



VARIÁVEIS MORFOSSINTÁTICAS

CONCORDÂNCIA NOMINAL NO PORTUGUÊS DE SÃO TOMÉ E NO PORTUGUÊS DE MOÇAMBIQUE

SILVIA FIGUEIREDO BRANDÃO

1. INTRODUÇÃO: POR QUE FOCALIZAR VARIEDADES AFRICANAS DO PORTUGUÊS

A concordância nominal talvez seja, juntamente com a concordância verbal, um dos temas mais focalizados no âmbito do Português do Brasil (PB). Em todas as variedades, as rurais e as urbanas, nos diferentes pontos do país, ela constitui uma regra variável, com fortes implicações sociolinguísticas, uma vez que o cancelamento da marca de plural tem caráter estigmatizante.

A observação do fenômeno, embora de forma assistemática, remonta aos primeiros trabalhos dialectológicos, em que já se chamava a atenção para o fato de a marca de número se encontrar predominantemente no primeiro constituinte do sintagma nominal, em exemplos como “*Estas carta não são as minha*”, dado por Amaral (1976 [1920]), ou “*grandes coisa/coisas grande*”, fornecido por Marroquim (1945). A partir da década de 1970, com a difusão no Brasil da Sociolinguística Variacionista, o tema é retomado, gerando uma série de estudos como os de Braga; Scherre (1976), Braga (1977), Scherre (1978, 1988), Guy (1981), para só mencionar os pioneiros nessa linha. A esses trabalhos seguiram-se vários outros, como os realizados, individualmente ou em conjunto, por Naro e Scherre, e por diversos doutorandos e mestrandos.

Essas primeiras pesquisas acabaram por demonstrar que uma das mais marcantes diferenças entre o PB e o Português Europeu (PE) reside na marcação variável de plural no SN, que, juntamente com a observação de outras variáveis, levaram à formulação de hipóteses sobre as origens do PB. Dentre elas, as mais difundidas são as de Naro; Scherre (2003; 2007) que defendem a tese da deriva, isto é, de que padrões de concordância variável, por exemplo, já se encontrariam em diversas fases da história em Portugal, e a de Lucchesi (2003) de que tais diferenças resultariam de um processo de aprendizagem irregular derivado do contato plurilinguístico que caracterizou (e ainda caracteriza) a sociedade brasileira desde sua formação.

Com a continuidade das pesquisas sobre o PB e também sobre o PE, na perspectiva não só sociolinguística mas também gerativista, foi-se formando a ideia de que para bem conhecer o PB seria fundamental observar o que ocorre em outras variedades do Português, em especial nas africanas, em que essa língua coexiste com várias outras: com crioulos de base portuguesa, no caso de São Tomé, e com diferentes línguas do grupo Banto, em Moçambique.

Como já se frisou na apresentação deste livro, na atualidade, os trabalhos de Petter (2007, 2015) e de Avelar; Galves (2014), entre outros, sugerem a possibilidade de se traçar um *continuum* afro-brasileiro, baseado em fenômenos comuns presentes nas variedades não europeias. Petter, que trabalha sobretudo com a variedade angolana, observa:

as diferentes situações de contato, em épocas diversas, mas envolvendo o português e um conjunto de línguas muito próximas, as do grupo banto, produziram alguns resultados semelhantes nos níveis fonológico, lexical e morfossintático, que nos permitem defender a existência de um *continuum* afro-brasileiro de português. Mesmo considerando que não existam entidades homogêneas identificáveis como “português africano”, “português moçambicano”, “português angolano” ou “português brasileiro”, a história do contato e os aspectos linguísticos comuns a essas variedades autorizam-nos a levantar a hipótese desse *continuum*. (PETTER, 2007: 9).

No âmbito do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*, tem-se buscado testar essa hipótese, como também já se ressaltou, com base no princípio do Uniformitarismo, sugerido por Labov (1972) e segundo o qual forças que concorrem, no presente, para a variação e a mudança linguísticas são do mesmo tipo das que operaram no passado.

Convergências entre variedades não europeias do Português podem ser aqui-latadas, no que se refere aos padrões de concordância de plural no SN, no quadro a seguir, em que se apresentam dados selecionados de *corpora* representativos do

Português do Brasil, de São Tomé, de Moçambique e de Angola, devendo-se lembrar que com eles coexistem SNs com todas as marcas de plural.

Quadro 1 Exemplos de padrões variáveis de concordância em quatro variedades do Português

Português do Brasil	Português de São Tomé
(i) [as <i>menina</i>] tiraram retrato	(i) [as <i>dificuldade</i>] que nós temos
(ii) tem [esses transportes <i>alternativo</i>]	(ii) passar [os tempos <i>livre</i>] em casa
(iii) com [as <i>minha</i> netinha]	(iii) n[os primeiro momento]
(iv) [os filhos tudo <i>formado</i>]	(iv) [as pessoas mais <i>velha</i>]
Português de Moçambique	Português de Angola
(i) [essas pessoa] trocam de zona	(ii) [as coisa] estão muito cara
(ii) não há [campos suficiente]	(ii) e buscar [as tuas fruta]
(iii) com [esse conflitos]	(iii) [os tais português], fomos [...]
(iv) [Tudo meus boi]	(iv) se [meus cliente] quer um pão, eu fia
Os exemplos relativos ao Português do Brasil e ao de Moçambique foram selecionados do <i>Corpus</i> do Projeto <i>Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português</i> (UFRJ) e os referentes ao Português de São Tomé e ao de Angola, do Projeto VAPOR do Centro de Linguística da Universidade de Lisboa.	

Embora as histórias sociolinguísticas desses países apresentem traços particulares (cf. capítulos 2 e 3 deste livro), os exemplos expostos no Quadro 1 permitem formular a hipótese de que motivações comuns, alicerçadas em questões relacionadas ao contato multilinguístico, sugerem a existência do referido *continuum*, como têm procurado mostrar os estudos de Brandão (2011a, 2011b, 2011c, 2013); Brandão; Vieira (2012a, 2012b, 2016); Vieira; Brandão (2014, 2015).

A análise aqui empreendida desdobra-se por mais oito seções. Na segunda seção, comentam-se alguns estudos sobre concordância nominal em quatro variedades africanas do Português e, na terceira, apresentam-se os objetivos e os procedimentos metodológicos que fundamentaram as pesquisas sobre o PST e o PM. Na seção 4, há uma breve apresentação do funcionamento dessa regra no PE, seguida das análises referentes ao PST (seção 5) e ao PM (seção 6). Na seção 7, expõem-se, brevemente, os resultados referentes às duas variáveis mais salientes para o cancelamento da marca no PB. A oitava seção está dedicada à síntese dos resultados e a nona, a breves considerações finais.

2. ALGUNS ESTUDOS SOBRE CONCORDÂNCIA NOMINAL EM VARIEDADES AFRICANAS DO PORTUGUÊS

Embora alguns linguistas tenham feito comentários sobre o tema, como, por exemplo, Gärtner (1996) e Marques (1983) em relação, respectivamente, a Moçambique e Angola, até há pouco, raros eram os estudos específicos sobre a concordância nominal em variedades africanas. Para se aquilatarem os resultados oriundos desses trabalhos, comentam-se aqui alguns deles.

(a) *São Tomé e Príncipe*

No que tange a São Tomé, tem-se notícia das análises que foram desenvolvidas por Baxter (2004, 2009) e por Figueiredo (2008, 2010).

Baxter (2009) compara os dialetos da comunidade dos Tongas à dos afrodescendentes de Helvécia, na Bahia, motivado por “uma série de fatos linguísticos e sociolinguísticos compartilhados” (p. 270-271) por ambas as variedades, analisando os dados de cada localidade por faixas etárias (20 a 40; 41 a 60; 61 a 80; 85 a 103 anos) e globalmente. Em todas as análises, a variável *posição do item em relação ao núcleo do SN* foi considerada a mais relevante, indicando que a presença da marca de número depende de uma “forte correspondência entre posição linear e classe gramatical”. Ele afirma que a posição pré-nuclear adjacente ao núcleo tem um papel chave como “âncora” para a introdução da marca de PL, o que se observa em todas as faixas etárias, mas está mais evidente na fala de indivíduos mais velhos. Com base nesses resultados, propõe um quadro (p. 293) evolutivo da atribuição de plural no SN.

Figueiredo (2010) analisa a aquisição de marca de plural no SN em uma variedade reestruturada do Português falada pela comunidade bilingue (crioulo e português) de Almoxarife, em São Tomé. Sua análise, bastante minuciosa, pautou-se pelos pressupostos da Sociolinguística Variacionista e baseou-se em 2.340 dados, dos quais 1.202 com a presença da marca (51%) e 1.138 (49%) sem ela. O autor observou, também, os sintagmas como um todo, concluindo que apenas 425 dos 1.488 SNs (29%) são produzidos com todas as marcas. Foram controladas 11 variáveis estruturais (três fonológicas, sete morfossintáticas e uma lexical) e quatro sociais (faixa etária, sexo, escolaridade e estadia – permanência fora da comunidade), algumas delas, por vezes, a outras amalgamadas.

A seguir, reproduz-se a tabela em que se expõem os resultados de Figueiredo referentes à atuação da posição linear do constituinte em relação ao núcleo na aplicação da marca de número. Essa tabela retrata, na fala da comunidade como um todo, o efeito dessa variável que seria, depois, controlada por faixas etárias (20-40 anos; 41-60 anos; mais de 60 anos) para verificar se, a exemplo do que mostrou o estudo de Baxter (2009), também existiriam gramáticas distintas para as diferentes gerações na comunidade de Almoxarife.

Tabela 1 Atuação da variável *posição linear e posição em relação ao núcleo do SN* na fala da comunidade de Almojarife, São Tomé

Efeito da relação entre <i>posição linear e posição em relação ao núcleo do SN</i> na marcação PL dos itens do SN do PA.				
<i>INPUT DESTA RODADA: 0,527</i>	<i>LOG-LIKELIHOOD: -952,926</i>		<i>SIGNIFICÂNCIA: 0,028</i>	
FACTORES SELECIONADOS	NR TOTAL	% MARCAÇÃO PL	% NO GRUPO	PESO RELATIVO
Anterior ao núcleo, não imediatamente	57/87	66	4	0,565
Imediatamente anterior ao núcleo	631/681	93	29	0,903
Núcleo na 1ª posição do SN	(25/29)	86	1	0,678
Núcleo na 2ª posição do SN	399/1248	32	53	0,300
Núcleo na 3ª posição do SN	41/117	35	5	0,210
Núcleo na 4ª ou outra posição do SN	32/88	36	4	0,213
Posterior ao núcleo (2ª, 3ª, 4ª ou outra posição)	17/90	19	4	0,069
Totais	1.202/2.340	51	100	–

Fonte: Figueiredo (2010), Tabela 21.7, p. 510

Sobre esse procedimento Figueiredo conclui que:

os resultados evidenciam dois aspectos a ter em conta: (i) a existência de gramáticas distintas nas diferentes gerações de Almojarife; (ii) perfis de marcação geracional muito semelhantes entre o PA [Português de Almojarife], o PT [Português dos Tongas] e HEL-Ba [Helvécia-Bahia] (...) Assim, as distintas gramáticas geracionais poderão ser observadas, por exemplo, pelo comportamento do factor *posição anterior ao núcleo do SN, mas não imediatamente*, que caminha no sentido da aquisição da concordância, pois começa por ser um elemento com escassas ocorrências na FE-3 [Faixa Etária 3] (apenas 19), para passar, depois, a ser mais usado na faixa etária intermédia, mas com um peso inibidor da marcação (pr. 0,365), e acabar por se fixar, na geração mais nova, como item que favorece moderadamente a marcação (pr. 0,685). Este aspecto, para além de revelar um padrão de desenvolvimento geracional oposto ao da deriva românica, evidencia também a preferência da geração mais idosa do PA pela realização de SN's que envolvem apenas uma posição pré-nuclear, isto é, de estrutura reduzida, um aspecto que encontra paralelo nas outras duas comunidades observadas. (FIGUEIREDO, 2010: 514)

(b) Moçambique

Sobre o PM, há os estudos desenvolvidos no âmbito da Universidade Eduardo Mondlane – UEM (cf. GONÇALVES, 2013, 2010; GONÇALVES; STROUD, 1997, 1998; GONÇALVES et al., 1998), com base no *corpus* PPOM (Panorama do Português Oral de Maputo) e, ainda, os de Jon-And (2010, 2011).

Moreno; Tuzine (1997), membros da referida equipe, realizaram estudo que, embora definido como sociolinguístico variacionista, só levou em conta as estruturas desviantes da norma padrão (as do PE). Na análise, compararam-se as ocorrências “efetivamente” observadas com “as que seriam de esperar”, com base em *corpus* selecionado da fala de 20 indivíduos que têm o Português como L2, distribuídos segundo a escolaridade, a idade, a profissão e o local de residência.

Os resultados, ancorados apenas em índices percentuais, demonstraram, em relação à concordância nominal, que os desvios à norma do PE se encontram entre os indivíduos com menor grau de instrução (76,7%) e mais de 45 anos (53,6%), cuja profissão se enquadra na categoria classificada como baixa (83,9%). Quanto à variável local de residência, os autores afirmam que não foi possível identificar fenômenos linguísticos característicos dos bairros que permitissem estabelecer uma oposição entre urbano e suburbano.

Jon-And, em sua tese de doutoramento (2010), focaliza a concordância confrontando duas variedades do Português L2: a de Moçambique, em geral L2 de falantes de línguas Banto e a de Cabo Verde, L2 de falantes de um crioulo de base lexical portuguesa sem substrato Banto. Os informantes tinham de 3 a 7 anos de escolaridade, motivo pelo qual não foi considerada na análise a variável nível de escolaridade, mas foi controlada a idade de aquisição do Português.

No resumo dos resultados da análise de 2875 dados do Português de Moçambique, a autora indica como favorecedores da presença da marca de plural (87%) – a variante tomada como valor de aplicação – a posição em relação ao núcleo/posição linear, a faixa etária do indivíduo, a idade de início de aquisição do português, a saliência fônica, marcas precedentes e o contexto fonológico posterior.

Em relação à variável estrutural mais relevante, verifica-se que elementos antecedendo o núcleo levam em maior grau marca de plural que elementos nucleares e pós-nucleares e que o núcleo em primeira posição apresenta, categoricamente, a marca de plural. Ao comentar esses resultados, Jon-And reporta-se à hipótese de Baxter (2004) de que a tendência a marcar mais o plural nos constituintes que antecedem imediatamente o núcleo do que nos que o antecedem não imediatamente seria influência das línguas Banto, que marcam o plural por meio de prefixos. Para a autora, embora todos os informantes levados em conta na pesquisa tenham uma língua Banto como L1, a explicação de Baxter, que “pare-

ceria à primeira vista interessante para os dados de Moçambique” (p. 117), não se aplica a essa variedade em razão de três dos resultados de seu estudo:

1. Nos dados de Moçambique, a diferença entre os elementos pré-nucleares adjacentes e não adjacentes ao núcleo não se mostrou estatisticamente significativa nos SNs em que o núcleo é precedido por dois elementos.
2. O grupo de fatores *classe nominal do núcleo em ronga* não tem nenhuma influência sobre a marcação de plural no material de Moçambique, o que seria esperado, se a hipótese de uma influência de línguas bantu fosse verdadeira.
3. A análise do material de Cabo Verde mostra efeitos do grupo de fatores *posição em relação ao núcleo/posição linear* extremamente semelhantes àqueles encontrados no material de Moçambique. O PCV [Português de Cabo Verde] não pode ter sido influenciado por línguas bantu, e é, portanto, preciso procurar uma outra explicação para a tendência de marcar plural em primeiro lugar no elemento que antecede imediatamente o núcleo (JON-AND, 2010: 117)

Quanto às duas variáveis sociais mais relevantes, ela conclui que

Os falantes que têm mais que sessenta anos marcam plural consideravelmente menos que os outros falantes. Isto indica uma tendência diacrónica para marcar o plural cada vez mais, o que pode ser explicado pelo aumento no uso do português na vida quotidiana de Maputo e pela consequente diminuição do uso de línguas bantu na mesma. Quanto à idade de início de aquisição do português, há uma tendência forte para as pessoas, que começaram a aprender português mais cedo na vida, marcarem plural em maior grau que as pessoas, que começaram mais tarde. A análise desses dois grupos de fatores extralinguísticos indica que a concordância variável de número em Moçambique é um resultado de contato linguístico (JON-AND, 2010: 116-117).

(c) *Cabo Verde*

Embora, neste capítulo, não se analise o Português de Cabo Verde (PCV) nem o de Angola (PA), julga-se importante, para fins comparativos, mencionar também o que se tem observado sobre a concordância nominal nessas duas áreas.

Jon-And (2010), no estudo recém-comentado, analisou 2359 dados referentes ao PCV. Na síntese dos resultados, ela indica sete variáveis como as mais relevantes para a aplicação da marca de plural, que foi da ordem de 82%: *posição em relação ao núcleo/posição linear; escolaridade/idade de início de aquisição do português; formação de plural; animacidade; flexionabilidade do correspondente lexical no crioulo cabo-verdiano; função sintática do SN e marcas precedentes.*

Os resultados referentes à primeira e mais significativa variável estrutural foram semelhantes aos obtidos para Moçambique, à exceção do que respeita aos elementos pré-nucleares. No PCV, os constituintes que antecedem imediatamente o núcleo levam marca de plural em maior grau que os constituintes que antecedem o núcleo não imediatamente. Jon-And novamente refere-se à explicação de Baxter (2004) para São Tomé e Helvécia, desta vez dizendo que ela, neste caso, também não se aplicaria, uma vez que, em Cabo Verde, nunca houve línguas Banto. Ela complementa, afirmando que “parece também difícil encontrar uma explicação para essa estrutura nos substratos do crioulo cabo-verdiano. É preciso, portanto, procurar uma explicação mais geral, provavelmente relacionada com o contato linguístico, mas sem relação com a transferência direta de estruturas de línguas particulares” (p. 137).

Quanto à variável amalgamada *escolaridadelidade de início de aquisição do português*, ela demonstra o que era de esperar: um nível de escolaridade mais baixo não favorece a marcação de plural, enquanto um nível de escolaridade mais alto a favorece. No entanto, em Cabo Verde, a maioria dos falantes só começa a aprender e só usa Português na escola, o que significa:

que um nível mais alto de escolaridade implica um processo de ASL [Aquisição de Segunda Língua] mais prolongado. O resultado obtido na análise da escolaridade indica, portanto, uma influência de ASL na marcação de plural. A *idade de início de aquisição do português* também tem um efeito sobre a marcação de plural, mas este efeito é subordinado ao efeito da escolaridade, razão pela qual os dois grupos foram cruzados para permitir uma leitura mais clara. O efeito da *idade de início de aquisição do português* sobre a marcação de plural é bastante leve, mas indica também o padrão esperado: Os falantes que começam a aprender português numa idade mais baixa marcam plural em maior grau que os falantes que começam a aprender português numa idade mais avançada. Este resultado indica também o papel importante da ASL para o nível de marcação de plural no PCV. O contato linguístico parece, portanto, influenciar a concordância variável de número no SN no PCV, assim como no PM (JON-AND, 2010: 136).

Cabe referir, ainda, a variável *flexionabilidade do item lexical correspondente no crioulo cabo-verdiano*, cujos resultados mostram a possibilidade “de transferência de estruturas de marcação de plural da L1 dos falantes” para o PCV, isto porque “os elementos que têm correspondentes lexicais com formas distintas para o singular e o plural no crioulo cabo-verdiano tendem a levar a marca de plural no PCV em maior grau que itens lexicais que têm correspondentes sem flexão de número no crioulo cabo-verdiano” (JON-AND, 2010: 137).

Jon-And, em suas conclusões, discute as diferentes interpretações postuladas para a origem da concordância variável de número plural à luz dos resultados por ela obtidos para o PM e o PCV.

(d) *Angola*

No que se refere ao PA vernacular, Inverno (2009) demonstra que, no SN, o núcleo raramente recebe a marca de número, estando esta presente no(s) seu(s) constituinte(s) não nuclear(es) mais à esquerda, o que se verifica principalmente na fala dos indivíduos mais velhos ou menos escolarizados ou no registro informal dos mais jovens ou mais instruídos. A hipótese de que tal tipo de marcação parece estar relacionada com o número de elementos no SN, segundo a autora, poderia ser reforçado pelo fato de o –s desaparecer em núcleos nos quais não desempenha a função de marcador, como no exemplo “*quando acabar as férias*”.

Para a explicação do fenômeno, Inverno, após algumas observações, acaba por adotar a proposta de Marques (1983), para quem o padrão acima descrito se deve à interferência das línguas Banto, que marcam a pluralidade por meio de prefixos (e não de sufixos) em todos os constituintes do SN. Segundo Marques (apud INVERNO, 2009), os angolanos teriam interpretado os nomes portugueses como invariáveis e os artigos como elementos equivalentes aos prefixos banto, tendo em vista que nenhuma mudança ocorre no início das palavras.

Inverno (p. 95-96), ao observar exemplos de SNs de mais de dois constituintes (ex.: “*os meus passatempo*”), que não se enquadrariam perfeitamente na interpretação de Marques, lembra que “são várias as classes nominais nas línguas banto, essencialmente as classes que incluem nomes que designam entidades abstractas, diminutivos, aumentativos, etc., em que a marcação de número é feita apenas nos elementos não nucleares do SN, mas não no núcleo”. Segundo ela, de fato “os elementos mais à esquerda no SN tendem a atrair a marcação de número, independentemente de se tratar do primeiro elemento de um nome composto, como em (10) [guardas-chuva] ou o núcleo do SN, como em (11) [eles próprio]” ou, ainda, em casos que ela caracteriza como excepcionais (“muitas línguas materna”).

3. OBJETIVOS E ORIENTAÇÃO TEÓRICO-METODOLÓGICA

Tem-se por objetivo focalizar, à luz dos pressupostos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968), a concordância de plural no SN nas variedades urbanas do Português de São Tomé (PST) e do Português de Moçambique (PM) com base nos *corpora* citados na Apresentação deste livro.

Para a consecução das análises, levadas a cabo com o auxílio do Programa Goldvarb-X, consideraram-se todas as ocorrências de SNs de um total de 27 entrevistas representativas do PST e 18, do PM. Os informantes foram distribuídos por sexo, três faixas etárias (18-35 anos, 36-55 anos e 56 a 75 anos) e três níveis de escolaridade (fundamental: 5 a 8 anos; médio: 9 a 11 anos; superior). Além

dessas variáveis sociais, foram consideradas, com base nas declarações dos próprios informantes (a) para o PST, a variável *frequência de uso do Forro*, com o objetivo de aquilatar se o não uso do Forro ou seu uso em maior ou menor grau influenciaria a performance dos indivíduos; (b) para o PM, o estatuto do Português (L1 ou L2) bem como o grau de conhecimento/uso de outra(s) língua(s) falada(s) em Moçambique.

Quanto às variáveis estruturais, controlaram-se (i) a saliência fônica, (ii) a tonicidade e (iii) o número de sílabas do item singular, (iv) a classe do vocábulo; (v) a posição linear e relativa do constituinte no SN; (vi) as marcas precedentes; (vii) o contexto fonológico subsequente e (viii) a animacidade do núcleo do SN.

Os resultados relativos a São Tomé foram selecionados dos trabalhos referidos no final do item 1; os relativos a Moçambique são, ainda, inéditos.

4. A NORMA DE REFERÊNCIA DAS VARIEDADES AFRICANAS: A CONCORDÂNCIA NOMINAL NO PE

Não obstante o grande número de línguas que coexistem nas áreas aqui focalizadas, só o Português tem estatuto de língua oficial, sendo, portanto, adotada nas escolas, nos meios de comunicação e nos documentos oficiais. Nem em São Tomé, em que o Forro ainda conta com um significativo número de falantes, nem em Moçambique, onde, nas províncias, dominam diferentes línguas Banto (algumas delas já descritas), são ainda em pequeno número as iniciativas no sentido de implementar o ensino bilingue de forma regular. Desse modo, não é de estranhar que o Português, considerada a língua de prestígio, venha, cada vez mais, se difundindo como a L1 ou a L2 da população, sobretudo nos centros urbanos.

Nesse sentido, ao que tudo indica, a norma de referência para o ensino nas escolas é a descrita nos manuais portugueses, cabendo lembrar que falantes santomenses e moçambicanos com nível superior e com mais de 50 anos, em geral, demonstram um desempenho linguístico similar ao de portugueses por terem feito sua formação em Portugal.

A análise não atomística realizada por Brandão (2013) em amostra do PE registrou 99,89% de marcação de plural. Apenas oito SNs não apresentaram marca de número em um de seus constituintes num total 6.952 SNs selecionados da fala de 54 indivíduos naturais de duas localidades da Região Metropolitana de Lisboa (Oeiras e Cacém) e de Funchal, na Ilha da Madeira, distribuídos segundo os parâmetros de sexo, faixa etária e nível de escolaridade acima indicados. A autora (p. 58-59) observa que a maior parte desses 0,11% de dados de cancelamento resultam de hesitação ou de alguma falha discursiva geralmente retificada pelo falante, como se pode perceber, respectivamente, em (1) e (2).

- (1) nada...ali nada tem **uma** pracetazinha... há **uma** ou [**duas pracet**a] onde os miúdos se entretêm...(CAC-C1h)
- (2) a maioria d[as pessoa] **das mulheres** era tra/ era em casa era a border (FNC-B1m)

Os resultados dessa amostra sugerem que, na variedade urbana do PE, a regra de concordância no SN é categórica, nos termos de Labov (2003), isto é, todos os constituintes flexionáveis do SN são marcados no plural, não havendo, portanto, variação sujeita a restrições de ordem social ou estrutural.

5. A VARIEDADE URBANA DO PORTUGUÊS DE SÃO TOMÉ

Os primeiros estudos sistemáticos sobre a concordância nominal na variedade urbana do PST foram realizados por Brandão (2011a, 2011b, 2013), com base em duas amostras, a primeira, aqui denominada Amostra 1, com 2612 dados referentes à fala de 17 indivíduos estratificados por sexo, faixa etária e os três níveis de escolaridade (2011a), a segunda com 2.375 dados selecionados da fala de 22 indivíduos de níveis fundamental e médio (2011b e 2013) – a Amostra 2.

Na Amostra 1, obtiveram-se 2439 ocorrências de aplicação da marca (93,4%) e apenas 173 (6,6%) de cancelamento. Os resultados indicaram o *nível de escolaridade* e a *posição linear e relativa dos constituintes no SN* como as duas variáveis mais relevantes para o cancelamento ou não da marca de número, sendo, ainda, selecionados a *animacidade do núcleo*, o *sexo* e o *contexto fonológico subsequente*.

Dos resultados dessa primeira análise interessa destacar a atuação da variável *nível de escolaridade* (Tabela 2), que demonstra que quanto menos escolarizado o indivíduo, mais ele tende a omitir o morfema de plural.

Tabela 2 Atuação da variável *nível de escolaridade* para o cancelamento da marca de número no SN (Amostra 1)

FATORES	APL./OCO	%	P. R.
1-5 a 8 anos (fundamental)	125/531	23,5	.91
2-9 a 11 anos (médio)	34/942	3,6	.51
3-12 a 15 anos (superior)	14/1139	1,2	.24
Input: .016		Significância: .023	

Para melhor aferir esses dados, procedeu-se ao cruzamento dessa variável com faixa etária, conforme se expõe na Figura 1, em que se observa que são os indivíduos de nível médio e os de nível fundamental os que apresentam maior grau de variação.

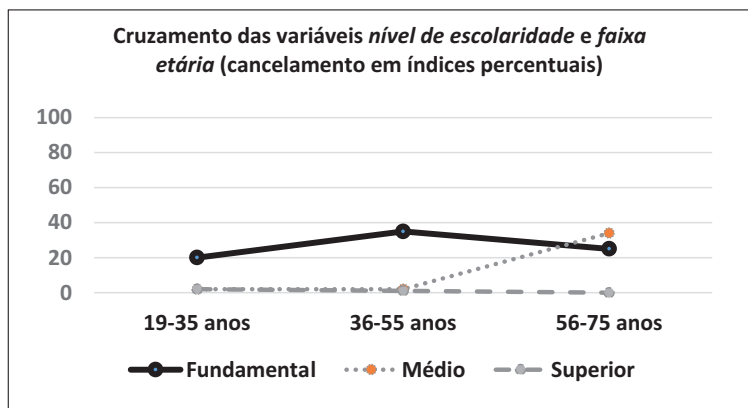


Figura 1 Cruzamento das variáveis *nível de escolaridade e faixa etária* (cancelamento em índices percentuais) no PST (Amostra 1)

Fonte: Brandão, 2011b

Sobre esse gráfico, Brandão (2011b: 38) faz as seguintes generalizações:

- nas diferentes faixas etárias, os indivíduos de nível superior de instrução apresentam os menores índices de cancelamento, 2%, 1% e 0%, respectivamente, nas faixas A, B e C, o que sugere que, nesse grupo, a regra de concordância nominal tem caráter semicategórico;
- entre os indivíduos de nível médio de instrução, os mais velhos apresentam índice de cancelamento (34%) contrastante com o dos jovens (faixas A e B), que não ultrapassa os 2%, o que aproxima o desempenho desse grupo da performance dos mais escolarizados;
- entre os menos escolarizados, são os falantes entre 36 e 55 anos os que apresentam os maiores índices de cancelamento (35%), ficando as faixas extremas com percentuais bem próximos (A: 20%, C: 25%);
- as curvas referentes aos indivíduos de nível fundamental e médio sugerem um movimento no sentido de incorporação da norma padrão de concordância.

Os altos índices de aplicação da marca de número entre os falantes de nível superior, certamente mais afinados com o PE, demonstraram que a regra de concordância tem, na norma desse segmento social, caráter semicategórico (Labov, 2003), o que levou à constituição da mencionada Amostra 2. Dessa amostra constam apenas dados selecionados da fala de indivíduos de níveis fundamental e médio

– os 12 informantes da Amostra 1 e mais 10 novos informantes desses níveis. (BRANDÃO, 2011c, em que se incluiu a variável *frequência de uso de um crioulo*)

Na Amostra 2, das 2.375 ocorrências, 305 correspondem à ausência de marca de número (12,8%) e 2070 (87,2%) à sua presença. Embora haja um aumento percentual não tão significativo do cancelamento nesta amostra (12,8%) em relação à anterior (6,6%), um maior número de dados permite melhor aquilatar as restrições que presidem à variação e, por outro lado, confirmar a hipótese de que, em variedades urbanas, há, em geral, maior tendência à concordância.

Cinco variáveis mostraram-se relevantes, duas de cunho social e três de natureza estrutural: *posição linear e relativa dos constituintes no SN*, *nível de escolaridade*, *frequência de uso de um crioulo*, *saliência fônica e animacidade* (significância .000), sendo .04 o *input* da regra de cancelamento. Na Tabela 3, expõem-se os resultados referentes à primeira dessas variáveis.

Tabela 3 Atuação da variável *posição linear e relativa do constituinte no SN* para o cancelamento da marca de número (Amostra 2)

POSIÇÃO LINEAR E RELATIVA DO CONSTITUINTE NO SN		APL/OCO	%	P. R.
PRÉ-NUCLEAR	1ª posição [<i>muita histórias</i>]	9/909	1	.14
	2ª/3ª posições [<i>todos esse livros</i>] [<i>os outros meus irmãos</i>]	6/104	5.8	.55
NUCLEAR	1ª posição [<i>curso noturnos de alfabetização</i>]	2/45	4.4	.36
	2ª posição [<i>as dificuldade</i>] com[<i>oito ano</i>]	228/1092	20.9	.75
	3ª/4ª posições [<i>umas pequena venda</i>] [<i>os seus quatro ano de idade</i>]	25/122	20.5	.76
PÓS-NUCLEAR	2ª posição [<i>trabalhos específico</i>]	11/44	25	.88
	3ª posição [<i>os tempos livre</i>]	16/44	36.4	.93
	4ª/5ª posições [<i>as pessoas mais velha</i>] n[<i>essas zonas assim mais distante</i>]	8/15	53.3	.95

A despeito de ser pouco significativo o grau de variação como um todo, observa-se que a primeira posição é a menos sujeita à não marcação (quando pré-nuclear: P. R.: .14; quando nuclear: P. R.: .36).

Em cada conjunto de constituintes, as primeiras posições são sempre mais suscetíveis às marcas. No âmbito dos elementos pré-nucleares, a segunda posição (P. R.: .55) apresenta mais cancelamento do que a primeira (P. R.: .14). Na área do núcleo, a segunda posição, a mais frequente tendo em vista que o *corpus* é constituído basicamente de SNs de dois constituintes, apresenta alta probabilidade de cancelamento (P. R.: .75) se comparada aos elementos pré-nucleares, o mesmo ocorrendo na terceira/quarta posição (P. R.: .76). Quanto aos elementos pós-nucleares, à medida que se afastam do núcleo, mais sujeitos estão a não serem marcados (segunda posição – adjacente ao núcleo – P. R.: .88; terceira: P. R.: .93 e quarta/quinta: P. R.: .95).

Tal configuração sugere a existência de *continua* intra e interconstituintes, a propósito dos quais se tecerão considerações na seção 8.

O *nível de escolaridade* (Tabela 4) e a *frequência de uso de um crioulo* (Tabela 5) são variáveis, de certo modo, interrelacionadas.

Tabela 4 Atuação da variável *nível de escolaridade* para o cancelamento da marca de número no SN (Amostra 2)

NÍVEL DE ESCOLARIDADE	APL./OCO	%	P. R.
Fundamental	252/930	27.1	.79
Médio	53/1445	3.7	.29

Fonte: Brandão, 2011c

De acordo com uma das hipóteses formuladas, os indivíduos de nível fundamental (P. R.: .79) apresentam maior tendência ao cancelamento do que os de nível médio (P. R.: .29). Observe-se que, a despeito do aumento do número de informantes, os percentuais permaneceram nos mesmos patamares da Amostra 1 (*fundamental*: Amostra 1: 23,5%, Amostra 2: 27,1%; *médio*: Amostra 1: 3,6%, Amostra 2: 3,7%).

Os pesos relativos, no entanto, sofreram uma redefinição (*fundamental*: Amostra 1: P. R.: .91, Amostra 2: P. R.: .79; *médio*: Amostra 1: P. R.: .51; Amostra 2: P. R.: .29), observando-se que o *range*, isto é, a diferença entre os dois níveis, passou de .27 para .50, o que sugere que indivíduos de nível médio têm uma performance mais próxima dos indivíduos de nível superior do que dos de nível fundamental.

Em geral, os indivíduos mais escolarizados são os que menos se utilizam do crioulo. Alguns declaram nunca usá-lo, certamente por terem consciência da sua valoração negativa, estigmatizante ou, ainda, como constantemente relatam, pelo fato de seus pais obrigá-los a falar, mesmo em casa, apenas o Português, para que, dominando as normas de prestígio, se tornem aptos a aproveitar as poucas oportunidades de ascensão socioeconômica.

Embora todos os informantes que concorreram para a formação da amostra tenham se declarado falantes de Português como L1, tem-se de levar em conta o caráter multilíngue da sociedade santomense. Os diferentes informantes retratam situações prototípicas de São Tomé: há indivíduos que (a) se expressam só ou fundamentalmente em Português; (b) se intercomunicam em Português, mas dominam um crioulo e dele fazem uso eventualmente; (c) falam Português, mas se expressam, regularmente, num crioulo. No primeiro caso (a), inscrevem-se nove informantes; no segundo (b), dez e, no terceiro (c), três.

Tabela 5 Atuação da variável *frequência de uso de um crioulo* para o cancelamento da marca de número no SN (Amostra 2)

FREQÜÊNCIA DE USO DE UM CRIOULO	APL./OCO	%.	P. R.
Zero/Baixa	75/1039	7.2	.41
Média	140/1117	12	.52
Alta	90/219	41.1	.75

Fonte: Brandão (2011c)

Independentemente da distribuição assimétrica, verifica-se (Tabela 5) que são os indivíduos que se comunicam preferencialmente num crioulo os que mais tendem a não implementar a marca de número (P. R.: .75). Os três 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.: .41) e os que o utilizam apenas eventualmente (P. R.: .52). Observe-se, ainda, que o cancelamento também incide de forma escalar a depender da frequência de uso do crioulo: alta > .75; média > .52; baixa/zero > .41.

A exemplo da *posição linear e relativa do constituinte no SN*, as duas outras variáveis estruturais selecionadas – *saliência fônica* e *animacidade do núcleo* – também se mostram relevantes em estudos sobre a concordância nominal no PB e no PM.

No que se refere à saliência fônica, adota-se a premissa de que as formas menos marcadas, isto é, aquelas em que a diferença fônica entre as formas do singular e do plural reside apenas na presença do morfema de número – como em *filho/filhos* – seriam mais suscetíveis a não apresentarem o morfema de número do que as mais marcadas – como *olho/olhos* (com abertura da vogal e -S).

Para o controle desta variável, levaram-se em conta os 1400 nomes (substantivos e adjetivos) constantes da amostra, tendo-se optado por contrapor as formas fonicamente menos marcadas às mais marcadas, comprovando-se, mais uma vez, a hipótese de que o cancelamento seria mais visível entre os nomes de menor saliência fônica (.54).

Tabela 6 Atuação da variável *saliência fônica* para o cancelamento da marca de número no SN (Amostra 2)

SALIÊNCIA FÔNICA	APL./OCO	%	P. R.
Menor saliência fônica <i>Nomes terminados em vogal oral ou nasal</i> <i>filho-filhos/homem/homens</i>	268/1.182	22.7	.54
Maior saliência fônica <i>Nomes terminados em -R, S, L e -ÃO e com marca dupla</i> <i>professor/professores</i> <i>mês-meses/atual-atuais</i> <i>razão-razões/pão-pães</i> <i>[o]lho-[o]lhos</i>	21/218	9.6	.25

Fonte: Brandão (2011c)

Quanto à *animacidade*, de acordo com Scherre (1988), núcleos constituídos por nomes que apresentam o traço [+humano] seriam mais suscetíveis à marcação do que os demais. Figueiredo (2010), que cita estudos (p. 371) no âmbito de pidgins e crioulos em que isso se comprova, também considerou esse grupo na sua análise da comunidade de Almojarife, em São Tomé, estabelecendo três fatores, os mesmos aqui inicialmente controlados: [+humano]; [-humano] [+animado] e [-animado], obtendo, respectivamente, índices de aplicação do morfema de 60%, 37% e 48%, o que comprova a hipótese.

Na amostra 2, houve apenas sete ocorrências de núcleos com os traços [-humano] [+animado]: *cabra, tubarão, lula* (3), *peixe, bicho*, em que cinco não apre-

sentaram marca. Diante do pequeno número de vocábulos com esse traço, optou-se por opor [+animado] a [-animado]. (cf. Tabela 7). Assim, dos 433 nomes referentes ao primeiro fator, 426 têm traço [+humano].

Tabela 7 Atuação da variável *animacidade do núcleo* para o cancelamento da marca de número no SN (Amostra 2)

ANIMACIDADE DO NÚCLEO	APL./OCO	%	P. R.
[+ animado] manter [os <i>filho</i>] na escola	59/433	13.6	.39
[-animado] comprar [uns <i>medicamento</i>]	195/826	23.6	.55

Fonte: Brandão, 2011c

Como se pode verificar, encontram-se, no PST, padrões variáveis de concordância: em um extremo, aqueles com marcação de plural em todos os constituintes, em outro, aqueles com a presença de marca apenas em um dos constituintes. Para a implementação de cada padrão, atuam fortes restrições não só de ordem estrutural, como a que diz respeito ao cancelamento da marca de plural segundo a posição linear e relativa do constituinte no SN, mas sobretudo de natureza social, entre as quais se destaca a maior ou menor exposição do indivíduo à educação formal.

6. A VARIEDADE URBANA DO PORTUGUÊS DE MOÇAMBIQUE

A análise sobre o PM baseia-se num *corpus* que retrata a grande complexidade linguística de Moçambique e, em especial, da capital, Maputo, que, como a mais representativa área urbana do país, apresenta o maior número de falantes de Português como L1 ou L2 e, ao mesmo tempo, congrega indivíduos de diversificada gama linguística oriundos de outras regiões. Embora, na amostra, se tenham controlado, de forma convencional, as variáveis *sexo*, *faixa etária* e *nível de escolaridade*, o mesmo não se conseguiu em relação às variáveis *estatuto do Português* e *língua de intercomunicação*, que não estão uniformemente distribuídas pelas células, mas podem servir de teste para fundamentar determinadas hipóteses. Há, apenas, 7 informantes utentes de Português como L2 e 11 que o têm como L1, conforme se demonstra no Quadro 2, a seguir, em que também se faz a distribuição dos indivíduos pelos fatores (indicados pelos números de 1 a 4) que compuseram a variável *língua de intercomunicação*.

- 1 – Fala apenas o Português (o indivíduo não entende nenhuma das línguas locais)
- 2 – Fala o Português e, em alguns contextos, uma ou mais das línguas locais
- 3 – Fala o Português e apenas compreende, mas não fala, uma ou mais das línguas locais
- 4 – Fala mais uma das línguas locais do que o Português

Quadro 2 Distribuição dos informantes pelas variáveis sociais

ESCOLARIDADE SEXO IDADE	NÍVEL 1 (FUNDAMENTAL)		NÍVEL 2 (MÉDIO)		NÍVEL 3 (SUPERIOR)	
	HOMEM	MULHER	HOMEM	MULHER	HOMEM	MULHER
FAIXA A (18-35 ANOS)	L2 (2)	L1 (2)	L1 (2)	L2 (2)	L1 (3)	L1 (3)
FAIXA B (36 A 55 ANOS)	L2 (2)	L2 (2)	L1 (2)	L1 (2)	L1 (3)	L1 (2)
FAIXA C (56-75 ANOS)	L2 (4)	L1 (4)	L2 (2)	L2 (2)	L1 (1)	L1 (3)

Observe-se que todos os indivíduos que compõem o nível 3 de escolaridade são falantes de Português como L1, enquanto os homens do nível 1 o usam como L2. No nível 1, entre as mulheres, e, no nível 2, entre ambos os sexos, falantes de L1 e L2 estão aleatoriamente representados.

A análise atomística dos dados da amostra recolhida em Maputo está ancorada em 4.707 constituintes flexionáveis do SN, em que apenas 140 (3%) não apresentam marca de número.

Levando-se em conta esse índice global, do ponto de vista quantitativo, se poderia dizer que a amostra indica tratar-se de uma regra semicategórica, nos termos de Labov (2003). No entanto, a observação detida dos dados, como a performance de determinados informantes e o comportamento dos núcleos em segunda posição, permite classificar a regra como variável, ainda que de forma incipiente para alguns deles.

Observe-se, à luz do Quadro 2 acima, a Tabela 8 com a distribuição dos 140 constituintes com cancelamento da marca de número na fala dos 18 informantes.

Tabela 8 Distribuição das 140 ocorrências de cancelamento da marca de número em constituintes flexionáveis dos SNs, pelos informantes da amostra

ESCOLARIDADE SEXO IDADE	NÍVEL 1 (FUNDAMENTAL)		NÍVEL 2 (MÉDIO)		NÍVEL 3 (SUPERIOR)	
	HOMEM	MULHER	HOMEM	MULHER	HOMEM	MULHER
FAIXA A (18-35 ANOS)	13/119 10,9%	7/225 3,1%	0/320 0%	9/229 3,9%	3/309 0,97%	4/305 1,3%
FAIXA B (36 A 55 ANOS)	8/188 4,2%	24/81 29,6%	2/152 1,3%	1/340 3%	2/160 1,25%	3/317 0,94%
FAIXA C (56-75 ANOS)	51/93 54,8%	4/440 0,9%	3/173 1,7%	1/171 0,58%	1/442 0,22	4/616 0,64

Entre os indivíduos de nível superior, todos falantes de Português como L1, a regra é semicatórica, com índices que variam de 0,22% a 1,3% de cancelamento. Entre os de nível médio, há desde os que não apresentam nenhum ou apenas um caso de cancelamento (o homem da faixa A, as mulheres das faixas B e C) em contraposição à mulher da faixa A, com 9 cancelamentos (3,9%). Parece ser entre os indivíduos de nível fundamental que a regra é variável, a despeito de a informante mais velha ter apresentado apenas 0,9% de cancelamento. Há índices de 54,4%, 29,6% e 10,9%, todos bastante significativos.

Por outro lado, os 4.707 constituintes estão distribuídos por 2.319 SNs, a maioria, como se observa na Tabela 9, com o núcleo na segunda posição (1.934, correspondentes a 90%), o que também demonstra que predominam SNs de dois constituintes. Há, ainda, 134 núcleos em primeira posição, todos com a marca, 230 em terceira, com apenas 3% de não marcação, e 12 em quarta, com marca de número. Conclui-se que a variação no âmbito dos núcleos se dá quando ele se encontra na segunda ou terceira posições.

Partindo-se do princípio de que se trata de uma comunidade altamente complexa do ponto de vista sociolinguístico, decidiu-se proceder à análise variacionista levando em conta os 18 informantes, mesmo sabendo-se que, na fala de alguns deles, a regra é categórica ou semicatórica (item 6.1). Em seguida, focalizaram-se, de forma quantitativa e qualitativa, os dados dos indivíduos que apresentam índice igual ou superior a 5% de cancelamento, com base apenas nas variáveis estruturais (item 6.2), uma vez que a atuação das variáveis sociais fica bem explicitada na primeira das análises. Este procedimento teve por intuito resgatar informações sobre algumas variáveis descartadas na análise global, como *marcas precedentes*, que poderiam melhor especificar o comportamento da

variável concordância nominal e expor todas as ocorrências referentes a esses três informantes.

Tabela 9 Índices relativos à (não)marcação do núcleo nos 3.219 SNs da amostra

ÍNDICES RELATIVOS À POSIÇÃO DO NÚCLEO NO SN					
POSIÇÃO LINEAR DO NÚCLEO NO SN	NÚCLEOS COM AUSÊNCIA DE MARCA		NÚCLEOS COM PRESENÇA DE MARCA		EXEMPLOS
	Nº	%	Nº	%	
Primeira	0/134	0	134/134	100	[<i>conflitos religiosos</i>]
Segunda	99/1943	5%	1844/1943	95	[<i>os objetivo</i>]
					[<i>algumas famílias</i>]
Terceira	8/230	3%	222/230	97	[<i>essas minhas amiga</i>]
					[<i>dos melhores funcionários</i>]
Quarta	0/12	0	12/12	100	[<i>os meus próprios textos</i>]

6.1 A análise geral quantitativa

A análise quantitativa, global, isto é, com os 18 informantes da amostra, como já se observou, levou em conta 4707 constituintes dos SNs, dos quais 140 (3%) não apresentaram marca de número.

Mostraram-se favorecedoras do cancelamento, com *input* .08 e significância .003, por ordem de seleção, as variáveis *língua de intercomunicação*, *posição linear e relativa do constituinte no SN*, *nível de escolaridade*, *estatuto do Português* e *saliência fônica*. Como ocorre no PST, as variáveis sociais são extremamente salientes para a implementação da regra, destacando-se, mais uma vez, entre as estruturais, a variável selecionada em segundo lugar.

Para o controle da variável *língua de intercomunicação*, amalgamaram-se os fatores indicados como 1 e 4 no Quadro 2 e que serão referidos como *fala só português/ fala só Português mas compreende línguas locais*, conforme se expõe na Tabela 10. Apesar de os informantes não estarem distribuídos de forma convencional pelas células, observa-se uma gradação em que parece haver a influência do contato com outras línguas para a efetivação do cancelamento, conforme uma das hipóteses que norteou esta pesquisa: se o indivíduo fala só português ou apenas entende uma das outras línguas (P. R.: .45), ele tende a cancelar menos a marca de número; por outro lado, se faz uso mais reiterado de outra língua (P. R.: .93), o

cancelamento é mais frequente. No ponto intermediário da escala estão os indivíduos que, por vezes, alternam o Português com uma outra língua (P. R.: .51).

Tabela 10 Atuação da variável *língua de intercomunicação* para o cancelamento da marca de número no SN no PM

LÍNGUA DE INTERCOMUNICAÇÃO	APL./OCO	%	P. R.
Fala só português ou fala só Português mas compreende línguas locais.	18/2273	0,8	.45
Fala Português e, em alguns contextos, uma ou mais das línguas locais	712323	3,1	.51
Fala mais uma das línguas locais do que o Português	51/93	54,8	.93
Input: .08		Significância: .003	

Embora só um dos informantes, o menos escolarizado, tenha declarado utilizar mais uma língua local do que o Português, o resultado aqui exposto, de certa forma, pode ser uma pista para uma situação bastante comum em outras áreas de Moçambique, em que as línguas locais são mais usadas do que o Português. Uma amostra que leve em conta indivíduos de comunidades mais afastadas de Maputo é fundamental para que se tenha uma visão mais abrangente sobre a dinâmica da concordância no PM.

O que se acaba de afirmar é corroborado pela seleção, em quinto lugar, da variável *Estatuto do Português*, em que se verifica que são os indivíduos para os quais o Português é L2 (P. R.: .76), os que mais se distanciam das normas do PE, que, teoricamente, lhes servem de referência.

Tabela 11 Atuação da variável *estatuto do Português* para o cancelamento da marca de número no SN no PM

ESTATUTO DO PORTUGUÊS	APL./OCO	%	P. R.
L1	31/3652	0,8	.41
L2	109/1053	10,4	.76
Input: .08		Significância: .003	

Intimamente relacionada a essas duas variáveis, está o *nível de escolaridade* (cf. Tabela 12), selecionado em terceiro lugar, uma vez que quatro dos sete informantes que têm o Português como L2 são de nível fundamental e, à exceção de um deles, se enquadram nos dois últimos casos explicitados na Tabela 10 quanto ao uso de outras línguas.

Os indivíduos de nível fundamental, portanto, são os que mais cancelam o morfema (P. R.: .75) em oposição aos de nível médio/superior (P. R.: .41), que foram considerados em conjunto pelo fato de terem apresentado, respectivamente, 3.2% e 3.3% de cancelamento, índices praticamente idênticos.

Tabela 12 Atuação da variável *nível de escolaridade* para o cancelamento da marca de número no SN no PM

NÍVEL DE ESCOLARIDADE	APL./OCO	%	P. R.
1 – Fundamental	107/1147	9,3	.75
2/3 – Médio/Superior	33/3560	0,9	.41
Input: .008			Significância: .003

Como tem ocorrido nas análises sobre o PB e o PST, a variável estrutural sistematicamente selecionada como a primeira ou a segunda mais relevante para a aplicação ou não da marca de número é a *posição linear e relativa do constituinte no SN*. Na amostra representativa do PM, apesar do baixíssimo índice de não marcação (3% equivalentes a 140 ocorrências), essa variável comportou-se segundo os padrões esperados (cf. Tabela 13) ratificando, mais uma vez, a proposta de *continuum* de marcação de plural no SN sugerida por Brandão (2013).

Tabela 13 Atuação da variável *posição linear e relativa do constituinte no SN* para o cancelamento da marca de número no SN, no PM

POSIÇÃO LINEAR E RELATIVA DO CONSTITUINTE NO SN		APL./OCO	%	P. R.
Pré-nuclear	1ª. posição	20/1889	1	.25
	2ª ou 3ª. posição	2/2015	0,9	.30
Nuclear	1ª. posição	0/143	0	–
	2ª. posição	97/1945	5	.70
	3ª. ou 4ª. posição	9/253	3,6	.72
Pós nuclear	2ª. posição	8/111	7,2	.87
	3ª. posição	2/127	1,6	.56
	4ª, 5ª. ou 6ª. posição	2/24	8,3	.94
Input: .08			Significância: .003	

Na Tabela 13, observam-se tendências similares às obtidas para o PST (Tabela 3):

- (a) no âmbito dos constituintes pré-nucleares, os menores índices de cancelamento encontram-se na primeira e segunda/terceira posições, com respectivamente, . P. R.: 25 e P. R.: 30;
- (b) entre os núcleos, os que estão em primeira posição não apresentam cancelamento, corroborando a atuação dos constituintes que normalmente ocupam a posição 1 no SN; em segunda e terceira/quarta posições, vai aumentando o índice de cancelamento, respectivamente, P. Rs.: .70 e .72 ; e
- (c) entre os constituintes pós-nucleares, em segunda posição, o índice de cancelamento suplanta o dos núcleos (P. R.: .87), diminuindo na segunda (P. R.: .56) e voltando a aumentar nas últimas posições (P. R.: 94).

Também reiteradamente relevante nas análises sobre o PB e o PST, a última das variáveis selecionadas foi a *saliência fônica*. Em função do pequeno número de dados com cancelamento, reuniram-se no fator *baixa saliência*, vocábulos terminados em vogal oral (*filho*) ou nasalizada (*homem*) e no fator *média/alta saliência*, os vocábulos terminados no singular por /S/ (*vez*), /R/ (*cor*), /l/ (*metical*) e -ão (*operação*).

Como se esperava, são as formas com baixa saliência fônica aquelas que mais predisõem ao cancelamento da marca de número (P. R.: .52).

Tabela 14 Atuação da variável *saliência fônica* para o cancelamento da marca de número no SN, no PM

SALIÊNCIA FÔNICA	APL./OCO	%	P. R.
Baixa	132/4264	3,1	.52
Média/Alta	8/442	1,8	.25
Input: .08		Significância: .003	

6.2 Estendendo a análise: outros aspectos quantitativo-qualitativos

Complementa-se a análise geral referente ao PM, em duas etapas, a primeira ainda de natureza quantitativa, a segunda de caráter especificamente qualitativo, considerando-se apenas as performances dos três informantes que, no Quadro 2, aparecem com índice de cancelamento igual ou superior a 5% – que, segundo Labov (2003), já caracteriza uma regra como variável. Num total de 293 dados, 30% (88 ocos) apresentam cancelamento. Os três indivíduos são falantes de Português L2, dois deles fazem uso eventual de outra língua e o mais velho e me-

nos escolarizado se expressa basicamente numa língua local. Como se verifica, a taxa de cancelamento cresce significativamente, o que baixa o índice de concórdância para 70%.

A primeira etapa da análise, que leva em conta apenas índices percentuais, concentra-se em três das variáveis descartadas da análise de cunho geral – *marcas precedentes*, *animacidade do núcleo* e *classe do vocábulo*, cujos índices foram reunidos na Tabela 15 – e retoma a variável *saliência fônica* (Tabela 16), apenas para exemplificar e comentar os casos de cancelamento registrados.

Na Tabela 15, é interessante verificar que os maiores índices de cancelamento se encontram nos núcleos com traço [+animado] [-humano] (66%), seguidos dos que apresentam traço [+animado] [+humano] (51,5%). Por último, vêm os de valor negativo para os dois traços (44%).

Tabela 15 Índices relativos a três variáveis não selecionadas na análise global na fala de três indivíduos (índices de cancelamento)

ANIMACIDADE DO NÚCLEO	APL./OCO	%	EXEMPLOS
[+ animado] [+ humano]	35/68	51,5	[aqueles <i>polícia</i>]
[+animado] [-humano]	4/6	66,7	[tudo meus <i>boi</i>]
[-animado] [-humano]	40/90	44,4	[nos <i>banco</i>]
Marcas precedentes			
Ausência de constituinte precedente	6/120	5	[<i>na</i> portas]
Marca formal no constituinte imediatamente precedente	54/121	44,6	[<i>das pessoa</i>] [<i>as empresa</i>]
Ausência de marca no constituinte imediatamente precedente	1/5	20	[<i>o</i> animais]
Numeral imediatamente precedente	27/46	58,7	[dezoito <i>ano</i>]
Marca formal e semântica precedentes	1/1	100	[três dias <i>consecutivo</i>]
Classes do vocábulo			
Artigo definido	4/61	6,6	[<i>o</i> animais]
Demonstrativo	0/18	0%	–
Possessivo	0/15	0%	–
Substantivo	79/163	48,5	[<i>as bomba</i>] [<i>as criança</i>]

Adjetivo	3/11	3,8	[turmas <i>pequeno</i>] [dos pontos <i>negativo</i>]
Quantificador indefinido	0/21	0	–
Quantificador universal	2/4	50%	[<i>todo</i> esses faculdade]

Marcas precedentes é uma variável tradicionalmente controlada para aferir o princípio de que marcas levam a marcas e zeros levam a zeros. Assim, no caso do SN, a presença de um constituinte com marca de número ou sem ela levaria, respectivamente, ao uso ou não da marca no constituinte subsequente. A Tabela 15 demonstra que tal variável não se aplica à variedade em foco, o mesmo podendo-se afirmar em relação às demais variedades não europeias.

É exatamente após um numeral (58,7%) e um constituinte com marca formal (44,6%) que se registram os maiores índices de apagamento, o que, conjugado à grande preponderância de SNs de dois constituintes, deixa claro que, nos padrões não canônicos de concordância nominal, é o primeiro constituinte aquele que deve conter a indicação de pluralidade. Corroboram essa premissa não só o baixo índice (5%) de cancelamento quando o constituinte é o primeiro do SN, mas também a alta taxa de marcação nos artigos (57 em 61 ocorrências) e nos vocábulos que normalmente ocupam a posição pré-nuclear. Quando o primeiro constituinte não é marcado, o segundo leva a marca (cf. tudo *meus* boi / todo *esses* faculdade).

Os dois exemplos acima referidos, os únicos SNs de três constituintes dos 57 que foram registrados na fala do informante PM-C1hO7, o mais velho e menos escolarizado da amostra, merece especial atenção pelo fato de conter uma característica ressaltada por Baxter (2009) em sua análise dos Tongas em São Tomé já comentada na seção 2. Segundo o autor, o elemento adjacente ao núcleo teria mais probabilidade de receber a marca de número, condicionamento que só pode ser comprovado em SNs de mais de dois constituintes, pouquíssimo frequentes na modalidade falada, em que predominam SNs de dois lugares. Observe-se, ainda sobre o mencionado informante, haver, em sua fala, como se verá adiante, variação de gênero, o que se comprova em [todo *esses* faculdade] e em [*minhas* cabrito], que alterna com [*meus* cabrito]. Caso semelhante também ocorre na fala de B1m5O: [turmas *pequeno*].

A Tabela 16 foi organizada para mostrar que, apesar de as formas fonicamente mais salientes serem as menos suscetíveis ao cancelamento, no PM há casos em que a marca deixa de ser aplicada, como nas três ocorrências de *televisor* e nas de *canal*, *pão* e *rapaz* (geralmente marcadas), fato que também se verifica em amostras do PB.

Tabela 16 Índices relativos à variável saliência *fônica* na fala de três indivíduos

SALIÊNCIA FÔNICA	APL./OCO	%	EXEMPLOS
Itens com marca dupla de plural	0/0	0/0	–
Itens terminados em vogal oral no sing.	81/264	30,7	[nas <i>novela</i>] [os nossos <i>pai</i>] [<i>todo</i> esses <i>faculdade</i>]
Itens terminados em vogal nasal no sing.	1/3	33,3	[aquelas <i>mamã</i>]
Itens terminados em L no sing.	1/6	16,7	[outros <i>cana</i>]
Itens terminados em -ÃO no sing.	1/4	25%	[quantos <i>pão</i>]
Itens terminados em R no sing.	3/5	60%	[nos <i>televisor</i>]
Itens terminados em S no sing.	1/11	9,1	[três <i>rapaz</i>]

A segunda etapa da análise busca mostrar como se estruturam, em relação às classes de palavras que ocupam as margens esquerda e direita do SN, todos as 88 ocorrências de cancelamento (80 SNs) presentes no *corpus* dos três referidos informantes, o que poderá servir de exemplo para o que, em geral, se verifica nas variedades não europeias em maior ou menor grau.

Inicia-se com a observação de dois informantes. Na fala de PM-A1h2O, há 119 constituintes, dos quais 13 (10,9%) não apresentam marca de número, conforme se explicita a seguir.

- (1) em [alguns *momento*]..tá agredir então a pessoa
- (2) se ajuda [três *pessoa*]
- (3) [quatro *pessoa*] depois vem a ser agredido
- (4) então é um [dos pontos *negativo*] que eu vejo
- (5) o senhor [três dias *consecutivo*] a pedir dinheiro
- (6) trocam de zonal/ [essas *pessoa*] trocam de zona para
- (7) pode ficar lá [dois *dia*] três dias até pode não
- (8) lá é provável que chegues [*na* portas]
- (9) quando encontra [as *pessoa*] na bicha
- (10) dar oportunidade [aquelas *mamã*] aquelas pessoas
- (11) sim depende [das *empresa*] da onde você vai trabalhar

(12) [das *família*] principalmente

(13) daquela idade [dezoito *ano*] encontra

Na fala de PM-B1m2O, em 81 constituintes, 24 (29,6%) não apresentam o morfema de número.

(14) só assistindo [nos *televisor*]

(15) nós subimos [aqueles *carro*]

(16) é assim tem... [turmas *pequeno*]

(17) porque [nos *televisor*] assistimos muito

(18) cinema e... [nas *discoteca*]

(19) [os *livro*] dão na escola

(20) dão [os *livro*] mas eu tenho que comprar [os *caderno*] [as *caneta*]

(21) não vou conseguir né sustentar [meus *filho*]

(22) o pouco que ganho consigo comprar comida pr[os meus *filho*]

(23) Uma mulher pode ter, [sete *filho*] [nove *filho*]

(24) e eu tenho [três *filho*] é porque estou a me prevenir

(25) [os nossos *pai*] ganhavam pouco

(26) ir à igreja n[os *domingo*]rezar

(27) [dezoito *hora*] estão em casa

(28) há [muitas *igreja*]... [muitas *igreja*]

(29) mas não é como eu e [meus *filho*]

(30) [dezoito *ano*] uma menina tá casada ainda não brincou

(31) só vejo [nos *televisor*] né n[as *novela*]

(32) pra você assistir [outros *canal*]

Comparando-se os dois informantes, verifica-se que:

- (i) dos 37 SNs arrolados, 33 (89,1%) são de dois constituintes;
- (ii) apenas quatro SNs são de três constituintes – cf. (4), (5), (22) e (25);
- (iii) só dois deles – (4) e (5) apresentam constituinte pós-nuclear;
- (iv) somente um SN não apresenta marca num constituinte pré-nuclear (8);
- (v) apenas um SN se inicia pelo núcleo, que é marcado (16);
- (vi) quando há dois constituintes flexionáveis antes do núcleo, ambos são marcados (22) e (25);

- (vii) numerais são os constituintes à esquerda que mais predis põem à ausência de marca no núcleo, como já se demonstrou na Tabela 15;
- (viii) entre os constituintes pré-nucleares, encontram-se artigos, predominantemente, além de possessivos, demonstrativos, pronomes indefinidos e quantificadores.

Tais características são prototípicas da variação encontrada em áreas urbanas, em que os índices gerais de concordância são altos, mesmo entre indivíduos de baixa escolaridade em qualquer das variedades não europeias. O informante PM-C1h4O, como se verá de (33) a (81), no entanto, destaca-se dos demais, aproximando-se da performance de muitos dos indivíduos de perfil rural no PB (na amostra de ALMEIDA, 1997, o índice de não marcação do núcleo em segunda posição é de 87%). Em 93 constituintes, 51 (54,8%) não apresentam marca de número.

Os enunciados de (33) a (81), a seguir, que correspondem a todas as ocorrências de SNs com um ou dois constituintes com cancelamento da marca de número na fala desse informante demonstram não apenas a vitalidade desse processo, mas também deixam entrever outras características como a já mencionada troca de gênero – (33), (39), (73), (77) – o uso de *todo* e *tudo*, em lugar de *todos* – (39), (53) – e a tendência à abertura de sílaba em casos de vocábulos com coda – *guardare* (33), que oscila com *guardar* (73); *mulhere* (81), entre vários outros casos que se registram na entrevista. Nesse sentido, é interessante notar que em (65), *animais*, o plural canônico, oscila com *meticales* (como plural de *metical*, a moeda do país), em contextos de concordância plena como em *três mil e seiscentos meticales*.

- (33) a ir guardare [esses *faculdade*]
- (34) a guarnecer com [aqueles *polícia*]
- (35) [**duas** *jipe*]
- (36) [os *carro*] ajudaram a mim de dia
- (37) com [**duas** *filha*] que eu tenho
- (38) com [tantos *ano*] que já tinha
- (39) [*todo* esses *faculdade*]
- (40) explodiu [as *bomba*] “bum”
- (41) [as *criança*] nasceu na minha vida
- (42) [os *guarda*] agora da antiga
- (43) [**oito** *mese*] acabamos lá
- (44) [**nove** *filho*]... onde que pá

- (45) com [setenta e cinco *ano*]
(46) [meus *filho*] já casou
(47) [**dois quarto**]... e sala
(48) [**dois quarto**]... e sala
(49) [**dois quarto**]... e sala
(50) deixou [**quatro criança**]
(51) deixou [**quatro criança**]
(52) deixou [**quatro criança**]
(53) [*tudo* meus *boi*]
(54) [os *cabrito*] aqui
(55) em [**três dia**]
(56) alimentar com [as *criança*]
(57) não sei agora [as *coisa*]
(58) deixou [**quatro criança**]
(59) [**quatro pessoa**]
(60) compra [quantos *pão*]
(61) assaltos aqui [nos *banco*]
(62) para [nossas *criança*]
(63) [as *criança*] que
(64) [as *criança*] que
(65) [o animais]
(66) [aos *funcionário*]
(67) [as *colônia*]
(68) são... [nossos *futuro*] aí
(69) [nossas *casa*]
(70) [**duas filha**] e eu só
(71) eu tenho [**nove filho**]
(72) eu tenho [**nove filho**]
(73) a guardar [minhas *cabrito*]
(74) guarnecer [meus *cabrito*]
(75) na época[*dos ouro*]
(76) são [muitos *católico*] aqui

- (77) já já [muitos igreja]
 (78) [nas coisa] dele
 (79) tinha [velhos colono]
 (80) saiu [esses colono]
 (81) [quatro filho] da minha mulhere

Assim, dos 49 SNs com cancelamento referentes a PM-C1h4O:

- (i) 47 são de dois constituintes, havendo apenas um encabeçado por vocábulo que não aparece prototipicamente em posição pré-nuclear (o adjetivo *velhos*, em (79));
 (ii) Em toda a amostra, há apenas dois SNs de três constituintes em que só um deles é marcado – (39) e (53);
 (iii) quando o numeral é o elemento pré-nuclear, a norma é a não marcação do núcleo, como se verifica nos 17 dados em negrito, entre os 23 presentes na amostra; os seis casos de concordância ocorreram combinados com o vocábulo *meticais*;
 (iv) os constituintes pré-nucleares são aqueles já mencionados na fala dos outros dois informantes.

Como se pôde verificar, por esta etapa da análise, o estudo da concordância em variedades africanas é complexo, pois, apesar de os diferentes padrões podem se repetir, com maior ou menor frequência, na fala de todos os indivíduos, a implementação de alguns deles depende de restrições de ordem social, tais como o fato de o indivíduo ser falante de Português como L1 ou L2, ter maior ou menor contato com outras línguas, ser mais ou menos escolarizado. A incidência dos padrões não marcados aqui considerados tem caráter escalar, atuando em falantes com menor nível de escolaridade (código 1), decrescendo segundo a faixa etária (C-B-A) e, sobretudo, em função do contato com outras línguas: PM-C1h4O declarou que usa sua língua materna, o Changana, em todas as situações comunicativas, só falando Português no trabalho, enquanto os dois outros só se expressam em Português, embora PM-B1m2O e PMA1h2O tenham, respectivamente, o Changana e o Emoninga como línguas maternas.

[+ cancelamento] PM-C1h4O \longrightarrow PM-B1m2O \longrightarrow PMA1h2O [-cancelamento].

Diante dos resultados obtidos nas análises do PST e do PM e tendo em vista o objetivo geral do estudo das variedades africanas explicitado na introdução (traçar um *continuum* afro-brasileiro referente à concordância nominal), cabe verificar resultados de análises concernentes ao PB.

7. A VARIEDADE URBANA DO PORTUGUÊS DO BRASIL: VISÃO PANORÂMICA

Seguindo a mesma metodologia empregada para o estudo do PM e do PST no que respeita às variáveis estruturais e às variáveis *sexo*, *faixa etária* e *nível de escolaridade*, Brandão (2011, 2013) analisou o PB com base em amostra selecionada de 36 entrevistas realizadas na Região Metropolitana do Rio de Janeiro nas localidades de Nova Iguaçu e Copacabana, focalizadas em separado.

Dos 3.716 constituintes do SN referentes a Copacabana e dos 3.777 concernentes a Nova Iguaçu, 284 (7,6%) e 338 (8,9%), respectivamente, apresentaram cancelamento da marca.

Como se verá melhor no Quadro 3 da Seção 8, restringem a atuação da regra tanto variáveis sociais quanto estruturais: em ambas as localidades, embora com diferente representatividade, *posição linear e relativa dos constituintes no SN*, *nível de escolaridade*, *faixa etária*, *saliência fônica*, e, ainda, *sexo* na fala de Copacabana e *animacidade do núcleo* na de Nova Iguaçu.

As mais representativas são a *posição linear e relativa dos constituintes no SN e nível de escolaridade*, sempre altamente salientes em todos os *corpora* até aqui analisados e sobre as quais se tecem breves comentários.

No que concerne à primeira dessas variáveis, elaborou-se a Tabela 17, por meio da qual se pode observar que o cancelamento da marca de número tem baixa incidência, sobretudo na primeira posição pré-nuclear (Copacabana: P. R.: .11; Nova Iguaçu: P. R.: .17) e aumenta consideravelmente a partir do núcleo em segunda posição (Copacabana: P. R.: .84; Nova Iguaçu: P. R.: .74), situação já observada no PST e no PM.

Tabela 17 Atuação da variável *Posição linear e relativa no SN* para o cancelamento da marca de número no SN no PB (Copacabana e Nova Iguaçu)

POSIÇÃO AMOSTRA	PRÉ-NUCLEAR		NUCLEAR			PÓS-NUCLEAR		
	POS. 1	POS. 2/3	POS. 1	POS. 2	POS. 3/4	POS. 2	POS. 3	POS. 4/5/6
Copacabana	8/1454	2/127	2/81	236/1698	8/164	5/75	77/89	1/26
	0,6%	1,5%	2,5%	13,9%	11%	6,7%	12%	3,8%
	.11	.31	.43	.84	.84	.71	.84	.54
Nova Iguaçu	19/1450	6/155	4/11	238/1628	18/194	9/95	34/109	10/35
	1,3%	3,9%	3,6%	14%	5,1%	9,5%	31,2%	28,6%
	.17	.47	.40	.74	.68	.72	.90	.89

No que respeita à atuação das demais variáveis estruturais, *saliência fônica* e *animacidade do núcleo* apresentam resultados semelhantes aos que se verificam nas demais variedades: menor *saliência fônica* (P. R.: .54 em Copacabana e .52, em Nova Iguaçu) e núcleos com traço [-animado] (P. R.: .55 em Copacabana e em Nova Iguaçu) condicionam o cancelamento.

Quanto às variáveis sociais, constata-se, na Tabela 18, a seguir, que o desempenho de indivíduos com nível superior (P. R.: .24, em Nova Iguaçu e .18 em Copacabana) contrasta com o dos indivíduos de nível fundamental e médio. Em Nova Iguaçu, estes últimos são os que mais se destacam quanto ao cancelamento (P. R.: .76), enquanto em Copacabana os de nível fundamental lideram o processo (P. R.: .86).

Tabela 18 Atuação da variável *nível de escolaridade* para o cancelamento da marca de número no SN no PB (Copacabana e Nova Iguaçu)

NÍVEL DE ESCOLARIDADE	PB					
	NOVA IGUAÇU			COPACABANA		
	OCO	%	P. R.	OCO	%	P. R.
5 a 8 anos (fundamental)	101/1175	8,6	.62	180/1175	14,5	.86
9 a 11 anos (médio)	193/1013	19,1	.76	76/1013	6,3	.63
12 a 15 anos (superior)	44/1589	2,8	.24	19/1589	1,2	.18

Fonte: Brandão (2013: 72), com alterações

Por esta última tabela, pode-se verificar, ainda, por meio dos índices percentuais, que a regra tem diferente estatuto a depender do nível de escolaridade: é semicategórica no âmbito dos indivíduos mais escolarizados (apenas 2,8% de cancelamento em Nova Iguaçu e 1,2% em Copacabana) e francamente variável entre os menos escolarizados de ambas as localidades, com índices que superam os 5% sugeridos por Labov (2003) e que oscilam de 6,3 a 19,1%.

Apesar de Copacabana e Nova Iguaçu pertencerem à mesma comunidade de fala (a da Região Metropolitana do Rio de Janeiro), verifica-se um diferente comportamento dos indivíduos consoante as faixas etárias. Em Copacabana, são os de meia idade os que mais promovem o cancelamento da marca (P. R.: .60), havendo uma nítida tendência à adoção da regra de concordância pelos mais jovens

(P. R.: .31), que apresentam índice de cancelamento que se distancia do dos mais velhos (P. R.: .46) por apenas .15 pontos. Em Nova Iguaçu, os indivíduos mais velhos são os mais predispostos ao cancelamento da marca (P. R.: .64), enquanto os de meia idade (.41) sobressaem quanto à adoção da marca de concordância, com índice bastante próximo do dos mais jovens (.48). Em ambas as localidades, não obstante a diferente performance por faixa, a curva etária é de variação estável.

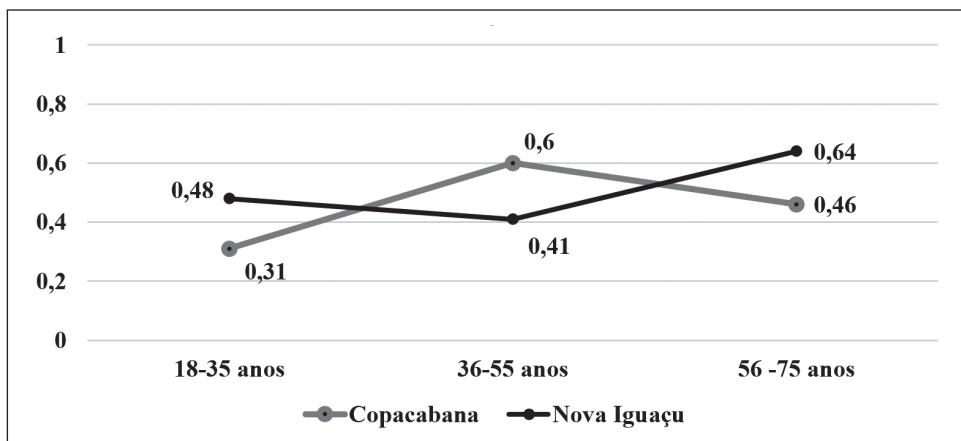


Figura 2 Atuação da variável *faixa etária* para o cancelamento da marca de número no SN, em pesos relativos, no PB (Copacabana e Nova Iguaçu)

Fonte: Brandão (2013: 82), com alterações

Já os resultados referentes à variável *sexo*, que se mostrou relevante apenas em Copacabana, atendem à expectativa inicial: os homens – embora com uma pequena margem (.04) de diferença – cancelam mais a marca (P. R.: .58) do que as mulheres (P. R.: .54), que, quando se trata de variáveis linguísticas sujeitas a avaliação social, tendem a usar as variantes mais prestigiadas.

8. SÍNTESE DOS RESULTADOS: ALGUMAS REFLEXÕES

A concordância nominal (e também a verbal) sempre esteve no cerne dos debates sobre as origens do PB. Padrões variáveis como os registrados neste estudo são considerados ora como reflexos do contato multilinguístico que caracterizou os primeiros séculos da formação do país, ora como resultado de um processo de deriva, ou, em outras palavras, como uma tendência estrutural (independente de injunções de ordem social). Daí, como também se observou de início, o interesse em verificar o que se passa nas variedades africanas do Português, em que esta língua, ainda nos dias atuais, coexiste com várias outras do grupo Banto ou com crioulos de base portuguesa.

Os resultados das análises sobre a concordância nominal realizadas nas seções 5 e 6 concernentes às variedades urbanas do PST e do PM demonstram haver uma notável semelhança com os obtidos nas análises do PB urbano, comentados na seção 7. Apesar dos aspectos divergentes que tipificam cada uma dessas variedades por conta de injunções quer sócio-históricas, quer estruturais, a regra de (não)marcação de número plural no SN apresenta vários traços convergentes, que distinguem do PE as variedades não europeias aqui descritas e entre as quais também se poderiam incluir o PCV e o PA com base nos relatos apresentados na Seção 2.

Ao longo do texto, fica evidente, de um lado, que os diferentes padrões de concordância (ou, em alguns casos, de simples indicação de pluralidade) são comuns a todas as variedades, de outro, que, como se observa no Quadro 3, as variáveis que os condicionam obedecem a restrições quer estruturais, quer sociais, também comuns.

Quadro 3 Variáveis atuantes para a (não)marcação de plural em três variedades não europeias do Português

PB		PST	PM
COPACABANA	NOVA IGUAÇU		
Nível de escolaridade	<i>Posição linear e relativa dos constituintes no SN</i>	Nível de escolaridade	<i>Língua de intercomunicação</i>
<i>Posição linear e relativa dos constituintes no SN</i>	Nível de escolaridade	<i>Posição linear e relativa dos constituintes no SN</i>	<i>Posição linear e relativa do constituinte no SN</i>
<i>Faixa etária</i>	<i>Faixa etária</i>	<i>Animacidade do núcleo</i>	Nível de escolaridade
Sexo	<i>Saliência fônica</i>	Sexo	<i>Estatuto do Português</i>
<i>Saliência fônica</i>	<i>Animacidade do núcleo</i>	<i>Contexto fonológico subsequente</i>	<i>Saliência fônica</i>

Se, de um lado, variáveis no nível sintático (*posição linear e relativa dos constituintes no SN*), semântico (*animacidade do núcleo*) e fonético (*saliência fônica*) são as que, do ponto de vista estrutural, caracterizam e definem os diferentes padrões de concordância, de outro, são variáveis de ordem social – *nível de escolaridade* (selecionada em todas as amostras), *frequência de uso de um crioulo* (selecionada na Amostra 2 do PST), *língua de intercomunicação* e *estatuto do Português* (concernentes ao PM), aquelas que determinam sua efetiva implemen-

tação, o que, no caso das três últimas, tudo indica dever-se a questões vinculadas ao contato multilinguístico.

O PB atual, cujas normas já estão consolidadas, apresenta padrões idênticos aos das variedades africanas, bem como restrições para a implementação da (não) marcação de plural bastante semelhantes às que nelas ocorrem, o que pode ser interpretado, como sugere Pagotto (2007: 468), como “evidências indiretas” do papel do contato interlinguístico ao longo do tempo:

No quadro atual dos estudos sobre a história do português do Brasil, em que a história social tem sido recolocada como uma preocupação central [...] o papel do contato interlinguístico é de fundamental importância, ainda que sobre ele só conheçamos evidências indiretas.

Levando em conta pesos relativos referentes às marcas de número obtidos nos diferentes constituintes do SN nas análises aqui realizadas, pode-se dizer que, nas variedades urbanas não europeias do Português (e mesmo nas de perfil rural), do ponto de vista estrutural, encontram-se as mesmas tendências, conforme a proposta de um *continuum* de (não)marcação de plural no SN já mencionada na seção 6.1 e aqui reformulada com o acréscimo dos resultados referentes a Moçambique. Na Figura 3, os pesos relativos referem-se à aplicação da marca de plural, não se levando em conta o núcleo em primeira posição.

	PRÉ-NUCLEAR		NUCLEAR		PÓS-NUCLEAR		
	POS. 1	POS. 2/3	POS. 2	POS. 3/4	POS. 2	POS. 3	POS. 4/5/6
PB-COP	.89	.69	.16	.16	.29	.16	.46
PB-NIG	.83	.53	.26	.32	.28	.10	.11
PST	.85	.50	.26	.24	.09	.12	.06
PM	.75	.70	.30	.28	.13	.44	.6
[+marcas]				[-marcas]			

Figura 3 *Continuum* de marcação de plural em constituintes do SN no PB, PST e PM (pesos relativos)

Fonte: Brandão (2013), com acréscimos

No *continuum*, consideram-se SNs prototípicos, isto é, aqueles compostos por constituinte(s) pré-nuclear(es), núcleo e constituinte(s) pós-nuclear(es). Em cada posição, indicam-se os PRs de presença da marca de plural obtidos nas análises gerais. Como observa Brandão (2013: 69).

[...] a primeira posição linear no 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 sua presença. [...] 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. Observe-se, como se ressaltou antes, que o deslocamento para a primeira posição de um núcleo ou de um constituinte que ocupe posição pós-nuclear concorrerá para que se torne mais suscetível a receber marca de número.

Nas variedades africanas, a situação de contato multilinguístico parece ter sido e ainda ser um elemento basilar para compreender os fatores que governam os diferentes padrões de concordância nelas registrados. O PM coexiste com uma série de línguas do grupo Banto, algumas delas circunscritas a determinadas regiões, mas todas suscetíveis de se difundirem para outras áreas por conta do deslocamento, sobretudo para Maputo, de indivíduos em busca de melhores oportunidades de vida. Os informantes da pesquisa naturais de Maputo que se declararam falantes de Português como L2 têm, em geral, o Changana ou o Ronga como língua materna, mas há os que, oriundos de outras regiões, são falantes do Xope, do Matsua ou do Sena.

É grande a complexidade também em São Tomé e Príncipe, em que o Português coexiste com crioulos de base lexical portuguesa, num espaço geográfico exíguo se comparado ao de Moçambique. No primeiro dos citados países, num total de 1001km², vivem 178.739 indivíduos, enquanto, no segundo, 28.861.863 habitantes espalham-se por uma área de 801.590km², 1.257.453 deles na capital Maputo.

Seria, portanto, de esperar que variáveis de cunho social se mostrassem altamente significativas para a (não)marcação de pluralidade no SN. Como se verificou no decorrer da análise, tanto em São Tomé quanto em Moçambique os indivíduos que mais fazem uso de outra língua são mais suscetíveis ao cancelamento da marca de plural. Em São Tomé, a variável *frequência de uso de um crioulo* mostra tal tendência entre aqueles que falam o Forro com frequência alta (P. R.:

.75) ou média (P. R.: .52). Em Moçambique, de acordo com a variável *língua de intercomunicação* – a mais relevante para a aplicação da regra – indivíduos que falam com mais frequência uma ou mais das línguas locais do que o Português (P. R.: .93), bem como aqueles que o fazem em determinadas situações (P. R.: .51), apresentam maiores índices de cancelamento, o que é corroborado pela variável *estatuto do Português*: L2, P. R.: .76; L1: P. R.: .41. Além dessas variáveis, outras, de caráter social, como *nível de escolaridade*, sempre são selecionadas nas análises variacionistas.

No Brasil, desde os primórdios da colonização, o contato interétnico e interlinguístico foi intenso e, ao que tudo indica, crucial para as especificidades que caracterizam o PB e o aproximam das variedades africanas. Como observa Pagotto (2007: 467) e as evidências aqui o comprovam,

sem o contato interlinguístico no Brasil Colônia e no Brasil Império, o português brasileiro não teria as características que teria hoje. Esta última formulação traz para o plano central o papel que ameríndios e africanos tiveram no processo de formação do português do Brasil, sem que caíamos na armadilha de supô-lo todo como fruto de um único pidgin e/ou uma única língua crioula.

Assim, pode-se dizer que “a similaridade de padrões de não marcação nas diferentes variedades sugere a existência de uma tendência de caráter universalizante, inerente às variedades do Português que emergem de situação de contato linguístico” (BRANDÃO, 2013: 89).

Por fim, cabe lembrar que há um comportamento similar entre as variedades não europeias também no que respeita à produtividade e ao estatuto da regra por segmentos sociais em oposição ao PE, em que a regra é categórica. O índice geral de cancelamento da marca de plural em variedades urbanas é, no PST, 6,6%; no PM, 3%; no PB, 8,3% (7,6% em Copacabana e 8,9% em Nova Iguaçu). No entanto, os indivíduos mais escolarizados apresentam índices bem pouco expressivos de cancelamento e os menos escolarizados índices que, em geral, ultrapassam os 5%.

No Quadro 4, busca-se fazer uma síntese do que se observou neste estudo e, ao mesmo tempo, mostrar que, mesmo com base apenas na análise de variedades urbanas, há notórias similaridades entre as variedades africanas e o PB. Mas traçar o tão almejado *continuum* afro-brasileiro só será possível com a ampliação das análises, considerando informantes de diferentes perfis e levando em conta outros fenômenos variáveis.

Quadro 4 Proposta de classificação do estatuto da regra de (não) marcação de plural no SN por variedade e segmento social em quatro variedades urbanas do Português, com base nas amostras analisadas

CONCORDÂNCIA NOMINAL	
ESTATUTO DA REGRA	VARIEDADE/SUBVARIEDADE
Catagórica	Português Europeu
Variável	Semicatagórica
	Português de São Tomé/Culta Português do Brasil/Culta Português de Moçambique/Culta
	Português de São Tomé Português do Brasil Português de Moçambique

Fonte: Brandão (2013: 90), com alterações

9. CONSIDERAÇÕES FINAIS

Tudo o que aqui se comentou, bem como as hipóteses decorrentes dos resultados obtidos, faz da indicação de pluralidade no âmbito do SN um tema instigante, não só pelas diferentes opções formais de que se podem servir os usuários do Português, mas também pelas pistas que fornecem em relação à constituição de normas vernáculas nos diferentes domínios pelos quais ele se expandiu.

Como observa Galves (2008: 161):

A transmissão irregular no quadro de uma aquisição de segunda língua com exposição insuficiente aos dados da língua-alvo, da qual temos uma imagem moderna em Moçambique e Angola hoje, parece dar conta do desenvolvimento histórico da variação encontrada no Brasil de hoje.

Tudo indica que sociedades multilíngues como as dos países africanos focalizados neste estudo muito têm a contribuir para que melhor se compreenda o processo de constituição do PB, tendo em vista que, sobretudo no passado, a coexistência de diversas línguas parece ter sido a pedra de toque para o surgimento de alguns dos aspectos mais salientes que o diferenciam do PE.

REFERÊNCIAS

ALMEIDA, Evanilda Marins. 1997. 145f. *A variação da concordância nominal num dialeto rural*. Dissertação. (Mestrado em Letras Vernáculas). Faculdade de Letras, Departamento Letras Vernáculas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1997.

AMARAL, Amaral. *O dialeto caipira*. 3. ed. São Paulo: Hucitec, Secretaria de Cultura, Ciência e Tecnologia, 1976 [1920].

AVELAR, Juanito; GALVES, Charlotte. O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística*, 30 (2): 241-288, 2014.

BAXTER, Alan. The development of variable NP plural agreement in a re-structured African variety of Portuguese. In: ESCURE, G.; SCHWEGLE, A. *Creoles, contact, and language change*. Amsterdam/Philadelphia: John Benjamins, 2004. p. 97-126.

BAXTER, Alan. A concordância de número. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Org.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009. p. 269-293.

BRAGA, Maria Luisa. *A concordância de número no sintagma nominal no Triângulo Mineiro*. 88f. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica, Rio de Janeiro, 1977.

BRAGA, Maria Luisa; SCHERRE, Maria Marta Pereira. A concordância de número no SN na área urbana do Rio de Janeiro. In: ENCONTRO NACIONAL DE LINGÜÍSTICA, 1, 1976, Rio de Janeiro. *Anais ...* Rio de Janeiro: PUC, 1976. p. 464-477.

BRANDÃO, Sílvia Figueiredo. Concordância nominal em duas variedades do português: convergências e divergências. *Revista Veredas*, 15 (1): 164-178, 2011a.

BRANDÃO, Sílvia Figueiredo. Concordância nominal na variedade urbana do Português falado em São Tomé. *Ciências Sociais e Humanas em Revista*, 33(1): 31-42, 2011b.

BRANDÃO, Sílvia Figueiredo. O cancelamento da marca de número nominal na variedade urbana não *standard* do Português de São Tomé. Comunicação apresentada ao XVI Congresso Internacional de la Asociación de Lingüística y Filología de la América Latina (ALFAL). Alcalá de Henares, Madrid, 6-9 de junho, 2011c.

BRANDÃO, Sílvia Figueiredo. Patterns of agreement within the Noun Phrase. *Journal of Portuguese Linguistics*, 12(2): 51-100, 2013.

BRANDÃO, Sílvia Figueiredo; VIEIRA, Sílvia Rodrigues. A 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*, 22(2): 7-41, 2012a.

BRANDÃO, Sílvia Figueiredo; VIEIRA, Sílvia Rodrigues. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa, Revista de Linguística*, 56(3): 1035-1064, 2012b.

BRANDÃO, Sílvia Figueiredo; VIEIRA, Sílvia Rodrigues. Para uma análise qualitativa de dados quantitativos: o caso da concordância em Português. Comunicação apresentada ao IV DIA: Repenser la variation linguistique. Zurique: Universidade de Zurique, 12-14 de setembro, 2016.

FIGUEIREDO, Carlos. Filipe Guimarães. A concordância variável no sintagma nominal plural do português reestruturado de Almoxarife. *Papia: Revista Brasileira de Estudos Crioulos e Similares* 18: 23-44, 2008.

FIGUEIREDO, Carlos. Filipe Guimarães. 2 v. 798 fls. *A concordância plural variável no sintagma nominal do português reestruturado da comunidade de Almoxarife, São Tomé*

(*Desenvolvimento das regras de concordância variáveis no processo de transmissão-aquisição geracional*). Tese (Doutorado em Linguística). Faculdade de Ciências Sociais e Humanas, Departamento de Português, Universidade de Macau, Macau, 2010.

GALVES, Charlotte. Concordância e origens do português brasileiro. In: SEDRINS, A. P.; CASTILHO, A. T.; SIBALDO, M. A.; LIMA, R. B. (Org.). *Por amor à linguística: miscelânea de estudos linguísticos dedicados à Maria Denilda Moura*. Maceió, EdUFAL, 2012. p. 123-149.

GALVES, Charlotte. O papel das línguas africanas na formação do português brasileiro: (mais) pistas para uma nova agenda de pesquisa. *Gragoatá* (24): 145-164, 2008.

GÄRTNER, Eberhard. Coincidências dos fenômenos morfo-sintáticos do substandard do português do Brasil, de Angola e de Moçambique. In: DEGENHARDT, R.; STOLTZ, T.; ULFERTS, H. (Eds.). *Afrolusitanistik – eine vergessene Disziplin in Deutschland? Dokumentation des 2. Bremer Afro-Romania Kolloquiums vom 27-29. Juni 1996*, Bremen: Universität Bremen, 1996. p. 146-180.

GONÇALVES, Perpétua. *A gênese do Português de Moçambique*. Lisboa: INCM, 2010.

GONÇALVES, Perpétua. O Português em África. In: RAPOSO, Eduardo B. P.; NASCIMENTO, M. F. B.; MOTA, M. A. C. da; SEGURA, Luísa; MENDES, Amália. *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian, 2013. v. 1, p. 157-178.

GONÇALVES, Perpétua; STROUD, Christopher (Org.). *Panorama do Português oral de Maputo*. v. I: objetivos e métodos. Maputo: Instituto Nacional do Desenvolvimento da Educação, 1997 (Cadernos de pesquisa do Inde n. 22).

GONÇALVES, Perpétua; STROUD, Christopher (Org.). *Panorama do Português oral de Maputo*. v. III; *Estruturas gramaticais do Português: problemas e exercícios*. Maputo: Instituto Nacional do Desenvolvimento da Educação, 1998 (Cadernos de pesquisa do Inde n. 27).

GONÇALVES, Perpétua; MORENO, Albertina; TUZINE, António; DINIZ, Maria João; MENDONÇA, Marisa. Estruturas gramaticais do Português: problemas e exercícios. In: GONÇALVES, Perpétua; STROUD, Christopher (Org.). *Panorama do Português oral de Maputo*. v. III; *Estruturas gramaticais do Português: problemas e exercícios*. Maputo: Instituto Nacional do Desenvolvimento da Educação, 1998. (Cadernos de pesquisa do Inde n. 27).

GUY, Gregory R. *Linguistic Variation in Brazilian portuguese: Aspects of the Phonology, Syntax, and Language History*. Ph.D. Dissertation. University of Pennsylvania, Philadelphia, 1981.

INVERNO, Liliana (No prelo) A transição de Angola para o português vernáculo: estudo morfosintático do sintagma nominal. In: CARVALHO, Ana M. (Ed.). *Português em contato*. Madrid, Frankfurt: Iberoamericana; Editorial Vervuert, 2009. p. 87-104.

JON-AND, Anna. 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*, 2: 28-50, 2010.

JON-AND, Anna. 2011. 157f. *Variação, contato e mudança linguística em Moçambique e Cabo Verde. A concordância variável de número em sintagmas nominais do português*. Tese (Doutorado). Department of Spanish, Portuguese and Latin American Studies, Stockholm University, 2011.

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

LABOV, William. *Principles of linguistic change: internal factors*. Oxford: Blackwell Publishers, 1994. v. 1.

LABOV, William. *Principles of linguistic change: social factors*. Oxford: Blackwell Publishers, 2001. v. 2.

LABOV, William. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (Ed.). *Sociolinguistics: the essential readings*. Blackwell Publishing, 2003. p. 234-250.

LUCCHESI, Dante. O conceito de transmissão linguística irregular e o processo de formação do Português do Brasil. In: RONCARATI, C.; ABRAÇADO, J. (Org.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003. p. 272-284.

MARQUES, Irene Guerra. Algumas considerações sobre a problemática linguística de Angola. In: INSTITUTO DE CULTURA E LÍNGUA PORTUGUESA (Ed.). *Actas do Congresso sobre a situação actual da língua portuguesa no mundo*. Lisboa: ICLP, 1983.

MARROQUIM, Mário *A língua do Nordeste (Alagoas e Pernambuco)*. 2. ed. São Paulo: Companhia Editora Nacional, 1945.

MORENO, Albertina; TUZINE, António. Distribuição social de variáveis lingüísticas no português oral de Maputo. In: STROUD, Christopher; GONÇALVES, Perpétua (Org.). *Panorama do Português Oral de Maputo: a construção de um banco de "erros"*. Moçambique: INDE, 1997. p. 71-91.

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

NARO, Anthony. Julius; SCHERRE, Maria Marta Pereira. *Origens do português brasileiro*. São Paulo: Parábola Editorial, 2007.

PAGOTTO, Emilio Gozze. Crioulo sim, crioulo não: uma agenda de problemas. In: CASTILHO, Ataliba; TORRES MORAIS, Maria Aparecida; LOPES, Ruth E. Vasconcellos; CYRINO, Sônia Maria Lazzarini (Org.). *Descrição, história e aquisição do Português Brasileiro*. São Paulo: FAPESP; Campinas: Pontes, 2007. p. 461-482.

PETTER, Margarida Maria Taddoni. Uma hipótese explicativa do contato entre o português e as línguas africanas. *Papia*, 17: 9-19, 2007.

PETTER, Margarida Maria Taddoni. Ampliando o continuum afro-brasileiro de português. *Papia: Revista Brasileira de Estudos Crioulos e Similares*, 25: 305-317, 2015.

SCHERRE, Maria Marta Pereira. 158f. 1978. *A regra de concordância de número no sintagma nominal em português*. Dissertação (Mestrado em Letras) – Rio de Janeiro, Departamento de Letras e Artes, Pontifícia Universidade Católica, 1978.

SCHERRE, Maria Marta Pereira. 1988. 555f. *Reanálise da concordância de número em português*. 1988. 2 v. Tese (Doutorado em Linguística). – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 1988.

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

VIEIRA, Silvia Rodrigues; BRANDÃO, Silvia Figueiredo. Tipologia de regras linguísticas e estatuto das variedades/línguas: a concordância em português. *Linguística* (30): 81-112, 2014.

VIEIRA, Silvia Rodrigues; BRANDÃO, Silvia Figueiredo (Eds.). Padrões de Concordância em variedades do Português. *Cuadernos de la ALFAL*, 7, 2015.

WEINREICH, Uriel, LABOV, William; HERZOG, Marvin I. Empirical foundations for a theory of linguistic change. In: LEHMANN, W.; MALKIEL, Y. (Org.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968. p. 97-195.