

COMPLEMENTANDO UM ESTUDO SOBRE CONCORDÂNCIA NOMINAL DE NÚMERO NA VARIEDADE URBANA DO PORTUGUÊS DE SÃO TOMÉ: ESTRUTURAS PREDICATIVAS/ PASSIVAS

Silvia Figueiredo Brandão

Universidade Federal do Rio de Janeiro/CNPq

Paulo Vitor Lima da Gama Soares

Universidade Federal do Rio de Janeiro/FAPERJ

RESUMO: Analisa-se, basicamente, a concordância de número em estruturas predicativas/passivas na variedade urbana do Português de São Tomé (PST), à luz dos pressupostos teórico-metodológicos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968) e com base em dados recolhidos de entrevistas de perfil sociolinguístico pertencentes ao *Corpus VAPOR* (Variedades do Português), do Centro de Linguística da Universidade de Lisboa. Com base no controle de quatro variáveis sociais e nove estruturais, busca-se determinar os fatores que atuam para a (não)marcação de plural, partindo-se da hipótese de que, nessa variedade, a concordância das referidas estruturas com o sujeito tem caráter variável a exemplo do que ocorre no âmbito do sintagma nominal (SN), cuja análise (BRANDÃO, 2011b) é retomada para fins de comparação. A observação

das estruturas passivas/predicativas, apesar do pequeno número de ocorrências, confirmou a hipótese e demonstrou que o cancelamento da marca de número é condicionado pela atuação de duas variáveis sociais (*nível de escolaridade do indivíduo e frequência de uso de um crioulo*) e pela variável estrutural *características formais do sujeito da construção*.

INTRODUÇÃO

A concordância de número (nominal e verbal) é dos temas mais focalizados no âmbito do Português do Brasil (PB), em virtude não só de ser um dos parâmetros mais emblemáticos para a distinção entre as variedades europeia e brasileira, mas também pelo fato de, nesta última, ocorrer como regra variável em todos os falares regionais e sociais. Além disso, no Brasil, sobretudo em áreas urbanas, a não marcação de número é objeto de avaliação negativa por alguns falantes, que muitas vezes a estigmatizam já que ocorre com mais frequência na fala de indivíduos de menor nível de escolaridade e/ou de origem urbana ou rural.

No que toca à concordância nominal no PB, desde os primeiros estudos que dela trataram numa perspectiva sociolinguística, como Braga (1978), Guy (1981), Scherre (1978, 1988), Almeida (1997), entre vários outros,¹ tem-se demonstrado que a par de padrões de marcação plena, como (i) *os meus irmãos*, ocorrem outras estruturas variáveis, como (ii) *aquelas coisas boa*, (iii) *as ideia interessante* e (iv) *meus filho/dois menino*, estes últimos, SNs de dois constituintes, os mais produtivos na modalidade falada.

Tal complexidade é determinada por fatores ligados, de um lado, à própria estrutura do SN e a alterações morfofonológicas decorrentes do mecanismo de flexão, e, de outro, à atuação de fatores vinculados a características dos falantes (entre as quais, sexo, idade, nível de escolaridade, área de origem ou de residência), à situação de intercomunicação (formal, informal), à modalidade de língua (falada, escrita). Como observa Scherre (1994):

Dos trabalhos realizados, conclui-se portanto, que o fenômeno da variação de número no português do Brasil pode ser caracterizado como um caso de variação linguística inerente,² tendo em vista que ocorre em contextos linguísticos e sociais semelhantes e apresenta tendências sistemáticas de variação altamente previsíveis (p. 38).

¹ Scherre (2005: p. 31-35) apresenta listagem bastante completa de textos até então inéditos e já publicados que têm por tema a concordância, entre as quais a nominal em EPPS e SNs, e que reitera e complementa a que foi apresentada em artigo de 1994 (cf. referências).

² Realce nosso.

Em 2008-2009, a perspectiva de observar a variação de número em variedades do Português de perfil urbano animou pesquisadores de três universidades – a de Lisboa, a Federal do Rio de Janeiro e a da Madeira – a realizarem pesquisas de cunho comparativo entre variedades brasileiras, africanas e europeias (continentais e insulares), iniciando pela concordância nominal e pela verbal de terceira pessoa do plural. Com base nas análises realizadas com *corpora*, em sua maioria, organizados por integrantes do Projeto que então se constituiu – *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias* –, verificou-se que, em contraposição ao Português Europeu (PE) urbano, em que a regra de concordância parece ser categórica no interior do SN (BRANDÃO, 2013) e semicategórica em relação ao verbo na terceira pessoa do plural (VIEIRA; BAZENGA, 2013), as variedades não europeias apresentam padrões variáveis semelhantes, em que a presença/ausência da marca de plural obedece a fortes restrições estruturais e sociais (BRANDÃO, 2011a, 2011b, 2013, 2018; BRANDÃO; VIEIRA, 2018, entre outros).

No âmbito do referido projeto, no entanto, não se contemplou a concordância de número que se estabelece (a) entre o particípio passivo e o sujeito da construção e (b) entre o SN ou o Sintagma Adjetival (na função de predicativos) e (i) o sujeito ou (ii) o objeto, devendo-se acrescentar que, no que tange ao PB, são poucos os estudos que tratam de tais estruturas, como se comentará mais adiante.³

Assim, neste estudo, no intuito de ampliar e complementar o conhecimento sobre a dinâmica da concordância, objetiva-se analisar, quantitativa e qualitativamente, a (não)marcação de número em estruturas predicativas/passivas (doravante EPPs) na variedade urbana do Português de São Tomé (PST), à luz dos pressupostos teórico-metodológicos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968), com base em dados recolhidos de entrevistas de perfil sociolinguístico realizadas em 2009 e pertencentes ao *Corpus VAPOR* (Variedades do Português), do Centro de Linguística da Universidade de Lisboa. Parte-se da hipótese de que, nessa variedade, a concordância das referidas estruturas com o sujeito tem caráter variável, a exemplo do que ocorre no âmbito do Sintagma Nominal (SN), focalizado por Brandão (2011b) e que também levou em conta apenas indivíduos de níveis fundamental e médio de escolaridade.

³ Uma mestranda e um graduando, orientados pela autora deste texto, estão desenvolvendo pesquisas sobre o tema em foco, respectivamente, na fala da Região Metropolitana do Rio de Janeiro e na de Moçambique.

São, portanto, focalizadas, na variedade urbana do PST, estruturas caracterizadas em (a) e (b-i) anteriormente e aqui exemplificadas, à exceção das duas últimas,⁴ com dados selecionados da amostra que serviu de base à análise.⁵

(a) eles eram massacrado (C-1-m-a)

meus filho foi batizado no católico (A-1-h-a)
com pé cruzado porque foram mortos (B-1-h-b)
todas essas coisas eram custeadas por ela (A-2-h-a)

(b-i) nós somos escrava deles (A-1-m-b)

nós somos compadre e fala dialeto (C-1-h-a)
as praias são muito bonitas (A-2-m-a)
as coisas ficaram mais cara (A-1-h-a)
acho que os manuais são importante (C-2-h-b)
as crianças chegaram assustadas / as pessoas chegaram muito cansada

O desenvolvimento deste estudo está distribuído por cinco outras seções: na primeira, a seguir, comentam-se resultados de algumas pesquisas que trataram da concordância em EPPs em diferentes variedades do PB; na segunda, faz-se uma breve caracterização da área da pesquisa, nela incluindo-se uma síntese do estudo de Brandão (2011b) que tratou da concordância no SN; na terceira, indicam-se os procedimentos metodológicos que nortearam a análise das EPPs e apresentam-se os seus resultados, que, na quarta seção, são comentados e confrontados aos que se obtiveram na referida análise de Brandão com os mesmos informantes da atual amostra; na quinta, tecem-se as considerações finais.

VISITANDO ALGUNS ESTUDOS SOBRE O TEMA NO PB

Como já se observou, poucos são os estudos que contemplam as EPPs no Português do Brasil se comparados aos que focalizam a concordância no SN. O primeiro trabalho de que se tem conhecimento – Scherre (1991) – será aqui comentado mais detalhadamente, pois na pesquisa sobre o PST consideraram-se todas as variáveis estruturais nele controladas.

⁴ Os dois últimos configuram possíveis ocorrências que, no entanto, não foram encontradas no *corpus*.

⁵ Para identificação das características dos informantes que forneceram as estruturas que servem de exemplo, ver Quadro 1, na seção referente aos procedimentos metodológicos.

O estudo de Scherre foi realizado com base em 64 entrevistas da amostra Censo. Os informantes distribuem-se segundo o sexo, a faixa etária e o nível de escolaridade, contabilizando-se 873 dados: 90% de predicativos e 10% de participios passivos. Deste total, 759 foram levados em conta na análise probabilística, sendo a marca de plural encontrada em 50% dos dados (380 ocorrências). A autora controlou, ainda, nove variáveis linguísticas, tendo-se mostrado estatisticamente significativas o *paralelismo formal das sequências de predicativos/participios no discurso*, *características formais do sujeito*, *características formais do verbo*, *estrutura do predicativo* e *processos morfofonológicos de formação do plural*. Por outro lado, *tonicidade dos itens singulares*, *ordem dos elementos na estrutura*, *material interveniente entre o verbo e o predicativo* e *tipo de estrutura* não se mostraram relevantes.

Com o controle da variável *paralelismo formal das sequências do discurso*, Scherre buscou evidências, no nível suprassentencial, para o Princípio do Processamento Paralelo, por ela proposto (1988). Foram codificados os predicativos/participios isolados contra os que ocorrem em série, precedidos por outros predicativos/participios marcados ou não marcados. A análise das variáveis *características formais do sujeito* e *características formais do verbo* buscaram evidências adicionais para o mesmo princípio, porém no nível sentencial. A primeira tinha o foco na presença/ausência do sujeito e nas marcas de sujeito, enquanto a segunda focalizava a presença/ausência do verbo e das marcas do verbo. Os resultados encontrados na análise das três referidas variáveis parecem convergir para a hipótese da autora: as formas precedidas de formas marcadas apresentam mais chances de serem marcadas e as formas precedidas de formas não marcadas apresentam mais chances de serem não marcadas. Em outras palavras, marcas levam a marcas e zeros levam a zeros.

O objetivo da análise da variável *estrutura do predicativo* era verificar a relevância da estrutura sintagmática do predicativo para a ausência/presença de marcas. A autora ressalta que a motivação para controlar essa variável são as estruturas com “tudo/todo”, em que quase não se encontram marcas de plural. Os resultados desta variável revelam que a estrutura substantiva favorece mais a presença de marcas do que a estrutura adjetiva.

As variáveis *processos morfofonológicos de formação do plural* e *tonicidade dos itens singulares* foram consideradas também por serem clássicas na literatura. A expectativa para as análises era chegar a resultados que convergissem com conclusões já existentes em estudos sobre concordância: formas mais salientes apresentam mais marcas de plural do que as formas menos salientes. Tal expectativa

foi alcançada, porém a primeira dessas variáveis foi considerada estatisticamente significativa, enquanto a segunda, não.

Seguindo com as análises de variáveis não estatisticamente significativas, com a finalidade de verificar se a ordem canônica favorece mais a presença de marcas de plural do que a *não canônica, foram analisadas as variáveis ordem dos elementos na estrutura e material interveniente entre o verbo e o predicativo*. Os resultados da primeira variável citada mostram, ainda que sem um peso relativo expressivo, o esperado: uma ordem não canônica, que interfere no processamento linear, favorece a ausência de marcas, enquanto os resultados da segunda indicam que a presença de material entre o verbo e o predicativo/particípio não interfere na presença/ausência de marca. Finalmente, a análise da variável *tipo de estrutura* pretendia observar a relevância estatística da estrutura passiva, que apresenta apenas 30 (37%), de um total de 81, casos marcados. Observou-se que a estrutura passiva desfavorece a presença da marca de plural, embora também sem peso relativo significativo, o que poderia estar relacionado à baixa ocorrência de formas passivas.

Sobre as variáveis sociais, os resultados apresentam dados normalmente interpretados como padrão de variação estável: as mulheres marcam mais o plural (variante de prestígio), que os homens; a presença das formas de prestígio é diretamente proporcional aos anos de escolarização e os jovens e os de mais idade desfavorecem as formas de prestígio, enquanto os de média idade as favorecem.

Além do estudo de Scherre, tem-se notícia dos realizados por Vazzata-Dias (1996, 2000) e Vazzata-Dias e Fernandes (2000), sobre a fala do Sul do Brasil; por Lucchesi (2008, 2009) por Salomão (2010), sobre a de São José do Rio Preto; Antonino (2007, 2012), o primeiro sobre o português popular do interior da Bahia, o segundo circunscrito a Salvador; e por Furtado (2017), sobre a fala de Fortaleza. A seguir, comentam-se três desses trabalhos.

Vazzata-Dias (2000) apresenta resultados de sua dissertação (1996), centrada na fala de Florianópolis, Chapecó e Irati, focalizando, em especial, variáveis de natureza extralinguística (sexo, escolaridade e faixa etária). Em sua análise todas as variáveis sociais foram selecionadas como relevantes, sendo que escolaridade se impôs, inclusive às de ordem estrutural. Ele mostrou que a concordância em predicativos se aplica, preferencialmente, entre indivíduos com nível colegial (P. R. .72), com mais de 50 anos (P. R. .59), do sexo feminino (P. R. 57), entre indivíduos de etnia italiana (P. R. .57).

Salomão (2010) tratou das estruturas predicativas na fala de São José do Rio Preto, com base em 153 dados selecionados de 95 entrevistas do tipo “narrativas

de experiência” do banco de dados Iboruna, com informantes distribuídos por gênero, quatro faixas etárias e quatro níveis de escolaridade. Na análise foram controladas as três referidas variáveis e as nove estruturais utilizadas por Scherre (1991). Mostraram-se como fatores relevantes para a aplicação da concordância, com índice de 61%: (a) quanto às marcas do sujeito: (i) o sujeito nulo – P. R. .84, (ii) sujeito explícito com todos os elementos nominais flexionáveis marcados; com os últimos elementos com a marca formal de plural; ou com a última marca neutralizada por contexto fonológico seguinte – P. R. .61; (b) marcas do verbo: (i) verbo com marca explícita de plural – P. R. .66, verbo nulo (elidido) – P.R. 54; (c) gênero: feminino – P. R. .72; (d) nível de escolaridade: (i) ensino superior – P.R. .65, (ii) médio – .60. A autora conclui seu estudo afirmando que

o conjunto de dados analisados ao longo deste trabalho permite concluir que a concordância no predicativo no português falado do Brasil está definitivamente arraigada na comunidade como uma autêntica regra variável [...] *a análise de gênero permite interpretar a distribuição dos dados como um padrão de variação estável, segundo o qual as mulheres favorecem mais as formas de prestígio do que os homens [...]. A análise do fator escolaridade permite deduzir a existência de uma relação diretamente proporcional entre acréscimo de educação formal e marcação de pluralidade no predicativo (p. 684).*

Furtado (2017), o estudo mais recente sobre o tema, é uma dissertação cujos dados foram selecionados de 48 entrevistas do Projeto NORPOFOR (Norma Oral do Português Popular de Fortaleza), com informantes estratificados por sexo, anos de escolaridade e tipo de registro (D2, diálogo entre dois informantes e EF, elocução formal). Com índice geral de 66,2% de marcação de número, as variáveis que se mostraram salientes para a aplicação da marca de plural foram o sexo do falante, o paralelismo formal, características formais do verbo da oração e características formais do sujeito da oração. O mestrando aplicou testes de reação subjetiva, por meio da ferramenta virtual *Google Forms*, a 300 alunos da Universidade Federal do Ceará, nativos de Fortaleza, controlando, ainda, se o aluno cursara o Ensino Médio em escola pública ou privada. Foram aplicadas 10 perguntas, que indicaram que alunos de ciências exatas tendem a ser mais conservadores em relação à concordância em predicativos e, ainda, que a não marcação de número em Fortaleza tem avaliação negativa, mesmo em situações informais.

SITUANDO BREVEMENTE A COMUNIDADE DE SÃO TOMÉ

Quanto a aspectos geo-sócio-linguísticos gerais

São Tomé é a capital da República Democrática de São Tomé e Príncipe, um arquipélago situado no Golfo da Guiné, na costa ocidental da África e constituído principalmente pelas duas ilhas que dão nome ao país. Com um total de 1.001 km², conta com 197.700 habitantes, segundo o Censo de 2017 (INE).⁶

Desabitadas até 1470, as ilhas foram descobertas pelos portugueses que as colonizaram do século XV (em 1493, iniciou-se seu definitivo povoamento) até a sua independência em 1975.

Nascimento (2018, p. 45) sintetiza aspectos da complexa história de São Tomé:

em cinco séculos de existência, vivenciou duas colonizações centrais, provenientes de dois ciclos econômicos dominantes e atuou como entreposto de escravos, sustentando o tráfico do atlântico e se submetendo a intensos fluxos migratórios. O atual panorama linguístico da região decorre de povoadamentos em massa, associados i) ao cultivo do açúcar, no século XVI; e ii) às culturas do café e do cacau, inauguradas no século XIX. A princípio, o convívio entre europeus e africanos, bem como a recorrência de entradas e saídas de cativos, culminaram com a formação e especiação do (proto)crioulo do Golfo da Guiné (*doravante* PCGG) (cf. seção 2.1.1). Tempos depois, a necessidade de um regime de contrato com trabalhadores de Cabo Verde, Angola, Moçambique, Benim e Gabão, contribuiu paulatinamente para a generalização do uso do português (*doravante* PST).

Atualmente, a maior parte da população (67,7%), segundo o INE, reside em áreas urbanas, sendo a taxa de alfabetização da ordem de 90,1%. No país, há quatro instituições de ensino superior, uma delas pública. Na capital, São Tomé, em que se encontra o principal porto do país, ocorrem as principais atividades comerciais, apesar de sua estrutura precária.

Em São Tomé e Príncipe, coexistem, além de quatro crioulos de base portuguesa – o Forro (ou Santomé) e o Angolar, ambos na Ilha de São Tomé; o Lung'ie (ou Principense) na Ilha do Príncipe e o Fa d'ambô (ou Anobonense), da Ilha de Ano Bom (província da Guiné Equatorial) – também o português dos Tongas, o Cabo-verdeano (crioulo de base portuguesa, nativo de Cabo Verde) e “resquícios de línguas do grupo Bantu” (HAGEMEIJER 2009: 1).

Para se ter ideia de aspectos da gramática do Forro, o crioulo que, no século XVIII, era “a língua materna de grande parte da população nativa de São Tomé”,

⁶ Todos os dados estatísticos aqui referidos foram obtidos no site do Instituto Nacional de Estatística, de São Tomé: <https://www.ine.st/>.

segundo Gonçalves e Hegemeijer (2015, p. 89), selecionaram-se de artigo de Espírito Santo (1998) algumas características que interessam ao tema aqui focalizado.

Entre outras observações,⁷ ele afirma que, nesse crioulo:

- (a) não há a forma singular do artigo definido, embora nomes de “deuses e pessoas” sejam “precedidos de *nen*, ou seja, *os* (artigo), que, aliás, nunca se aplica a objectos e, realmente, a tais seres quando estão no singular” (p. 51);
- (b) “quanto ao artigo indefinido, resume-se a *ũa* ou *um, uma*”, havendo “todavia, falantes que utilizam a expressão *ũa nen* (uns e umas), que somente se refere a seres humanos e divinos no plural” (p. 51);
- (c) “*ũa* surge por vezes combinado com as preposições *de* e *em*, dando origem às formas *dũa* e *nũa*, que também se aplicam a nomes do género masculino e feminino, mas apenas no singular” (p. 55);
- (d) “os substantivos forros não variam em número; a pluralização é obtida através do quantitativo que os acompanha; os falantes antepõem o quantitativo *montxi* ou *iô*, que significam *muito*, aos substantivos, os quais, no entanto, precedem *lumadu*, *bastante*, *muito*” (p. 55); e
- (e) “tal como os substantivos, os adjectivos não se flexionam em número, permanecendo no singular” (p. 57).

Desse modo, o interesse no estudo do PST reside, entre outros aspectos, em observar, nas normas em construção dessa variedade, um possível impacto do contato do Português com os crioulos, em especial, com o Forro, uma vez que, dentre os países africanos que têm o Português como língua oficial, essa é a única variedade falada como L1 ou L2 pela maioria da população – 98, 4%, em 2012–sendo, hoje, a L1 da maior parte dela, segundo Araújo (2020), que afirma que o país caminha para o monolinguismo, como já salientara Hagemeijer em 2009:

Embora o Português, a língua oficial e de prestígio em S. Tomé e Príncipe, siga oficialmente a norma do Português europeu, existem, na prática, diversos registos de Português, uns próximos dessa norma, outros com maior ou menor grau de influência dos crioulos (Afonso 2008; Lorenzino 1996a), muitas vezes determinado por factores tais como o nível de escolaridade, nível económico e o ambiente de inserção social (urbano/rural). Esta variação reflecte o conflito entre a norma oficial e a prática local e um passado recente em que o português era L2 para a maioria dos habitantes das ilhas (HAGEMEIJER, 2009, p. 19).

⁷ Fizeram-se algumas alterações nas citações para maior clareza.

Quanto à marcação de plural no SN

Embora, em relação ao PST, haja estudos sobre outras variáveis (concordância verbal de 3a. pessoa do plural, róticos, lateral /l/, ditongos, vogais médias em contexto pretônico, entre outras), destacam-se, aqui, os de Brandão (2011a e 2011b), que tratam, atomisticamente, a marcação de plural em constituintes do SN, numa perspectiva sociolinguística variacionista. No primeiro deles (2011a), consideraram-se 17 informantes distribuídos por sexo, três níveis de instrução e três faixas etárias, em que se pôde observar que, entre os indivíduos de escolaridade superior, a regra de concordância é semicategórica: em 1.139 ocorrências apenas 14, isto é, 1,2% (P. R. .24), não apresentavam marca de plural, em contraste com os demais informantes, em que a regra se mostrou variável. Respalhada por esse resultado, Brandão (2011b) realizou nova análise, considerando apenas indivíduos dos níveis fundamental e médio (12 dessa amostra e outros 10, cf. Quadro 1, na próxima seção), os mesmos 22 informantes em que se apoia o estudo sobre as EPPs.

No cômputo geral, num total de 2.375 constituintes flexionáveis do SN, predominantemente de dois constituintes, houve 2.070 ocorrências com o morfema de plural (87,2%) e 305 sem ele (12,8%), tendo-se mostrado significativas para a não marcação de número as variáveis *posição linear e relativa do constituinte no SN, nível de instrução, frequência de uso de um crioulo, saliência fônica e animacidade do núcleo*.

Quanto às variáveis estruturais,

- (a) a *posição linear e relativa do constituinte* foi a mais significativa para a determinação dos padrões variáveis observados:
 - (i) na margem esquerda do SN, contexto por excelência dos determinantes, a marca de número se fez presente, de forma praticamente categórica, na primeira posição (P. R. .14, ex.: [as dificuldade]) e, em menor escala na segunda (P. R. .55, ex.: [todos esses livro]);
 - (ii) no núcleo do SN, na segunda posição (P. R. 75, ex: com [oito ano]) ou na terceira/quarta (P. R. .77, ex: [umas pequena venda] / [os seus quatro ano]) a marca se torna menos presente;
 - (iii) nos constituintes do SN à direita do núcleo (pós-nucleares), e quanto mais afastados dele, maior é a probabilidade de não ocorrer marca: na segunda posição, P. R. .88 (ex.: trabalhos específico), na terceira, P. R. .93 (ex.: os tempos livre] e na quarta/quinta: P. R. .95 (ex.: as pessoas mais velha / n[essas zonas assim mais distante].

- (b) fatores relacionados à saliência fônica e a traços semânticos do núcleo também concorrem para a variação observada:
- (i) em vocábulos cuja forma plural é fonicamente menos saliente em relação à do singular (filho/filhos, homem/homens), há menos marcas de número – P.R. .54 – do que entre aqueles em que há maior saliência fônica (mês/meses, razão/razões): P. R. .25;
 - (ii) núcleos com traço [-animado] estão mais sujeitos à não implementação da marca (P. R. .55, ex: [uns medicamento]) do que os de traço [+animado] (P.R. .39, ex: [os filho]).

No que se refere às variáveis extralinguísticas,

- (a) *nível de instrução*, única variável extralinguística sistematicamente selecionada nas diferentes etapas de análise, sugere que, em áreas urbanas de São Tomé, há uma situação polarizada, que, de certo modo, retrata um maior ou menor domínio da norma de referência (no caso, a do PE), por parte dos indivíduos: entre os falantes de nível médio de instrução houve 3,7% de não marcação, enquanto entre os de nível fundamental, 27,1%, com pesos relativos, respectivamente, de .29 e .79; o *range* entre esses dois índices – .50 – é uma evidência dessa polarização.
- (b) *frequência de uso de um crioulo* reforça o que se indicou no item anterior a este: os indivíduos que se comunicam preferencialmente num crioulo, o que ocorre, em geral, entre os falantes de nível fundamental de instrução, são os que mais tendem a não implementar a marca de número (41,1%; P. R. .75), seguidos pelos que o utilizam apenas eventualmente (12,5%; P.R. .52) e os que nunca ou pouco se expressam em crioulo (7,2%; P.R. .41).

Deve-se salientar que os resultados expostos, à exceção da última variável comentada, são muito similares aos que se encontram em análises de outras variedades urbanas não europeias do Português, como a brasileira (BRANDÃO, 2013) e a moçambicana (BRANDÃO, 2018), embora, por vezes, em diferente ordem de importância.

ANALISANDO AS ESTRUTURAS PREDICATIVAS/PASSIVAS (EPPS) NO PST

Procedimentos metodológicos

A pesquisa enquadra-se na linha sociolinguística de inspiração laboviana (LABOV, 1972, 1994, 2001, 2003) e tem por finalidade determinar, com apoio

no programa Goldvarb-X, as restrições estruturais e/ou sociais que condicionam a não implementação da marca de plural nas EPPs em dados eliciados de 22 entrevistas realizadas em 2009 e pertencentes ao referido *Corpus VAPOR*, as mesmas que fundamentaram o estudo de Brandão (2011b) sobre a concordância entre constituintes do SN.

Os informantes estão distribuídos por sexo/gênero, três faixas etárias (18-35, 36-55 e 56-75 anos) e dois níveis de escolaridade (fundamental e médio), dois por célula, à exceção da faixa C, em que só há um informante masculino de nível fundamental e uma mulher de nível médio de escolaridade.

Controlou-se, ainda, a variável *frequência de uso de um crioulo* (em geral, o Forro) para verificar se a sua maior ou menor frequência implicaria maior ou menor probabilidade de não marcação de plural. A variável ficou constituída por 3 fatores: Frequência 0 (F0), zero ou baixa, – os indivíduos expressam-se fundamentalmente em Português; Frequência 1 (F1), média – os indivíduos expressam-se em Português, mas dominam um crioulo e dele fazem uso eventual; Frequência 2 (F2), alta – os indivíduos, embora falem o Português, se expressam, regularmente, num crioulo.

Além das quatro variáveis extralinguísticas já descritas e retratadas no Quadro 1, controlaram-se as nove variáveis estruturais consideradas por Scherre (1991), algumas delas já comentadas na seção 1: *estrutura do predicativo; paralelismo formal das sequências de predicativos/participios no discurso; processos morfofonológicos de formação de plural; tonicidade do item no singular; características formais do sujeito da construção; características formais do verbo da construção; ordem dos elementos na estrutura; material interveniente entre o verbo e o participio/predicativo; tipo de estrutura.*

Nas 22 entrevistas, cuja duração varia de 25 a 55 minutos, obtiveram-se apenas 75 dados, o que se deve à baixa produtividade, em entrevistas sociolinguísticas, dos tipos de estruturas aqui retratadas. Basta dizer que, na fala do mesmo conjunto de informantes, ocorreram 1.259 SNs, uma média de 57 por indivíduo (BRANDÃO, 2011b). No Quadro 1, que apresenta a distribuição dos informantes segundo as variáveis extralinguísticas consideradas, observa-se que duas das células não foram integralmente preenchidas: faixa C, nível 1 do sexo masculino e faixa C do nível 2, do sexo feminino, que só contam com 1 informante.

Quadro 1 – São Tomé (ST): distribuição dos informantes conforme as variáveis extralinguísticas

Escolaridade	Nível fundamental (1)		Nível médio (2)	
	Homem (h)	Mulher (m)	Homem (h)	Mulher (m)
Sexo				
Faixa etária				
18-35 anos A	ST-A-1-h-a (F1-A)	ST-A-1-m-a (F1-B)	ST-A-2-h-a (F1-C)	ST-A-2-m-a (Fø -D)
	ST-A-1-h-b (F1-E)	ST-A-1-m-b (F2-F)	ST-A-2-h-b (Fø -G)	ST-A-2-m-b (Fø -H)
36-55 anos B	ST-B-1-h-a (F1-I)	ST-B-1-m-a (Fø -J)	ST-B-2-h-a (F1-K)	ST-B-2-m-a (F1-L)
	ST-B-1-h-b (F1-M)	ST-B-1-m-b (F1-N)	ST-B-2-h-b (Fø -O)	ST-B-2-m-b (Fø -P)
56-75 anos C	ST-C-1-h-a (F2-Q)	ST-C-1-m-a (F1-R)	ST-C-2-h-a (Fø -S)	ST-C-2-m-a (Fø -T)
	-----	ST-C-1-m-b (F2-V)	ST-C-2-h-b (F1-X)	-----
Ao final do código do informante, a e b individualizam o informante da mesma célula.				
Entre parênteses, encontra-se a classificação segundo a variável <i>frequência de uso de um crioulo</i> (Fø = baixa; F1 = média; F2 = alta) e o código individual do falante.				

Fonte: Brandão (2011b), com alterações.

Em estudo anterior (SOARES, 2021), com 17 informantes (os doze sombreados no Quadro 1 e os 5 de nível superior que foram considerados na primeira análise do SN), o número de dados chegou a 87, 47 deles produzidos pelos indivíduos de nível superior, registrando-se apenas dois de não marcação: “não são grandes coisa” (ST-A-3-h-a) e “pra nós por exemplo que somos crescido” (ST-B-3-h-a), o que permite classificar a regra de concordância nesse segmento social, como semicatórica (LABOV, 2003), constatação que não surpreende tendo em vista que dele fazem parte os indivíduos que têm mais acesso aos bens culturais e viajam para Portugal com certa regularidade, estando, portanto, mais expostos à norma europeia do Português.

Na análise de Soares, inclusive, a variável *nível de escolaridade* foi a mais saliente para a (não)aplicação da regra (cf. Tabela 1), o que determinou, mais uma vez, a opção por se observar, nesta nova etapa de pesquisa, apenas indivíduos de níveis médio e fundamental (estes últimos com alto índice de não concordância – P. R. .97), como se expõe na Tabela 1.

Tabela 1 – Atuação da variável nível de escolaridade para a não marcação de número em EPPs no PST (SOARES, 2021)

Nível de escolaridade	Apl./Oc.	%	PR
Fundamental	11/15	73,3	.97
Médio	5/25	20	.56
Superior	2/47	4,3	.21
Input: .06	Significância: .000		

Fonte: Soares (2021, slide 8).

RESULTADOS

Antes de passar aos resultados, é importante dizer que dos 22 informantes, quatro – ST-A-1-m-a, ST-B-1-m-a, ST-B-2-m-a e ST-C-1-m-b – não apresentaram nenhuma estrutura passiva/predicativa em, respectivamente, 25,5; 23,14; 25,18 e 14,2 minutos de entrevista, as mais breves do acervo. Por outro lado, o número de ocorrências por indivíduo variou de um a dez, não necessariamente em função do tempo de gravação, mas dos temas focalizados.

A análise dos dados submetidos ao programa Goldvarb-X indicou que, no cômputo geral, na fala popular urbana de São Tomé, a ausência da marca de número em EPPs incide em 25 das 75 ocorrências (33,3%), sendo condicionada pela atuação das variáveis *nível de escolaridade*, *características formais do sujeito da construção* e *frequência de uso de um crioulo*.

Tendo em vista, de um lado, a reduzidíssima amostra e, de outro, o número expressivo de variáveis, foi necessário agrupar fatores (como foi o caso em relação às segunda e terceira variáveis selecionadas), e realizar uma análise mais pontual, de caráter qualitativo.

No que tange ao *nível de escolaridade*, também selecionada em primeiro lugar no estudo de Soares (2021), são os indivíduos de nível fundamental os que menos aplicam a marca, como se constata na Tabela 2.

Tabela 2 – Atuação da variável *nível de escolaridade* para a não marcação de número em EPPs na fala de indivíduos de níveis fundamental e médio de escolaridade

Nível de escolaridade	Apl./Oc.	%	PR
Fundamental	17/29	58,6	.695
Médio	8/46	17,4	.373
Input: .395		Significância: .001	

Fonte: Autoral

A segunda variável selecionada, de caráter estrutural, era inicialmente composta por 8 fatores, dois a mais do que os sugeridos por Sherre (1991). A detida observação do número de dados e das frequências de marcação e não marcação de plural no âmbito desse grupo autorizou que se reduzissem os fatores a apenas dois, como se indica na Tabela 3.

Tabela 3 – Reorganização da variável *características formais do sujeito da construção*

Novos fatores	Antigos fatores	OCOs
(1) Sujeito explícito marcado	(a) Sujeito explícito com todos os elementos nominais flexionáveis marcados	33
	(b) Sujeito explícito com marca semântica de plural (numerais)	2
	(c) Sujeito explícito pronominal (nós/eles/elas)	4
	(d) Sujeito explícito com os últimos elementos com marca	0
(2) Sujeito (não)explícito não marcado	(a) Sujeito explícito sem a(s) última(s) marca(s) formal(ais) de plural	2
	(b) Sujeito zero	29
	(c) Sujeito representado por pronome relativo	2
	(d) Sujeito explícito com marca formal de plural totalmente neutralizada	3

Fonte: Autoral

Os exemplos a seguir, ilustram os casos referidos na Tabela 3.

1-a. *os filhos* estão...estão separados de mim (A-1-h-a)

1-b. *dois* era meu (B-1-m-b)

1-c. *nós* somos escrava deles (A-1-m-b)

2-a. é que *meus filho* foi batizado no católico (A-1-h-4-a)

2-b. em que hoje em dia [] estamos submetido (A-2-h-a)

2-c. ajudar as crianças *que* é crianças de zonas urbanas (C-2-h-b)

2-d. *elas* são professoras (C-1-m-a)

Para a definição dessa variável remodelada, partiu-se da hipótese de que maior seria a tendência ao cancelamento da marca de número nas EPPs quanto mais o sujeito (elemento à sua esquerda) apresentasse marca morfológica ou semântica, o que acabou por se confirmar (cf. Tabela 4).

Tabela 4 – Atuação da variável *características formais do sujeito da construção* para a não marcação de número em EPPs na fala de indivíduos de níveis fundamental e médio de escolaridade

Fatores	Apl./Oco	%	P. R.
Sujeito explícito marcado	20/39	51,3	.751
Sujeito (não)explícito não marcado	5/36	13,9	.232
Input: .395	Significância: .001		

Fonte: Autoral

O cruzamento dessas duas variáveis, é também elucidativa quanto a serem os falantes menos escolarizados aqueles que lideram o cancelamento da marca se esta já está explícita à esquerda (13/16, 81%), o que vai ao encontro do que se observa no âmbito do SN, em que a marca de plural tem no *locus* esquerdo seu contexto por excelência, sendo menos frequente quanto mais à direita.

Tabela 5 – Não marcação de plural: cruzamento das variáveis nível de escolaridade e *características formais do sujeito da construção*

	Nível Fundamental		Nível Médio	
	Apl/Ocos	%	Apl/Ocos	%
Sujeito explícito marcado	13/16	81%	7/23	30%
Sujeito (não)explícito não marcado	4/13	30%	1/23	4%

Fonte: Autoral

Outra variável que se tem mostrado significativa nas análises do PST (vide, anteriormente, a síntese sobre sua atuação no SN) é *frequência de uso de um crioulo*, composta por três fatores que buscam aquilatar o maior ou menor contato do indivíduo com outras línguas, em especial, o Forro. No entanto, cabe fazer, à luz do Quadro 1, uma reflexão sobre os indivíduos que concorreram para o resultado exposto a seguir, tendo em vista que alguns informantes não apresentaram ocorrências das estruturas aqui focalizadas.

Nenhum dos informantes que apresentou dados de EPPs e declarou falar fundamentalmente o Português (Fø), deixou de implementar a marca de plural (0/24

ocorrências). A variação se encontra entre os indivíduos que, embora falantes de Português, fazem uso eventual de um crioulo (F1) ou se expressam regularmente numa dessas línguas (F3), conforme se expõe na Tabela 6.

Tabela 6 – Atuação da variável *frequência de uso de um crioulo* para a não marcação de número em EPPs na fala de indivíduos de níveis fundamental e médio de escolaridade

Fatores	Apl./Oco	%	P. R.
Frequência 1 (F1), média	20/44	45,5	.492
Frequência 2 (F2), alta	5/7	71,4	.553
Input: .395	Significância: .001		

Fonte: Autoral

Os pesos relativos estão no limiar da neutralidade, o que se deve basicamente ao fato de apenas os informantes ST-C-1-h-a e ST-C-1-m-b, um homem e uma mulher do segmento mais velho e do nível fundamental, serem os responsáveis pelo índice referente à **alta** frequência de uso de um crioulo, enquanto a frequência média está representada por 7 informantes.

Para dar uma visão mais ampla sobre as estruturas que constam do *corpus* cabe mencionar, em termos percentuais, dentre as variáveis não selecionadas, resultados referentes às duas outras de caráter social e algumas de cunho estrutural.

A variável *sexo*, que foi selecionada no trabalho de Soares (2021), mostrou que são os homens os principais implementadores da regra de não marcação (19/48 ocos: 39,6%), em relação às mulheres (6/27 ocos: 22,2%), corroborando o resultado anterior. No que se refere à *faixa etária*, o quadro é de variação estável, com os indivíduos mais jovens (16/39 ocos: 41%) e os mais velhos (6/17 ocos: 35,3%), apresentando índices aproximados, em contraposição aos de idade mediana com 15,8% (3/19 ocos) de não marcação.

Todos os 75 dados estão na ordem canônica. Quanto à *estrutura do predicativo*, a maior parte delas é formada por adjetivos (35 ocos) seguidos por 20 formas participiais e 13 predicativos nominais de um elemento, que apresentaram, respectivamente, 34,3%, 40% e 30,8% de não aplicação da marca. As demais 7 ocorrências correspondem a 3 casos de predicativo pronominal/quantificador (com uma ocorrência de ausência de marca), 1 de predicativo de mais de um elemento com os itens *todo/tudo* e 3 de predicativos de mais de um elemento, nos quais a presença de marca foi categórica.

O efeito do *paralelismo formal*, que se pretendia observar, foi prejudicado pelo fato de 58 das 75 ocorrências (77,3% da amostra) constituírem construções

isoladas, exatamente aquelas em que ocorreu o maior número de não aplicação da marca (21 ocos, 32,6%).

Nos 8 casos de primeira construção de uma série, apenas uma não apresentou a marca (12,5%); já nas ocorrências precedidas de predicativo/particípio com todas as marcas de plural ou sem elas, respectivamente, duas (25%) e uma (100%) apresentaram cancelamento.

Outra variável sempre observada em análises sobre concordância nominal ou verbal é a *saliência fônica* que parte do princípio de que quanto menor a diferença fônica entre as formas singular e plural, maior a tendência a cancelar a marca de número. Na amostra, 66 dados correspondem a formas regulares de plural, com 22 ocorrências (33,3%) de ausência de marca. Entre os 9 casos de formação irregular, 3 de não marcação (também 33,3%), encontram-se duas ocorrências de *fácil* (ambas sem marca); duas de *difícil* (uma com marca, outra sem marca), duas de *morto* e duas de *português*, estas com as marcas de plural (*m[ɔ]rtos, portugueses*).

A última variável que cabe comentar diz respeito à forma do verbo: construções que se seguem a verbos no plural (67 ocos) ou no singular (8 ocos) apresentam, respectivamente, 19 (28,4%) e 6 (75%) casos de ausência de marca, o que, de certa forma, sugere, em certo grau, a atuação do princípio de *coesão estrutural* apontado por Lucchesi (2000, *apud* ANTONINO, 2012, p. 93-95): verbo com marca de plural → maior tendência à EPP com marca de plural; verbo sem marca de plural → maior tendência à EPP sem marca de plural).

COMPARANDO E COMPLEMENTANDO ANÁLISES: A NÃO MARCAÇÃO DE PLURAL EM EPPS E EM SNS NO PST

Para complementar a análise das EPPs, alicerçada em um diminuto número de dados (75), procedeu-se a uma nova observação da amostra de Brandão (2011b), por informante, desta vez levando em conta cada SN como um todo por número de constituintes e indicando os que apresentavam pelo menos uma marca de não aplicação do plural. Verificou-se que o *corpus* desse estudo era constituído de um total de 1.251 SNs, sendo 1.061 (84,%) de 2 constituintes, 168 (13,4%), de três e 22 de quatro ou cinco (1,75%).

Assim como nas amostras de modalidade falada de quaisquer variedades do Português, na amostra em foco o número de SNs de três ou mais constituintes é escasso, predominando os de dois constituintes, sendo também reduzido, como já se observou, o número de EPPs em entrevistas de perfil sociolinguístico. O pequeno número de dados, no entanto, permite a observação não apenas quantitativa, mas

também qualitativa do fenômeno, em especial em relação à performance de cada indivíduo que concorreu para a amostra.

Retomando o Quadro 1, agora organizado em duas tabelas, expõem-se os índices de não marcação de número por informante de nível fundamental (Tabela 7) e médio (Tabela 8) de instrução nas EPPs aqui focalizadas e nos SNs que embasaram o referido estudo de Brandão (2011b), considerados não atomisticamente (os espaços tracejados indicam que não se obtiveram ocorrências).

Tabela 7 – Índices de não marcação de número em EPPs e SNs (segundo o número de seus constituintes) na fala de 11 indivíduos de nível fundamental de instrução (PST)

Nível 1 (Fundamental)						
Faixa	Sexo	Informante	EPP	SN-2	SN-3	SN-4
A	H	ST-A-1-h-a	7/9 = 77,8%	18/95 = 18,9%	12/17 = 70,5%	3/4 = 75%
		ST-A-1-h-b	1/1 = 100%	12/33 = 36,6	3/4 = 75%	1/1 = 100%
	M	ST-A-1-m-a	----	24/44 = 54,5%	0/6 = 0%	----
		ST-A-1-m-b	3/5 = 60%	39/57 = 68,4%	8/12 = 66,6%	2/2 = 100%
B	H	ST-B-1-h-a	1/1 = 100%	16/19 = 84,2%	-----	----
		ST-B-1-h-b	0/3 = 0%	6/29 = 20,6%	1/3 = 33,3%	----
	M	ST-B-1-m-a	----	11/29 = 37,9%	0/4 = 0%	----
		ST-B-1-m-b	2/4 = 50%	21/63 = 33,3%	7/10 = 70%	1/1 = 100%
C	H	ST-C-1-h-a	2/2 = 100%	28/43 = 65,1%	2/3 = 66,6%	----
	M	ST-C-1-m-a	1/4 = 25%	3/20 = 15%	0/5 = 0%	----
		ST-C-1-m-b	----	7/11 = 63,6%	----	----
Totais parciais			17/29	185/443	33/39	7/8
			58,6%	41,7%	84,6%	87,5%

Fonte: Para os SNs, Brandão (2011b).

Tabela 8 – Índices de não marcação de número em EPPs e em SNs (segundo o número de seus constituintes) na fala de 11 indivíduos de nível médio de instrução (PST)

Nível 2 (Médio)						
Faixa	Sexo	Informante	EPP	SN-2	SN-3	SN-4
A	H	ST-A-2-h-a	5/10 = 50%	12/107 = 11,2%	1/25 = 4%	0/4 = 0%
		ST-A-2-h-b	0/5 = 0%	2/38 = 5,2%	1/4 = 25%	----
	M	ST-A-2-m-a	0/6 = 0%	5/50 = 10%	1/12 = 8,35	1/3 = 33,3%
		ST-A-2-m-b	0/3 = 0%	1/37 = 2,7%	1/10 = 11,1%	----
B	H	ST-B-2-h-a	0/5 = 0%	2/33 = 6%	0/3 = 0%	0/1 = 0%
		ST-B-2-h-b	0/3 = 0%	0/24 = 0%	0/4 = 0%	0/1 = 0%
	M	ST-B-2-m-a	----	2/76 = 2,6%	1/9 = 11%	2/3 = 66,6%
		ST-B-2-m-b	0/2 = 0%	2/59 = 3,3%	1/7 = 14,2%	----

C	H	ST-C-2-h-a	0/2 = 0%	3/82 = 3,6%	3/17 = 17,6%	----
		ST-C-2-h-b	3/7 = 42,9%	7/66 = 10,6%	3/10 = 30%	1/1 = 100%
	M	ST-C-2-m-a	0/2 = 0%	1/46 = 2,1%	0/3 = 0%	0/2 = 0%
Totais parciais			8/46	37/618	12/104	4/15
			17,4%	5,9%	11,5%	26,0%

Fonte: Para os SNs, Brandão (2011b).

A observação de cada tabela demonstra, em seus totais parciais, que, em qualquer das estruturas focalizadas, são os indivíduos de nível fundamental de instrução os que mais concorrem para a não aplicação da marca de plural, quer no âmbito das EPPs (58,6%), quer no dos SNs (SN-2: 41,7%; SN3: 84,6%; SN 4/5 = 87,5%).

Entre os falantes de nível médio, cujas EPPs apresentam 17,4% de cancelamento da marca, os índices relativos aos SNs mantêm a mesma escala progressiva de cancelamento a depender do número crescente de constituintes do SN: SN2: 5,9%; SN3: 11,5%; SN3: 26%.

Por outro lado, no que se refere especificamente às EPPs:

- (a) apenas um indivíduo de nível fundamental – ST-B-1-h-b – aplicou a marca de número nas 3 ocorrências com que contribuiu para a amostra;
- (b) entre os 10 informantes de nível médio que apresentaram ocorrência de EPPs, oito aplicaram a marca categoricamente (ST-B-2-h-b, inclusive, em todos os tipos de SNs) e apenas dois a cancelaram – ST-A-2-h-a (50%) e ST-C-2-h-b (42,9%), com índices bastante semelhantes.

CONSIDERAÇÕES FINAIS

Neste estudo, apesar do reduzido número de dados referentes às EPPs, pôde-se verificar que, na fala de São Tomé, a marcação de plural é variável, corroborando os resultados obtidos na análise no âmbito do SN, em que fatores de ordem estrutural e social são atuantes, entre os últimos destacando-se os relacionados a *nível de instrução e frequência de uso de um crioulo*, a primeira e a terceira variáveis selecionadas na análise variacionista. Fica claro que são os indivíduos que menos frequentaram a escola os que menos implementam a marca de plural, o mesmo ocorrendo com aqueles que têm maior contato com um crioulo.

Quanto à aferição dos condicionamentos estruturais, mostrou-se relevante a variável *características formais do sujeito da construção*, selecionada em segundo lugar no estudo de Scherre (1991), embora sem que se observasse a ação

do paralelismo formal. A reorganização dos fatores da variável indica que há maior probabilidade de não aplicação da marca quando o sujeito está explícito e apresenta marca semântica e/ou morfológica de pluralidade.

Análises sobre a concordância no SN em variedades africanas do Português têm demonstrado comportamento similar ao que se constata em diferentes variedades do PB, quanto aos padrões observados e às variáveis que os condicionam, devendo-se salientar que as de cunho social (sobretudo *nível de escolaridade*) sempre se mostram altamente significativas, retratando o que, em relação ao PB, Lucchesi (2015, entre outros de seus trabalhos) tem denominado de polarização sociolinguística, reflexo do contato multilíngue que ocorreu (e ainda ocorre) na fase da colonização.

Embora a norma de referência do PST seja, em termos oficiais, o Português Europeu, em que a regra de concordância no SN, na variedade urbana, é categórica, como sugere Brandão (2013), a nascente norma santomense dela se afasta, certamente em função da história sociolinguística do país, marcada pelo contato multilíngue, em que o Forro, certamente, representa papel importante.

Em trabalhos de Hagemer (2009, 2018), bem como em Araújo (2020), há indicações sobre como se vêm delineando (ao que tudo indica, de forma acelerada) as normas que regem o uso do Português em São Tomé, cada vez mais adotado em detrimento dos crioulos, sobretudo do Forro, a língua antes mais difundida e caracterizada por alguns informantes da pesquisa como “a língua nacional”.

Hagemer (2009), que já havia comentado a diversidade de normas que coexistem no PST (cf. citação no item referente à caracterização da comunidade de São Tomé), sintetiza (2018, p. 182) o quadro atual, bem diferente daquele de um “passado recente em que o português era L2 para a maioria dos habitantes das ilhas”:

Os anos coloniais lançaram as bases para o boom do Português, após a independência em 1975. O Português tornou-se o incontroverso idioma oficial, que é usado na administração, educação, mídia, etc. A maior mobilidade social aumentou a exposição a esta língua, que se tornou amplamente utilizada em domínios formais e informais. Como consequência, o número de falantes de Português como L1 cresceu rapidamente e levou a um aumento do monolinguismo em Português. Este caso de mudança massiva para a antiga língua colonial e sua consequente nativização é atípica no contexto africano, mas também pode ser observado em outras ex-colônias portuguesas em África, em particular em Angola e, a um ritmo mais lento, em Moçambique.⁸

⁸ The colonial years laid the foundations for the boom of Portuguese, following independence in 1975. Portuguese became the uncontroversial official language, which is used in administration, education, the media etc. Higher social mobility increased exposure to this language, which

Apesar dos parcos (*maus*) dados no que tange às EPPs, deles buscou-se fazer o melhor uso possível como uma contribuição para o conhecimento da dinâmica da (não)marcação de plural no PST e, por extensão, na das demais variedades não europeias do Português, que, quanto aos temas aqui abordados, parecem apresentar significativas convergências, entre as quais as que dizem respeito ao contato multilinguístico como um dos fatores determinantes da construção de suas normas.

REFERÊNCIAS

- ALMEIDA, Evanilda Marins. *A variação da concordância nominal num dialeto rural*. 1997. 145 f. Dissertação (Mestrado em Letras Vernáculas). Faculdade de Letras, UFRJ, Rio de Janeiro, 1997.
- ANTONINO, Vivian. *A concordância nominal no predicativo do sujeito e estruturas passivas no português popular no interior do estado da Bahia*. 2010. 121 f. Dissertação (Mestrado em Letras e Linguística). Instituto de Letras, Universidade Federal da Bahia, Salvador, 2010.
- ANTONINO, Vivian. *Português popular de Salvador: uma análise da concordância nominal em predicativos do sujeito e em estruturas passivas*. 2012. 189f. Tese (Doutorado em Letras e Linguística). Instituto de Letras, Universidade Federal da Bahia, Salvador, 2012.
- ARAUJO, Gabriel Antunes. Portuguese language expansion in São Tomé and Príncipe: an overview. *Diadorim*, Rio de Janeiro, v. 22, n° 1, p. 57-78, 2020.
- BRAGA, Maria Luiza. *A concordância de número no sintagma nominal no Triângulo Mineiro*. 1978. Dissertação (Mestrado em Letras). Departamento de Letras e Artes, Pontifícia Universidade Católica-RJ, Rio de Janeiro, 1978.
- BRANDÃO, Silvia Figueiredo. Concordância nominal em duas variedades do português: convergências e divergências. *Revista Veredas*, v. 15, n. 1, p. 164-178, 2011a.
- BRANDÃO, Silvia Figueiredo. *O cancelamento da marca de número nominal na variedade urbana não standard do Português de São Tomé*. Comunicação

became widely used in both formal and informal domains. As a consequence the number of L1 Portuguese speakers grew fast and led to increasing monolingualism in Portuguese. This case of massive shift to the former colonial language and its consequent nativization is atypical in the African context, but can also be observed in other former Portuguese colonies in Africa, in particular in Angola and, at a slower pace, in Mozambique.

apresentada ao XVI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina (ALFAL). Alcalá de Henares, 6-9 de junho, 2011b.

BRANDÃO, Silvia Figueiredo. Patterns of agreement within the Noun Phrase. *Journal of Portuguese Linguistics*, v. 12, n. 2, p. 51-100, 2013.

BRANDÃO, Silvia Figueiredo. Concordância nominal no Português de São Tomé e no Português de Moçambique. In: BRANDÃO, Silvia Figueiredo. *Duas variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher, 2018. p. 203-244.

BRANDÃO, Silvia Figueiredo; VIEIRA, Silvia Rodrigues. The agreement continuum in urban samples of African, Brazilian and European varieties of Portuguese. In: LÓPEZ, Laura Álvarez; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de. *The Portuguese language continuum in Africa and Brazil*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2018. p. 267-289.

ESPÍRITO SANTO, Carlos. O crioulo Forro. *Revista Camões*, Lisboa, v. 1, p. 54-59, 1998.

FURTADO, Bárbara Amaral de Andrade. *A concordância de número em predicativos do sujeito: variação linguística em Fortaleza*. 2017. 153 f. Dissertação (Mestrado em Linguística) Programa de Pós-Graduação em Linguística, Universidade Federal do Ceará, Fortaleza, 2017.

GONÇALVES, Rita; HAGEMEIJER, Tjerk. O português num contexto multilíngue: o caso de São Tomé e Príncipe. *Revista Científica UEM: Série Letras e Ciências Sociais*, Maputo, v. 1, n. 1, p. 87-107, 2015.

GUY, Gregory Riordan. *Linguistic variation in Brazilian portuguese: aspects of the Phonology, Syntax, and Language History*. 1981. 391 f. Ph.D. dissertation. (Ph.D in Linguistics) Philadelphia: University of Pennsylvania, 1981.

HAGEMEIJER, Tjerk. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, Macau, v.1, n.1, p. 1-27, 2009.

HAGEMEIJER, Tjerk. From creoles to Portuguese: Language shift in São Tomé and Príncipe. In: LÓPEZ, Laura Álvarez; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de. *The Portuguese language continuum in Africa and Brazil*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2018. p. 169-184.

INSTITUTO NACIONAL DE ESTATÍSTICA – INE, São Tomé e Príncipe. Disponível em: <https://www.ine.st/>. Acesso em: 15 abr. 2012.

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

LABOV, William. *Principles of linguistic change*. V. I: Internal factors. Oxford: Blackwell, 1994.

LABOV, William. *Principles of linguistic change*. V. II: Social factors. Oxford: Blackwell, 2001.

LABOV, William. Some sociolinguistic principles. In: PAULSTON, Christina Bratt; TUCKER, Richard. *Sociolinguistics: the essential readings*. Massachusetts: Blackwell Publishing, 2003. p. 234-250.

LUCCHESI, Dante. A concordância nominal em estruturas passivas e de predicativo do sujeito em comunidades rurais afro-brasileiras isoladas no contexto da história sociolinguística do Brasil. In: VOTRE, Sebastião; RONCARATI, Cláudia. *Anthony Naro e a linguística no Brasil: uma homenagem acadêmica*. Rio de Janeiro: 7 Letras; FAPERJ, 2008. p. 148-168.

LUCCHESI, Dante. A concordância em estruturas passivas e de predicativo do sujeito. In: LUCCHESI, Dante; BAXTER, Alan.; RIBEIRO, Ilza. *O Português afro-brasileiro*. Salvador: EDUFBA, 2009. p. 373-387.

LUCCHESI, Dante. *Língua e sociedade partidas: a polarização sociolinguística do Brasil*. São Paulo: Contexto, 2015.

NASCIMENTO, Fabiane Rodrigues. São Tomé e Príncipe: aspectos históricos, econômico-sociais e linguísticos. In: BRANDÃO, Sílvia Figueiredo. *Duas variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher, 2018. p. 43-73.

SALOMÃO, Mircia Hermenegildo. A marcação de pluralidade no SN em contexto predicativo do noroeste paulista. *Estudos Linguísticos*, São Paulo, v. 39, n. 2, p. 671-685, maio-ago. 2010.

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

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

SCHERRE, Maria Marta Pereira. A concordância de número nos predicativos e nos participios passivos. *Organon*, Porto Alegre, v. 5, n. 18, p. 52-70, 1991.

SCHERRE, Maria Marta Pereira. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua portuguesa*, Lisboa, v. 12, p. 37-49, 1994.

SCHERRE, Maria Marta Pereira. *Doa-se lindos filhotes de poodle: variação linguística, mídia e preconceito*. São Paulo: Parábola, 2005.

SOARES, Paulo Vitor Lima da Gama. *Concordância de número em estruturas predicativas/passivas no Português de São Tomé: resultados preliminares*. Comunicação apresentada à 42ª Jornada Giulio Massarani de Iniciação Científica, Tecnológica, Artística e Cultural. Rio de Janeiro, Universidade Federal do Rio de Janeiro, março de 2021.

VAZZATA-DIAS, Juçá Fialho. *A concordância de número nos predicativos e participios passivos na fala da região sul: um estudo variacionista*. 1996. 157 f. Dissertação (Mestrado em Letras/ Linguística). Universidade Federal de Santa Catarina, Florianópolis, 1996.

VAZZATA-DIAS, Juçá Fialho. A concordância de número nos predicativos/participios passivos na fala do Sul do Brasil – motivações extralinguísticas. *Letras de Hoje*, Porto Alegre, v. 35, n. 1, p. 209-228, 2000.

VAZZATA-DIAS, Juçá Fialho; FERNANDES, Marisa. A inter-relação da concordância nominal e da concordância nos predicativos/participios passivos, sob o enfoque da Teoria da Variação e Mudança Linguística. *Organon*, Porto Alegre, v. 14, n. 28-29, p. 115-131, 2000.

VIEIRA, Silvia Rodrigues; BAZENGA, Aline. Patterns of third person verbal agreement. *Journal of Portuguese Linguistics*, Lisbon, v. 12, n. 2, p. 7-50, 2013.

WEINREICH, Uriel; LABOV, William; HERZOG, Martin. Empirical foundations for a theory of linguistic change. In: LEHMANN, Winfred.; MALKIEL, Yakov. *Directions for historical linguistics*. Austin: University of Texas Press, 1968. p. 97-195.

