *versus* cancelamento nos dados do AFeBG. Fonte: Lima (2006)

Aqui, também os índices demonstram a alta incidência de cancelamento (47,2%) quando o /S/ é marca de número. Esses dados, que dizem respeito a falantes de nível fundamental, moradores de áreas de diferente perfil – umas mais urbanas, outras, mais propriamente rurbanas – levam a formular a hipótese de que o cancelamento da marca de número e o cancelamento do segmento travador de sílaba possam estar sujeitos a diferentes restrições. Não há evidências que permitam afirmar que a tendência à simplificação da estrutura silábica esteja na base da redução flexional, como propõem Naro & Scherre, embora ambas as tendências se encontrem em variedades que se formaram em meio a forte contato multilinguístico.

4. À GUIA DE CONCLUSÃO

Neste estudo que se caracterizaria mais propriamente como um “exercício especulativo”, buscou-se trazer dados para debate sobre as motivações que estariam na gênese dos padrões variáveis de concordância que tipificam as variedades não europeias do Português aqui focalizadas. Para tanto, lançou-se mão, de um lado, de uma variável, que, no plano sincrônico, fornecesse indícios para testar a hipótese da interferência do contato multilinguístico na marcação defectiva de plural no SN, e, de outro, de dois trabalhos de cunho geo-sociolinguístico que permitissem testar a hipótese de que a tendência à simplificação da estrutura silábica – tida como fato de deriva – constituiria a motivação de base para a não implementação da marca de número em todos os constituintes do SN.

Com base em dados sincrônicos, verificou-se, de um lado, que a simplificação flexional parece estar fortemente vinculada ao contato multilinguístico; de outro, que o cancelamento do /S/ não morfêmico, não ocorre, mesmo entre indivíduos de baixa escolaridade, com a significativa frequência do cancelamento de /S/ marca de número nominal: tudo indica serem processos condicionados por diferentes restrições, embora, eventualmente, possam ocorrer na fala de um mesmo indivíduo.

O referido exercício carece de mais evidências, mas, em contrapartida, permite que se formulem, de forma indireta, outras hipóteses. Tendo em vista as diferentes variedades do Português atual (considerando a dicotomia variedade europeia x variedades não europeias), caberia verificar, por exemplo, no que tange à concordância e à simplificação da estrutura silábica, até que ponto seria válido falar em deriva. A deriva – a tendência – em português seria a que seguiu no sentido de preservar a coda e a marcação plena no SN, como se verifica no PE, ou a que redundou na simplificação, silábica e flexional, como se observa nas variedades não

européias? Caberia, ainda, verificar se falantes que cancelam o /S/ marca de número nominal cancelariam, na mesma proporção, o /S/ em coda quando este não tivesse valor morfêmico.

Pode ser que novos *corpora* sincrônicos (a serem constituídos em diferentes comunidades, como as da Ilha da Madeira, por exemplo), bem como pesquisas em tempo real de longa duração possam trazer novos e inesperados elementos que impliquem ajustar ou reformular as análises já realizadas. Por enquanto, aqui ficam algumas especulações a serem confirmadas ou contestadas por futuros desdobramentos da pesquisa.

REFERÊNCIAS BIBLIOGRÁFICAS

- Almeida, Fabiana da Silva Campos. 2008. *Micro Atlas Fonético do Estado do Rio de Janeiro: uma contribuição para o conhecimento dos falares fluminenses*. 2 v. Tese de Doutorado. Rio de Janeiro: Faculdade de Letras, Universidade Federal do Rio de Janeiro.
- Almeida, Letícia Barros de. 2007. A aquisição das codas mediais por uma criança bilíngue em português e francês. *Actas do XII Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa, APL: 73-85.
- Brandão, Sílvia Figueiredo. 2011a. Concordância nominal em duas variedades do português: convergências e divergências. *Veredas*, v. 15, n. 1: 164-178.
- Brandão, Sílvia Figueiredo. 2011b. Concordância nominal na variedade não *standard* do português falado em São Tomé. Comunicação apresentada ao XVI Congresso Internacional de ALFAL Alcalá de Henares, Espanha. Universidad de Alcalá, 6-9 jun.
- Brandão, Sílvia Figueiredo. 2013. Patterns of plural agreement within the Noun Phrase. *Journal of Portuguese Linguistics*, 12 (2): 51-100.
- Brandão, Sílvia Figueiredo. 2014. A múltipla face da concordância, em português, em A. M. Dios (ed.) *La lengua portuguesa: estudos lingüísticos*, Salamanca, Ediciones Universidad de Salamanca, v. II: 135-146.
- Brandão, Sílvia Figueiredo e Sílvia Rodrigues, Vieira. 2012a. Concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística. *PAPIA: Revista Brasileira de Estudos Crioulos e Similares*, v. 22, n. 2: 7-39.
- Brandão, Sílvia Figueiredo e Sílvia Rodrigues Vieira. 2012b. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa: Revista de Linguística*, v. 56, n. 3: 1035-1064.
- Capellari, Elaine T. C. e Anna M. S. Zilles. 2002. A marcação de plural na linguagem infantil - estudo longitudinal. *Revista da ABRALIN*, vol. 1, n. 1: 185-218.
- Castro, Ivo. 1991. *Curso de história da língua portuguesa*, Lisboa, Universidade Aberta.
- Guy, Gregory. 1981. *Linguistic variation in Brazilian Portuguese: aspects of the Phonology, Syntax, and Language History*. Philadelphia: University of Pennsylvania. PhD Dissertation.
- Guy, Gregory. 1989. On the nature and origins of vernacular Brazilian Portuguese. *Estudios sobre el Español de América y Linguística Afroamericana*. Bogotá: Instituto Caro y Cuervo: 226-244.
- Hagemeijer, Tjerk. 2009. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola 1*, 1: 1-27.
- Holm, John. 2001-20012. The social and linguistic origins of Brazilian Vernacular Portuguese: parallel cases of partial restructuring. *Revista Portuguesa de Filologia*, XXIV: 123-145.
- Jon-And, Ann. 2010. Concordância variável de número no SN no português L2 de Moçambique – algumas explicações sociais e linguísticas. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, n. 2: 28-50.
- Jon-And, Ann. 2011. *Variação, contato e mudança linguística em Moçambique e Cabo Verde. A concordância variável de número em sintagmas nominais do português*. 167 fls. Tese (Doutorado). Department of Spanish, Portuguese and Latin American Studies, Stockholm University, Stockholm.
- Labov, William. 1972. *Sociolinguistic patterns*, Philadelphia, University of Pennsylvania Press.
- Labov, William. 1994. *Principles of linguistic change: internal factors*, Oxford: Blackwell Publishers, v.1.
- Labov, William. 2001. *Principles of linguistic change: social factors*, Oxford, Blackwell Publishers, v. 2.
- Labov, William. 2003. Some sociolinguistic principles, em C. B Paulston e G. R. Tucker (eds) *Sociolinguistics: the essential reading*, Cambridge, Blackwell Publishing: 234-250.
- Lima, Luciana Gomes de. 2006. *Atlas Fonético do entorno da Baía de Guanabara-AFeBG*. Dissertação (Mestrado em Letras Vernáculas). Rio de Janeiro: Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2v.
- Lucchesi, Dante. 2003. O conceito de transmissão linguística irregular e o processo de formação do Português do Brasil, em C. Roncarati e J. Abraçado, J. (org.). *Português brasileiro: contato linguístico, heterogeneidade e história*, Rio de Janeiro, 7 Letras: 272-284.

- Lucchesi, Dante. 2012. A diferenciação da língua portuguesa no Brasil e o contato entre línguas. *Estudos de Linguística Galega*, 4: 45-65.
- Mezommo, Carolina Lisboa. 2004. Interação entre a aquisição da fonologia e da marca de plural no português brasileiro: o domínio de /s/ nas codas lexical e morfológica. *Anais do VI Encontro do Círculo de Estudos Linguísticos do Sul – CELSUL*, Florianópolis: CELSIL, 2004.
- Naro, Anthony Julius e Maria Marta Pereira Scherre. 2007. *Origens do português brasileiro*, São Paulo, Parábola Editorial.
- Nobre, Wagner Carvalho de Argolo. 2011. *Introdução à história das línguas gerais no Brasil: processos distintos de formação no período colonial*. 229 fls. Dissertação (Mestrado em Letras) – Instituto de Letras, Universidade Federal da Bahia, Salvador.
- Rodrigues, Aryon Dall’Igna. 1986. As línguas gerais em A. D. Rodrigues. *Línguas brasileiras: para o conhecimento das línguas indígenas*, São Paulo, Loyola: 93-109.
- Teyssier, Paul. 1997. *História da língua portuguesa*. Tradução Celso Cunha. São Paulo, Martins Fontes.
- Vieira, Silvia Rodrigues e Brandão, Silvia Figueiredo. 2014. Tipologia de regras linguísticas e estatuto das variedades/línguas: a concordância em português. *Linguística*, 30 (2), dezembro: 81-112.
- Weinreich, Uriel, Labov, William e Herzog Mikail. 1968. Empirical foundations for a theory of linguistic change, em W. Lehmann e Y. Malkiel (orgs.) *Directions for historical linguistics*, Austin, University of Texas Press: 97-195.