

Uso variável do artigo definido no português da comunidade de Almojarife, São Tomé¹

Carlos Filipe Guimarães Figueiredo

Universidade de Macau²

O uso variável do artigo definido no sintagma nominal do português da comunidade de Almojarife, São Tomé, apresenta respaldos no substrato deste, o crioulo santome (Figueiredo 2010, 2014). Fazem-se aqui comparações com similares perfis de variação do português dos Tongas, Roça Monte Café, São Tomé, do dialeto rural de Helvécia, Bahia, e das falas culta e popular da cidade de Salvador, também Bahia. O estudo contribui com mais dados para o conjunto de variáveis que testemunham a reestruturação morfossintática quer de variedades de português de África quer de variedades de português do Brasil.

Palavras-chave: Aquisição e uso variável do artigo definido; português de Almojarife, São Tomé; falas cultas e populares do Brasil; crioulos atlânticos de base portuguesa.

1. Introdução

Observações comparadas acerca do artigo definido no português dos Tongas, Roça Monte Café, São Tomé (PT) (Baxter & Lopes 2005, 2006) e variedades de português faladas no Brasil, urbanas (Silva 1996; Baxter & Lopes 2004a) ou rurais (Ferreira 1984; Baxter & Lopes 2005), levam Baxter & Lopes (2009: 330) a inferir que o seu uso no português rural de Helvécia, Bahia (PHEL), reflete aquisição diacrónica em direção ao artigo pleno, por influência de variedades do Brasil mais próximas da norma padrão.

Paralelamente, os autores apontam semelhanças da tipologia do SN sem artigo do PHEL relativamente aos crioulos de Cabo Verde e de São Tomé (neste

¹ Este trabalho resulta da reflexão e expansão de uma apresentação no Encontro Conjunto do XIII Encontro da ACBLPE e VII Encontro da ABECS, 30 de julho a 03 de agosto de 2012, Universidade de São Paulo, Brasil. O autor agradece o apoio da Universidade de Macau para a apresentação, bem como as sugestões dos pareceristas, isentando estes de quaisquer imprecisões que, eventualmente, persistam no trabalho.

² <https://fah.umac.mo/staff/staff-portuguese/carlos-filipe-figueiredo-guimaraes/>

caso, o santome) e seus substratos nígero-congoleses.³ Quanto a Ferreira (1984), refere a possibilidade de o SN sem artigo do PHEL representar referência definida específica.

Considerando que as variedades de português de África permitem tanto analisar fenómenos emergentes do contato entre o português e línguas africanas como perceber a formação do próprio português do Brasil (PB), analisámos, com recurso ao GOLDVARB X (Sankoff, Tagliamonte & Smith 2005), a influência de variáveis linguísticas e extralinguísticas na realização de SN's do português da comunidade de Almojarife, São Tomé (PA),⁴ com ou sem artigo definido. A análise justifica-se dado observarem-se realizações com semelhanças tipológicas seja em relação às do substrato santome⁵ seja relativamente às de outros crioulos portugueses do Atlântico, estas últimas apontadas por Alexandre & Hagemeyer (2007) e Baptista (2007). Estabelecemos também comparações com resultados do PT (Baxter & Lopes 2004a), da norma culta de Salvador (NURC) (Baxter & Lopes 2004a) e do PHEL (Baxter & Lopes 2004a, 2004b, 2009), buscando contribuir com mais dados para o conjunto de variáveis que testemunham a reestruturação morfossintática quer do PA quer de variedades do português afro-indígena do Brasil.⁶ Os resultados permitem conclusões sobre: (i) uso do nome nu no PA,

³ Para designação dos grupos linguísticos, famílias e línguas africanas, adoptou-se a proposta ortográfica de Fiorin & Petter (2008: 13).

⁴ O PA é falado pela comunidade de forros estabelecida na parte sul da Baía de Almojarife, zona Este da ilha de São Tomé. Trata-se de uma variedade de português resultante do contato com o santome e sujeita a posterior *transmissão linguística irregular*, nos séculos XIX/XX (Figueiredo 2010: 85-87).

⁵ À data da coleta dos dados deste trabalho, o santome era uma língua de comunicação generalizada em São Tomé, falada por 72,4% da população (Instituto Nacional de Estatística de São Tomé e Príncipe, 2003/2005). Os falantes entrevistados eram todos bilingues em português e santome, mas os dados coletados revelam que as gerações mais novas, por preconceito linguístico, já estigmatizavam fortemente este crioulo (Figueiredo 2010: 92-93), contribuindo também para a situação que o coloca atualmente em perigo de extinção. O santome, embora detenha léxico maioritariamente de origem portuguesa, possui características morfossintáticas que estabelecem paralelismos com as L1's africanas faladas pelos ancestrais escravizados africanos transplantados (Hagemeyer 1999, 2009), ou seja, falantes dos grupos Cuá (Benin) e Banto (Guiné, Gabão e Angola) (Ferraz 1976, 1979). Recentes pesquisas evidenciam que o proto-crioulo do Golfo da Guiné, a raiz dos quatro crioulos desta região (Hagemeyer 1999, 2009), exhibe traços quer do edóide quer dos idiomas bantóides da área H, especialmente das variedades congo (Lorenzino 1998; Hagemeyer 1999, 2007, 2009; Schang 2000; Tomás *et al.* 2002; Güldemann & Hagemeyer, este volume).

⁶ Define-se o português afro-indígena como parte do conjunto de variedades que formam o *continuum dialetal português vernacular brasileiro*, contendo o português quilombola num dos extremos e os falares urbanos não-padrão no outro extremo (Figueiredo & Oliveira 2013: 15).

comparativamente aos crioulos portugueses do Atlântico observados por Alexandre & Hagemeyer (2007) e Baptista (2007); (ii) uso variável do artigo definido no PA; (iii) possibilidade de o português adquirido por contato por falantes de um crioulo português do Atlântico manifestar variação idêntica ao português adquirido por falantes cuja língua ancestral não é um crioulo português (cf. Figueiredo 2010).

Na Seção 2, apresentaremos os informantes cujos dados linguísticos foram atomística e quantitativamente tratados. Nesta mesma seção, abordaremos também a metodologia aplicada, que possibilitou comparar as tendências de uso do artigo definido no PA com as de outras variedades de português de África e do Brasil. Para tanto, criaram-se variáveis dependentes e independentes (linguísticas e extralinguísticas), para as quais se constituíram fatores e se levantaram hipóteses a confirmar.

Na Seção 3, efetuaremos: (i) análise dos resultados; (ii) comparação destes com os obtidos em trabalhos de outras variedades de português; (iii) enquadramento justificativo das ocorrências em função de teorias apresentadas em vários estudos. Encerra-se o trabalho com conclusões gerais do estudo.

2. Metodologia e hipóteses

Sendo uma variedade reestruturada por *transmissão linguística irregular* (Figueiredo 2010: 98), o PA evidencia variação morfossintática ternária (cf. Lucchesi 2019) que, no caso do artigo definido, se traduz em:

(i) Omissão do determinante (nomes nus):

(1) PA:

E amigo meu tinha uma mulher. [OSVALH1]⁷

‘E o meu amigo tinha uma mulher.’ [PE]

(ii) Aplicação do determinante do português “o(s)/a(s)”:

(2) PA:

É os meus pai que me deu essa ideia. [OSVALH1]

‘Foram os meus pais que me deram essa ideia.’ [PE]

⁷ Neste trabalho, os informantes são identificados com um código, salvaguardando-se, assim, as suas identidades.

- (iii) Realização do determinante sob a forma do pronome pessoal do português, 3ª pessoa “ele(s)” (exemplos (3a) e (3b)), originando estruturas que estabelecem paralelismo semântico-formal com as do substrato santome:⁸

(3) PA:

- a. Dentro de dez minuto **ele seôr** já stá morto. [ABILH2]
[Dentro de dez minutos o senhor já está morto. – PE]
- b. Eu, quê zê, nós temos **eles** nosso documentos. [ABILH2]
[Eu, quer dizer, nós tínhamos os nossos documentos. – PE].

2.1. Caraterização dos informantes

O estudo incidiu, atomística e sincronicamente, sobre a língua-E de um *corpus* de 18 entrevistas recolhidas em Almojarife entre fevereiro e maio de 1998,⁹ tendo os informantes sido equitativamente repartidos por sexo e idade. A tabela do Anexo 1 fornece uma visão dos informantes.¹⁰

2.2. Variáveis constituídas

Considerando as dificuldades para constituição de amostras de perfil sociolinguístico em comunidades africanas, sobretudo por questões sócio-

⁸ Esta é uma das particularidades que faz com que o PA apresente distinções relativamente ao português de São Tomé. O uso do relativo do PA também exhibe variação ternária (Figueiredo, 2014), podendo ser realizado com “que”, “ku” e “com”:

1. PA:

dia **ku** ele tem, ele dá, mas dia **que** ele não tem, ele não tem mesmo [SALVH3]
‘nos dias (em) **que** ele tem, ele dá, mas nos dias (em) que ele não tem, ele não tem mesmo’ [PE] (Figueiredo 2014: 281, dado renumerado)

2. PA:

ê tem um rapaz mais grande **com** trabalha na Andamede [MINISM3]
‘eu tenho um rapaz mais velho **que** trabalha em Andamede’ [PE] (Figueiredo 2014: 283, dado renumerado)

Segundo o autor, o uso de “ku” configura transferência do relativo do substrato santome determinada por aspetos pragmáticos e sociopragmáticos, enquanto a forma “com” resulta de similaridades fónicas entre as três formas e que ocorreu durante o processo de reanálise e aquisição do relativizador “que” do português (Figueiredo 2014: 305).

⁹ O *corpus* faz parte do projecto “*Semi-creolization: testing the hypothesis against data from Portuguese-derived languages of São Tomé (Africa)*”, dirigido por Alan N. Baxter (University of Saint Joseph, Macau, China), que integrava, à época, os quadros de La Trobe University, Victoria, Austrália, e Universidade Federal da Bahia, Brasil. Na coleta colaboraram Dante Lucchesi (Universidade Federal da Bahia, Brasil) e Nara Barreto (La Trobe University, Victoria, Austrália, à época).

¹⁰ As restrições à extensão do trabalho levaram-nos a colocar as tabelas maiores como anexos.

culturais, estamos convictos que o número de informantes é suficiente para uma análise robusta e obtenção de resultados não-viciados. As variáveis dependente e independentes constituídas para análise constam da tabela do Anexo 2.

2.2.1. Variável dependente e caracterização dos dados

Para a variável dependente, codificaram-se SN's de referência definida com ausência (exemplo (1)) ou presença (exemplos (2) e (3)) de artigo definido. Desconsideraram-se SN's configurando locuções adverbiais correntes, como “às vezes”, visto representarem estruturas lexicalizadas e cunhadas por condicionalismos pragmáticos que viciariam os resultados.

Contrariamente, codificámos as estruturas lexicalizadas que não contêm artigo definido na norma padrão, mas que o registam no PA (exemplo (4a)), pois as mesmas são entendidas como desvio:

(4) PA:

- a. e, **do modo que**, eles também matou. [MAURIH2]
- b. ele não tem essas possibilidade de dirigir a moça, **de modo** ele procura de dizer aquilo que ele não é [CASTEH1]

2.2.2. Variáveis independentes constituídas e variáveis analisadas

Para se precisarem as motivações que determinam a ausência/inserção do artigo definido no PA, avaliou-se o efeito exercido pelas três variáveis independentes sociais, três variáveis independentes semânticas e sete variáveis independentes estruturais e funcionais listadas na tabela do Anexo 2. Constituídas a variável dependente binária e as variáveis independentes, elaboraram-se os fatores envasados nestas últimas e aos quais faremos referência ao longo do trabalho.

As restrições à dimensão do trabalho impõem que se observem unicamente as variáveis consideradas estatisticamente significativas pelo GOLDVARB X, analisando-se também apenas os resultados destas.

2.2.2.1. Variáveis independentes sociais

Em Almoxarife, depreende-se que os homens realizem menos variação na aplicação do artigo definido, devido à sócio-história da comunidade. À data da recolha dos dados, os informantes tinham nascido todos na época colonial portuguesa. Neste período, o maior contato dos homens com o mundo exterior à comunidade, preconizado pelos fluxos de migração para os centros urbanos, integração de mancebos no exército ou maior escolarização dos membros deste

sexo (Figueiredo 2010: 665-666) serão fatores que potenciam o uso do artigo, visto os homens deterem um contato mais próximo com as formas da norma padrão. Quanto às mulheres, mais fixadas na comunidade, deverão apresentar maior variação. Considerando-se estes aspetos, formulámos a seguinte hipótese para a variável *Sexo*: *devido às condicionantes sociolinguísticas de Almojarife, os falantes do sexo masculino apresentam tendência para realizarem menos variação do que as falantes do sexo feminino.*

Como os homens se ausentavam mais, não foi necessário constituir a variável *Estadia*, normalmente utilizada para constatar se os falantes que se afastam da comunidade por mais de seis meses produzem menos variação.

Para a variável *Escolaridade*, considerámos que a difusão dos modelos normativos facilita a mudança linguística, uma vez que expõe os falantes à norma culta, quer da fala quer da escrita. No caso de Almojarife, tal é relevante porque os dados são de 1998, tendo a quase totalidade dos informantes escolarizados (gerações mais novas) frequentado o ensino numa fase em que São Tomé e Príncipe era colónia portuguesa. Porém, após a independência desta, o ensino passou a ser ministrado por professores falantes do português de São Tomé, uma particularidade que consideraremos quando analisarmos os resultados.

Para esta variável, distribuámos os falantes pelos fatores: (i) Analfabetos; (ii) Escolarização baixa (frequência do ensino primário); (iii) Escolarização média (ensino primário concluído); (iv) Escolarização alta (frequência pós-primária). Pretende-se confirmar a hipótese: *quanto mais alta é a escolaridade, maior é a inserção de artigo definido no PA.*

A variável *Idade*, observada concomitantemente com outros padrões sociais (p.e. sexo, escolarização ou estrato social), permite detetar mudanças linguísticas em curso na gramática da comunidade, verificáveis no chamado *tempo aparente* (Labov 1972) ao contemplar-se um período de pelo menos 20 anos para cada faixa etária.

Para esta variável, constituímos os fatores: (i) faixa etária FE-1 (20/40 anos); (ii) faixa etária FE-2 (41/60 anos); (iii) faixa etária FE-3 (+60 anos). Foi também estabelecida a hipótese: *a idade e a aquisição da regra de inserção do artigo definido estabelecem uma relação proporcional, com as gerações mais novas (e mais escolarizadas) favorecendo o uso deste artigo.*

2.2.2.2. Variáveis independentes estruturais e funcionais

A constituição das variáveis estruturais e funcionais (todas linguísticas) prende-se com evidências em estudos sobre o SN e/ou nomes nus quer de crioulos

portugueses do Atlântico (Ferraz 1979; Castro & Pratas 2006; Baptista 2007; Alexandre & Hagemeyer; 2007) quer de variedades de português africano (Baxter 2004; Figueiredo 2008, 2010) quer ainda de variedades rurais e urbanas de português faladas no Brasil (Lopes 2001; Baxter & Lopes 2004a, 2004b, 2006, 2009).

O crioulo de Cabo Verde (CCV) e o santome possuem línguas nígero-congolesas como substrato, estabelecendo paralelismos formais entre si. Estes evidenciam-se quer no uso de nomes nus (exemplos (5) e (6)) quer na marcação plural (exemplos (7) e (8)), que pode ser feita com inserção de marca apenas no item pré-nominal, permanecendo inalterado o radical dos restantes itens do SN.

(5) CCV (Santiago):

Es tene odju regaladu na kabésa. (Pereira 2006: 47)
eles tem olho arregalado na cabeça.
'Eles estão espantados.'

(6) Santome:

ome plejidu (Hagemeyer 2000: 118)
'O homem preguiçoso.'

(7) CCV:

Góra, nu ta kume uns banana. (Pereira, et al 2006: 12)
agora 1PL ASP comer IND banana
'Agora, nós vamos comer umas bananas.'

(8) Santome:

inen ome plejidu (Hagemeyer 2000: 118)
3sg homem preguiçoso
'Os homens preguiçosos.'

Estes fenómenos podem ser conotados às L1's dos ancestrais aprendentes do português e, no caso concreto dos nomes nus do santome (o substrato do PA), eles têm distribuição e interpretação livre a nível gramatical, mas restringidas pela componente extra-gramatical. Licenciando-se na *interface* entre nível gramatical e extra-gramatical, os nomes nus relacionam o discurso à pragmática (Alexandre & Hagemeyer 2007: 49) e a interpretação de definitude acontece por ser previamente licenciada no discurso.

Relativamente ao sistema de marcação estrutural, nota-se tendência para a transposição do sistema de línguas do grupo nígero-congolês quer para o

português de Cabo Verde (PCV) (exemplo (9)) quer para o PA (exemplo (10)). Sendo a marca formal de plural inserida apenas no primeiro item da cadeia sintagmática, considera-se a probabilidade de as distintas variedades de português adquiridas em situação de contato por falantes de diferentes crioulos portugueses do Atlântico manifestarem idênticos padrões de variação entre si.

(9) PCV:

As pessoa não interessem para enfeitar a rua (Jon-And 2011: 16)
'As pessoas não se interessam em enfeitar a rua.'

(10) PA:

Vem do pé todas dia paqui [FRANCM2]
'Vem a pé todos os dias para aqui.'

Por outro lado, o paralelismo deste tipo de marcação com o de variedades do português afro-indígena do Brasil e do próprio PB argumenta a favor da influência das ancestrais línguas africanas no sistema de pluralização deste último (cf. Guy 1981: 301-302), decorrente de um *continuum* afro-brasileiro de português (Petter 2008: 193; Figueiredo & Oliveira 2013: 175). A ideia deste *continuum* assenta na observação de comportamentos morfossintáticos semelhantes em certas variedades de português, como a angolana, a moçambicana e a brasileira, tendo como ponto em comum o contato, ancestral ou atual, entre essas variedades de português e línguas do grupo banto.

No conjunto de variáveis independentes, não foram consideradas estatisticamente significativas: (i) *Presença de outros modificadores no SN*; (ii) *Determinante no SN anafórico prévio*; (iii) *Função sintática do SN*; (iv) *Presença de preposição dos adjuntos do sintagma verbal e do SN*; (v) *Presença de advérbio locativo no SN*. Como tal, iremos apenas discutir as questões em torno das variáveis: (i) *Número do SN*; (ii) *Tipo de artigo*.

Para a constituição e estudo da variável *Número do SN*, foram criados os fatores: (i) SN singular; (ii) SN plural; (iii) SN singular com semântica de plural.

No PA, são frequentes nomes singulares com referência semântica de plural, originando dois padrões que se demarcam do sistema mais comum de marcação plural no SN desta língua (exemplo (10)), ou seja, a pluralização com recurso à marcação formal única no elemento que antecede o nome núcleo do SN: (i) a estrutura com nome nu como núcleo singular mas com referência plural (exemplo (11a)); (ii) a estrutura DET+N, em que nenhum dos constituintes carrega a marca de plural (exemplo (11b)):

(11) PA:

- a. **polícia** cercaram ele e pegaram-lhe [CARMOM1]
‘os polícias cercaram-no e pegaram-no/agarraram-no.’
- b. **a roça** praticamente stá tudo bandonado [ABILH2]
‘as roças estão praticamente todas abandonadas.’

Assim, constituiu-se o fator *SN no singular com semântica plural* para se entender o seu efeito sobre a variável dependente, ou seja, perceber qual a tendência dos falantes inserirem artigo neste tipo de nomes, já que, ao que tudo indica, o SN com marca formal de plural revela propensão para favorecer a inserção. Importa ainda referir que os SN's cuja concordância plural é parcial, por norma atribuída à categoria DET (o elemento pré-nuclear adjacente), foram tratados como SN's plural. Para a variável *Número do SN* levantou-se a hipótese: *o SN plural favorece a inserção do artigo definido, enquanto o SN singular a desfavorece.*

A variável *Tipo de artigo definido* foi motivada por estudos sobre o modo como ocorre a aquisição da categoria SDET. Para tanto, estabelecemos os seguintes 16 fatores: (i) Artigo definido masculino singular; (ii) Artigo definido masculino plural; (iii) Artigo definido masculino singular desviado da norma; (iv) Artigo definido masculino plural desviado da norma; (v) Ausência de artigo definido masculino singular; (vi) Ausência de artigo definido masculino plural; (vii) Artigo definido feminino singular; (viii) Artigo definido feminino plural; (ix) Artigo definido feminino singular desviado da norma; (x) Artigo definido feminino plural desviado da norma; (xi) Ausência de artigo definido feminino singular; (xii) Ausência de artigo definido feminino plural; (xiii) Pronome pessoal singular, substituto do artigo definido masculino; (xiv) Pronome pessoal plural, substituto do artigo definido masculino; (xv) Pronome pessoal singular, substituto do artigo definido feminino; (xvi) Pronome pessoal plural, substituto do artigo definido feminino.

Godinho (2005: 132-138), observando a aquisição do português por estudantes chineses, aponta como ocorre a inserção da categoria funcional SDET na interlíngua dos alunos, pressupostos que Figueiredo (2010) aplicou também para observar como acontece essa aquisição no santome e no PA. Tais pressupostos permitem perceber que o N pode ocorrer sem DET (SN's nus), sem invalidar a hipótese de existir projeção do SDET (Baxter, Lucchesi & Guimarães 1997: 8, nota 14, *apud* Godinho 2005: 132). Meisel (1994: 92), por seu lado, constatou que, no estado inicial da aquisição bilingue, não ocorrem categorias funcionais. Estas emergem apenas no tempo, quando se desenvolvem as

categorias associadas à flexão (concordância, tempo e finitude). Portanto, apesar de ser foneticamente vazio, o SDET encerra em si as propriedades referenciais de especificação do SN (Godinho 2005: 132).

As aquisições de L1 e L2 exibem aspetos comuns, apesar de orientadas por operações particulares. Na aquisição do SDET da L1, as categorias funcionais apenas detêm função gramatical, subordinando a si as categorias lexicais (Godinho 2005: 135). No entanto, falantes de L1's do grupo nígero-congolês em aquisição do português L2 terão de adquirir a morfossintaxe da língua alvo (LA), ou seja, aprender que os nomes desta não são marcados com afixos temáticos como nas suas L1's, o que resulta em variação na interlíngua (o santome) determinada pela aquisição parcial das categorias funcionais da LA. Considerando que esta aquisição pressupõe ainda a aquisição de traços e formas morfofonológicas, infere-se que os falantes do PA também não adquirem plenamente os traços que não foram inseridos no santome.

Por outro lado, ocorrem duas particularidades idênticas em aquisição de L1 e de L2: “a concordância em género é primeiro estabelecida com o determinante e depois com o adjetivo. O número de erros também é menor com determinantes do que com outros modificadores” (Godinho 2005: 148). Este segundo aspeto foi também observado em estudos sobre aquisição de género no PHEL (Lucchesi 2000) e no PA (Figueiredo 2009), fazendo pressupor que o género masculino encerra o valor de “*default*”. Portanto, o falante aplica este quando não tem a certeza sobre o género do nome.

Estudos sobre aquisição de género em L2 mostram que a aquisição de concordância do masculino é a mais problemática (White *et al* 2004). Assim, apesar de os falantes almorarifanos dominarem a categoria funcional SDET e evidenciarem aquisição de N e DET+N numa mesma fase (cf. Montrul 2004: 42-43; cf. Godinho 2005: 161-162), poderão não realizar a categoria funcional SDET. Tal resulta de dificuldades na perceção fonológica, que determinam uma construção mental de concordância temática e não uma construção do tipo estrutural, já que esta efetua concordância a nível sintático.

Ao constituir-se a variável *Tipo de artigo definido* pretendemos trazer mais luz sobre os processos de formação das estruturas sintagmáticas do PA, já que esta variável, observada em concomitância com as variáveis *Idade* e *Escolaridade*, deixa entender como ocorre a inserção geracional do determinante definido. Para esta variável formulou-se a hipótese: *considerando que a aquisição das categorias funcionais pressupõe a aquisição quer de traços quer de formas morfofonológicas, os falantes de Almorarife inserem mais o artigo definido masculino do que o feminino, em virtude de o primeiro encerrar o valor de “default”. Paralelamente, aplicam a marcação de número plural*

nestes artigos com a fricativa /s/ numa fase mais tardia à da aplicação do número singular.

2.2.2.3. Variáveis independentes semânticas

O pacote estatístico considerou significativas as variáveis: (i) *Animacidade do substantivo*; (ii) *Referência do SN*.

Para constituição da variável *Animacidade do substantivo*, considerou-se que vários trabalhos atribuem ao traço [+humano], por ser discursivamente mais saliente (Mollica 1977), um papel relevante em variados fenómenos (Lyons 1979). Tal acontece também em variedades de português, favorecendo a marcação plural no SN (Scherre 1988), ou a conservação da flexão de género em nomes adjetivados do tipo *macho/fêmea* ou *homem/mulher* (Ferraz 1979; Lucchesi 2000; Figueiredo 2009). Portanto, as propriedades semânticas interagem com a função sintática, levando os nomes em posição de sujeito a serem marcados, visto nessa posição surgirem maioritariamente agentes do tipo animado (Young 1991).

Atentando-se à variação dos exemplos (3a) e (3b), abaixo renumerados como (12a) e (12b) para mais fácil leitura, constata-se que ela ocorre com nomes do tipo [+humano] ou extensão de [+humano]:

(12) PA:

- a. Dentro de dez minuto ele seô já stá morto. [ABILH2]
'Dentro de dez minutos o senhor já está morto.' [PE]
- b. Eu, quê zê, nós temos eles nosso documentos. [ABILH2]

Ao que sabemos, esta é uma especificidade do PA que respaldará no substrato santome (Figueiredo 2010), conforme exemplo (9), abaixo renumerado como (13):

(13) Santome:

- | | | | |
|--------------------------|------------|----------------|------------------------|
| <i>inen</i> | <i>ome</i> | <i>plejidu</i> | (Hagemeijer 2000: 118) |
| 3sg | homem | preguiçoso | |
| 'Os homens preguiçosos.' | | | |

A marcação plural do santome recorre ao determinante definido *inen* do edo, homófono do pronome pessoal 3ª pessoa plural, colocado pré-nominalmente como marcador de flexão plural em nomes [+humanos], o mesmo acontecendo no PA (exemplo (12b)). Dada esta similaridade entre o PA e o santome, e tendo

em conta que no passado das comunidades de Almojarife e de Helvécia terá havido “uma fase em que variedades de português L2 faladas por africanos serviram de dados linguísticos primários para aquisição do português L1, processo que teria facilitado a incorporação de características crioulas” (Baxter & Lopes 2009: 321), seja na realização de nomes nus seja na inserção da pluralização com recurso aos itens *inen/eles*, importa verificar se as motivações que levam a tal no PA seguem o mesmo padrão de interpretação que acontece no santome.

Alexandre & Hagemeyer (2007: 52) estabelecem uma ponte formal entre as noções semânticas de [+definitude] e [+plural] no santome, advogando que ambas estão inseridas na categoria DET *inen*. Assim, o nome nu núcleo do SN, sobretudo quando é representado por substantivos de traço [+humano], tem referência singular e é entendido como definido singular.

Quanto à marcação plural no CCV, Baptista (2002: 41, 2007: 84-86) aponta idêntico vínculo entre definitude e animacidade, observando que existe articulação entre SN definido singular e nomes nus. Para esta autora, os nomes nus no CCV são também motivados pela animacidade do substantivo, uma particularidade igualmente defendida por Baxter & Lopes (2004a) em estudo sobre esta variável na NURC. Efetivamente, nesta fala os nomes nus são também motivados pelo traço [+humano], enquanto o traço [-humano] potencia a inserção do artigo definido.

Contudo, Figueiredo (2010) ao observar a influência da variável *Animacidade* no desenho da concordância plural no SN do PA, indica que esta é mais condicionada por questões estruturais do que semânticas, uma possibilidade que importa confirmar quanto ao uso do artigo definido. Para se analisar estas questões, codificaram-se os substantivos de acordo com o seu traço semântico e colocou-se a seguinte hipótese para a variável *Animacidade do substantivo*: *os nomes de traço [-humano] favorecem a inserção do artigo definido, enquanto os substantivos de traço [+humano] tendem a desfavorecê-la*.

Quanto à variável *Referência do SN*, foi constituída considerando-se que, no CCV e no santome, o uso de nomes nus é facilitado pela pragmática (Baptista 2007: 74; Alexandre & Hagemeyer 2007: 53-54). Assim, criaram-se os fatores: (i) Referência [+única]; (ii) Referência [-única]; (iii) Referência ao SN do documentador.

Baxter & Lopes (2004a, 2005, 2009) testaram o peso desta variável, única ou específica, na inserção do artigo definido do PHEL. Os resultados confirmaram que “o conhecimento compartilhado pelos interlocutores facilita bastante o uso do artigo definido \emptyset nos SNs que têm referentes claramente

identificáveis" (Baxter & Lopes 2009: 323),¹¹ enquanto a referência [+única] favorece a inserção. Este aspeto levou-nos também a constituir um fator nunca antes observado, *Referência ao SN do documentador*, pois considera-se que algumas realizações poderem espelhar automonitorização e autocorreção, ou seja, os falantes reproduzem, pontualmente, realizações do discurso do documentador, logo mais próximas da norma:

(14) PA:

DOC: E lá você ficô 1 ano. **A vida lá é boa?**

INF: **A vida lá é boa**, sim. [CARMOM1]

Para a variável *Referência do SN*, elaborámos a hipótese: *o fator referência [-única] facilita a ocorrência de nomes nus, enquanto o fator referência [+única] e a automonitorização potenciam a inserção do artigo.*

3. Análise dos resultados

Depreenderam-se 13220 SN's de referência definida, com baixa inserção do artigo (16,2%) e altíssima taxa de não inserção (83,8%). A rodada geral para seleção das variáveis estatisticamente significativas forneceu os seguintes valores: *Input*: 0.026; *Log-likelihood*: -1057.493; *Significância*: 0.082. As variáveis estatisticamente significativas são:

- (i) Sociais: *Sexo; Idade; Escolaridade.*
- (ii) Estruturais: *Número do SN; Tipo de artigo definido.*
- (iii) Semânticas: *Animacidade do substantivo; Referência do SN.*

3.1. Variáveis independentes sociais

Observaremos agora os resultados quantitativos das variáveis sociais estatisticamente significativas, estabelecendo-se comparações com os valores registados no PT (Baxter & Lopes 2005, 2006), no PHEL (Baxter 2009) e na NURC (Baxter & Lopes 2004a, 2005).

Visto que variável *Sexo* não foi observada nestes estudos, apenas apresentaremos e analisaremos os resultados do PA (Tabela 1).

¹¹ Exemplo de Baxter & Lopes (2009: 323): “*tava oiano na lua assim ói, lua tá traz do sol uma bandinha*”.

Tabela 1: Sexo: Inserção do artigo definido no PA

Sexo	Dados	Pr.	Inserção	Grupo
Masculino	1580/7669	0.573	20.6%	58.0%
Feminino	566/5551	0.400	10.2%	42.0%
Totais	2146/13220		16.2%	100.0%

Os homens produziram mais ocorrências (7669 = 58%) do que as mulheres (5551 = 42%). Os valores mostraram ainda que os primeiros usam a forma mais próxima da norma (pr. 0.573), cabendo a maior variação às mulheres (0.400).

Observe-se também o perfil geracional de inserção por sexo (Tabela 1.1).

Tabela 1.1: Idade e Sexo: Inserção do artigo definido no PA

Faixa etária	Sexo	Dados	Pr.	Inserção	Grupo
FE-1	M	497/2323	0.684	21.4%	60.5%
	F	249/1519	0.235	16.4%	39.5%
FE-2	M	683/3030		22.5%	57.3%
	F	154/2258		6.8%	42.7%
FE-3	M	400/2316	0.656	17.3%	56.6%
	F	163/1774	0.301	9.2%	43.4%
Totais:		2146/13220		16.2%	100.0%

A inserção do artigo é liderada pelo sexo masculino. Efetivamente, ainda que a variável *Sexo* não seja estatisticamente significativa na FE-2, os percentuais de inserção (M = 22,5%; F = 6,8%) acompanham os das outras faixas etárias, revelando que o sexo feminino apresenta acentuável variação, enquanto os homens favorecem ligeiramente a inserção (FE-3: M = pr. 0.656; F = pr. 0.301; FE-1: M = pr. 0.684; F = pr. 0.235). Portanto, confirma-se a hipótese de que os homens realizam menos variação, dadas as condicionantes sociolinguísticas específicas da comunidade já apontadas (ponto 2.2.2.1).

Contudo, em sequência da independência de São Tomé não só os colonos e exército colonial se retiraram como também as estruturas socioeconômicas, o sistema de ensino e a rede rodoviária de São Tomé se desarticularam, levando a comunidade de Almojarife a isolar-se ainda mais (Figueiredo 2010: 88-89). Devido à falta de emprego generalizada, os membros do estrato masculino recorreram a alternativas de sobrevivência, dedicando-se a atividades novas, como a pesca e a lavoura de subsistência, e fixando-se mais na comunidade. Portanto, será necessário observar dados das gerações mais

novas para se confirmar se o atual padrão de variação dos dois sexos está agora mais nivelado.

Vejamos agora os resultados da variável *Idade*, para a qual não há resultados da NURC (Tabela 2):

Tabela 2: *Idade*: Inserção do artigo definido [3 variedades de português]

Faixa etária	PA			PT		
	Dados	Pr.	%	Dados	Pr.	%
FE-1	746/3842	0.656	19.4%	402/419	0.93	96.0%
FE-2	837/5288	0.445	15.8%	251/409	0.35	61.0%
FE-3	563/4090	0.406	13.8%	140/403	0.11	35.0%
Totais	2146/13220		16.2%	793/1231		64.4%

Faixa etária	PHEL		
	Dados	Pr.	%
FE-2 (35/45 anos)	1090/1147	0.750	95.0%
FE-3 (60/70 anos)	514/657	0.479	78.2%
FE-4 (+80 anos)	423/666	0.212	63.5%
FE-5 (+100 anos)	102/185	0.143	55.1%
Totais	2129/2655		80.2%

No PA verifica-se diferença de pesos relativos entre as três gerações. A FE-1, provavelmente devido à sua maior escolaridade, favorece moderadamente o uso do artigo definido (pr. 0.656). Quanto às faixas FE-2 (pr. 0.445) e FE-3 (pr. 0.406), revelam nivelamento entre si, inibindo ligeiramente a inserção. Porém, os valores permitem inferir que as gramáticas das três faixas etárias não registam distinções acentuadas, apresentando um padrão de inserção geracionalmente homogêneo. Assim, infere-se a possibilidade de mudança consumada,¹² a qual configura uma característica distintiva relativamente à variedade padrão.

Comparando os resultados do PA, do PT e do PHEL, constata-se que os informantes tongas estão distribuídos por faixas etárias idênticas às do PA, mas que os estratos etários de Helvécia variam ligeiramente. Todavia, estes aspetos não impedem uma análise comparada, pois permitem observar o perfil de evolução geracional nas três comunidades.

¹² O isolamento acentuado de Almoxarife no período pós-independência reforça a possibilidade de mudança consumada. Esta questão é debatida na análise dos resultados da variável *Escolaridade*.

O PT e o PHEL apresentam aquisição diacrónica semelhante, denotando um crescente uso do artigo definido, ou seja, aquisição em direção à nivelção com a variante padrão. Este perfil não ocorre em Almojarife, dado o seu isolamento temporal mais prolongado e que terá conduzido a uma estabilização da variação. Esta reflete-se na inserção inócua do artigo definido e no uso acentuado de SN's nus, uma particularidade ancorando no substrato santome. Portanto, a gramática da comunidade tem-se mantido geracionalmente apegada ao sistema da sua fase inicial de vida, na qual prevaleciam prolíficos dados da interlíngua.

Quanto à hipótese colocada, confirma-se com reservas. Apesar de a idade e a aquisição da regra de inserção estabelecerem uma relação proporcional, com a FE-1 favorecendo mais a inserção, acontece que os níveis de inserção estão estabilizados entre a FE-3 e a FE-2 e pouco crescem entre esta e a geração mais nova. Portanto, infere-se que o padrão de variação é geracionalmente conservador, reforçando a probabilidade de mudança consumada ou, pelo menos, de variação estabilizada.

Como a variável *Idade* revela perfil geracionalmente homogêneo, favorecendo o recurso aos nomes nus, avaliou-se também o peso da variável *Escolaridade* (Tabela 3) na regra de inserção. Esta variável foi apenas observada por nós. Verifica-se uma relação direta entre idade e escolaridade, pois a FE-3 apresenta elevado nível de analfabetismo, com apenas dois informantes tendo frequentado o ensino primário, mas sem o concluírem. Quanto à FE-1, é toda escolarizada, tendo 50% dos informantes frequentado o ensino pós-primário.

Tabela 3: Escolaridade: Inserção do artigo definido no PA

Escolaridade	Dados	Pr.	Inserção	Grupo
Analfabetos	538/4831	0.573	11.1%	36.5%
Baixa	510/3126	0.490	16.3%	23.6%
Média	603/3656	0.486	16.5%	27.7%
Alta	495/1607	0.780	30.8%	12.2%
Totais	2146/13220		16.2%	100.0%

Ainda que os indivíduos menos escolarizados e mais idosos de Almojarife apliquem inocuamente a regra de inserção (pr. 0.573), por oposição aos grupos com frequência da escolarização primária, parcial (pr. 0.490) ou total (pr. 0.486), os perfis de realizações destes três grupos exibem variação diacrónica bastante nivelada, sempre com valores próximos do ponto neutro. Assim, o ápice da

inserção ocorre apenas na geração mais escolarizada e menos idosa, ou seja, aquela que frequentou o ensino pós-primário e revela um ligeiro incremento da inserção da regra em direção à variante padrão (pr. 0.780).

Figueiredo (2012: 68-69) observou também a influência da escolaridade na aquisição da regra de pluralização no PA, apontando que condicionalismos específicos dos períodos colonial e pós-colonial marcaram o sistema educativo de São Tomé, em geral, e da comunidade de Almojarife, em particular, determinando o modo como esta variável actua sobre a pluralização. Consideraremos também esta particularidade na análise da regra de inserção do artigo definido, já que, segundo o autor, o fato de o corpo docente contratado após a independência do país ser composto por falantes nativos, que detêm um padrão mais próximo do dos próprios alunos do que da variante padrão, contribuiu para estabilizar a variação.

Assim, o desenvolvimento da regra de concordância plural não apresenta uma linha curvilínea ou de inclinação progressiva, confirmando-se que o PA, se não se encontra em estágio de mudança consumada, está já, pelo menos, em situação de variação estabilizada.

Os resultados do nosso estudo permitem confirmar parcialmente o pressuposto levantado para a variável *Escolaridade*, pois apenas a frequência do ensino pós-primário faz os falantes inserirem mais o artigo definido. Contudo, considerando que a variação diacrónica entre as três faixas etárias é nivelada, bem como os pressupostos de Figueiredo (2012) acerca da nivelção entre o padrão dos atuais docentes e alunos, é possível inferir que os traços não-interpretáveis na interlíngua transmitida irregularmente como L1 estarão a contribuir para a fixação da actual parametrização entre os falantes do PA (cf. Gonçalves 2004: 235). Esta parametrização poderá conduzir a uma fossilização suave (cf. Figueiredo 2012: 69, *apud* Long 2003: 490), a qual, por seu lado, determinará que o actual perfil de inserção do determinante venha a configurar mudança, caso os falantes não recebam *input* corretivo segundo a norma padrão.

3.2. Variáveis independentes semânticas

Os resultados das variáveis extralinguísticas não confirmam totalmente as hipóteses levantadas, pois argumentam a favor da variação estável ou mudança consumada, que favorece a realização de nomes nus no PA. Para se confirmar esta possibilidade, analisemos as variáveis semânticas estatisticamente significativas.

Os resultados da variável *Animacidade do substantivo* (Tabela 4) só serão comparados com os da NURC, pois não há valores para o PT nem o PHEL.

Tabela 4: Animacidade: Inserção do artigo definido [2 variedades de português]

Traço	PA			NURC		
	Dados	Pr.	%	Dados	Pr.	%
[+humano]	461/4479	0.417	10.3%		0.37	
[-humano]	1685/8741	0.543	19.3%		0.62	
Totais	2146/13220		16.2%			

Os dois padrões evidenciam características comuns, pois são os nomes de traço [-humano] que propiciam a inserção do artigo definido (PA = pr. 0.543; NURC = pr. 0.62), enquanto os substantivos de traço [+humano] o inibem (PA = pr. 0.417; NURC = pr. 0.37). Contudo, a diferença entre ambos não é muito significativa, fazendo pressupor que a inserção do artigo será mais determinada por condicionalismos sócio-estruturais do que pelos de carácter semântico, uma particularidade também observada por Figueiredo (2010: 695) em estudo sobre a inserção plural no SN do PA.

Por outro lado, os valores permitem ver que o PA e a NURC seguem os padrões do santome e do CCV quanto à interpretação semântica do nome nu ser motivada pelo traço [+humano] (Alexandre & Hagemeijer 2007: 49; Baptista 2007: 86). Confirma-se, então, a hipótese de que os nomes com traço [-humano] favorecem mais a inserção do artigo definido, enquanto os substantivos com traço [+humano] a desfavorecem.

A variável *Referência do SN* (Tabela 5) não foi considerada estatisticamente significativa no PT nem na NURC, pelo que apenas compararemos os resultados do PA e do PHEL. Lembre-se ainda que o fator *Referência ao SN do documentador* foi somente testado no PA.

Tabela 5: Referência do SN: Inserção do artigo definido [2 variedades de português]

Referência	PA			PHEL		
	Dados	Pr.	%	Dados	Pr.	%
[+única]	2045/12386	0.509	16.5%	2040/2524	0.516	80.8%
[-única]	73/648	0.279	11.3%	89/130	0.218	68.5%
Documentador	28/186	0.703	15.1%			
Totais	2146/13220		16.2%			

O PA e o PHEL evidenciam perfis condizentes, com o fator referência [+única] exercendo peso neutro na inserção (PA = pr. 0.509; PHEL = pr. 0.516) e a referência [-única] contribuindo para se realizarem nomes nus no SN subsequente (PA = pr. 0.279; PHEL = pr. 0.218).

Paralelamente, os informantes revelam também tendência para a automonitorização, inserindo frequentemente o artigo quando repetem estruturas do documentador com a presença do mesmo (PA = pr. 0.703).

Os resultados analisados vão ao encontro dos pressupostos de Alexandre & Hagemeyer (2007: 53-54) para o santome e de Baptista (2007: 74) para o CCV, que advogam que o uso de nomes nus é potenciado pelo fator pragmático. Confirma-se, assim, a hipótese levantada para a variável *Referência do SN*, pois o conhecimento compartilhado pelos interlocutores facilita a ocorrência de nomes nus no SN subsequente, enquanto o fator pragmático de referência [+única] e a automonitorização propiciam mais a inserção do artigo. Contudo, os valores de favorecimento divergem entre si, já que os do fator referência [+única] são praticamente neutros (pr. 0.509), enquanto os da automonitorização favorecem acentuadamente a inserção (pr. 0.703).

Os resultados das variáveis semânticas, se se excetuar o peso exercido pelo fator *Referência ao SN do documentador*, indicam que todos os fatores favorecedores da inserção do artigo definido possuem valores neutros, confirmando-se que este tipo de variáveis não exerce peso significativo na regra da inserção. Assim, esta deverá ser mais condicionada por fatores sócio-estruturais, tal como sucede no desenho da concordância de número variável no SN do PA (cf. Figueiredo 2010: 695), um aspeto que procuraremos confirmar analisando o efeito de variáveis estruturais estatisticamente significativas sobre as variáveis dependentes.

3.3. Variáveis independentes estruturais

Tabela 6: Número do SN: Inserção do artigo definido [4 variedades de português]

SN	PA			PT		
	Dados	Pr.	%	Dados	Pr.	%
SN SG	1859/11735	0.520	15.8%	592/1004	0.40	59.0%
SN PL	242/558	0.628	43.4%	201/227	0.84	89.0%
SN SG com semântica PL	45/927	0.211	4.9%			
Totais	2146/13220		16.2%	793/1231		64.4%
SN	PHEL			NURC		
	Dados	Pr.	%	Dados	Pr.	%
SN SG	1863/2369	0.458	78.6%		0.45	
SN PL	266/286	0.800	93.0%		0.71	
SN SG com semântica PL						
Totais	2129/2655		80.2%			

A variável *Número do SN* (Tabela 6) permite observar o efeito do número sobre a variável dependente, bem como a relação entre o SN definido singular e o artigo definido de referência zero. Refira-se ainda que o fator *SN singular com semântica plural* foi analisado apenas neste estudo.

O *SN plural* favorece a inserção do artigo (pr. 0.628), enquanto o *SN singular* tem peso neutro na mesma (pr. 0.520). Este padrão é semelhante em todas as variedades (favorecimento do *SN plural*: PT = pr. 0.84; PHEL = pr. 0.800; NURC = pr. 0.71 / peso neutro do *SN singular*: PT = pr. 0.40; PHEL = pr. 0.458; NURC = pr. 0.45) e respalda nos pressupostos sobre a ponte formal entre as noções de [+definitude] e [+plural], levando ambos os traços, no caso do santome, a serem inseridos no DET *inen* (cf. Alexandre & Hagemeyer 2007: 52). Assim, é atribuída referência singular ao nome nu, sobretudo em substantivos de traço [+humano]. Este respaldo nos traços crioulos vai também ao encontro dos pressupostos de Baptista (2002: 41; 2007: 84-86), pois no CVV ocorre igualmente um vínculo entre o SN definido singular e a aplicação de nomes nus.

O uso de nomes singulares com referência semântica de plural no PA (exemplos (11a) e (11b)) levou-nos a constituir o fator *SN singular com semântica plural*, observando-se o seu peso sobre a variável dependente. Apesar de o fator *SN plural* favorecer o uso do artigo, a semântica plural do nome, por si só, não é suficiente para determinar a inserção, já que a tendência é para a inibição acentuada: pr. 0.211. Verifica-se, novamente, que as particularidades semânticas não exercem peso na regra da inserção do artigo definido.

O facto de os falantes das L1's do grupo nígero-congolês em aquisição do português L2 terem de fixar novas propriedades sintáticas na interlíngua pode estar na origem desta situação. Estudos sobre aquisição da concordância do género (Name & Corrêa 2003; Ferrari-Neto, Corrêa & Augusto 2005) e do número (Marrero & Aguire 2003) evidenciam que o seu processamento formal ocorre no SDET. Os estudos evidenciam também que a informação de género está expressa no DET, mas que o traço interpretável de número se encontra na categoria N (Chomsky 1996) ou numa projeção funcional independente (Ritter 1991: 61). Assim, no caso do género, o falante identifica-o em palavras novas sem necessidade de motivação semântica (Ferrari-Neto, Corrêa & Augusto 2005: 1077).

Quanto ao número, a interpretação expressa-se numa desinência flexional do nome, ganhando visibilidade na *interface* fonética. Já os restantes elementos do SDET (determinantes, possessivos, adjetivos...) apresentam traços não-interpretáveis de número, estabelecendo concordância com o nome ou com a projeção funcional NUM. Portanto, na aquisição de número, os falantes não

fazem inicialmente a distinção entre formas singular/plural, recorrendo apenas ao traço interpretável do N e realizando nomes nus (exemplo (11a)), para, posteriormente, aplicarem o artigo, embora no singular (exemplo (11b)), Posteriormente, inserirem a marca formal de plural, inicialmente na classe DET e, seguidamente, nos outros elementos do SDET, mas com variação na aplicação da regra de concordância (Figueiredo 2010: 207). A evolução da regra de inserção do artigo pode, então, ser explicitada de acordo com o esquema seguinte:

Figure 1: Inserção do artigo definido e desenvolvimento da pluralização na estrutura DET+N do PA

NSG ⇔ DETSG+NSG ⇔ DETPL+NSG ⇔ DETPL+NPL

Este esquema justifica também o fato de o determinante sob a forma do pronome pessoal do português, 3ª pessoa (exemplos (3a) e (3b)), poder ocorrer no singular ou no plural, enquanto no santome se realiza apenas com a forma plural *inen*. Paralelamente, o esquema não invalida a possibilidade de o nome nu, em determinadas situações, poder ser pluralizado, particularidade com motivações que importa verificar futuramente. Por agora, limitar-nos-emos a observar como ocorre a formação geracional das estruturas sintagmáticas do PA com inserção do artigo definido, cruzando o grupo de fatores *Número do SN* com as variáveis *Tipo de artigo* e *Faixa etária*. Isentamo-nos de apresentar aqui as tabelas destes resultados, dado serem bastante desenvolvidas e este trabalho ter limitações quanto à extensão.

Quanto à hipótese levantada para a variável *Número do SN*, confirma-se parcialmente, pois o *SN plural* apenas favorece a inserção do artigo definido quando possui marcação formal visível. Por seu lado, o *SN singular* exerce peso neutro na inserção.

A variável *Tipo de artigo definido* foi apenas observada por nós, pelo que apresentaremos resultados percentuais do seu cruzamento com as variáveis *Idade* e *Número do SN* somente para o PA. Ocorrem 11735 SN's singular no *corpus*, assim repartidos por género: 5906 SN's género masculino; 5829 género feminino. Nos primeiros, o artigo definido masculino singular registou 4947 ausências (83,8% SN's nus) e 941 inserções (15,9%) segundo o seguinte padrão:

- (i) 792 vezes consoante norma do português (13,4%);
- (ii) 148 variações de género (2,5%);

- (iii) 11 variações de número (0,2%);
- (iv) 7 variações de género e número (0,1%);
- (v) 1 vez com o pronome pessoal “ele” em vez do artigo definido.

Confirma-se a tendência de os SN's singulares favorecerem a realização de SN's nus, sendo o percentual de variação mais acentuado na identificação do género do que do número.

Relativamente ao artigo definido feminino singular, verificaram-se 4929 inibições (84,6% SN's nus) e 900 inserções (15,4%) conforme o seguinte padrão:

- (i) 807 consoante norma do português (13,8%);
- (ii) 76 variações de género (1,3%);
- (iii) 11 variações de número (0,2%);
- (iv) 5 variações de género e número (0,1%);
- (v) 1 vez com o pronome pessoal “ele” em vez do artigo definido.

Os padrões de favorecimento e inibição são idênticos quanto ao uso dos artigos definidos singulares masculino e feminino. Porém, regista-se uma ligeira acentuação percentual na inserção desviada do género, favorecendo levemente o uso do definido masculino, o que confirma a hipótese do seu valor de “*default*”.

Geracionalmente, os padrões da FE-3 e FE-2 são semelhantes entre si e relativamente ao da comunidade, registando aproximações percentuais de favorecimento e de inibição no uso do artigo. Por seu lado, a FE-1 realizou 3376 SN's singulares (1567 do género masculino e 1809 do género feminino), apresentando a seguinte variação na identificação do género e número:

- (i) 10 no género masculino (0,6%) e 13 no feminino (0,7%);
- (ii) 2 no número masculino (0,1%) e 5 no feminino (0,3%);
- (iii) 2 no género e número masculino (0,1%) e nenhuma no género e número feminino.

Ainda assim, o percentual de SN's nus é elevado, comprovando-se que a aquisição da regra de inserção do artigo definido não representa um parâmetro categórico da gramática da comunidade: 80,1% de inibição face a 19,1% de realizações com masculino singular; 82,6% de inibição em oposição a 16,4% de inserções com feminino singular. Assim, mantém-se a tendência geracional para maior inserção do artigo masculino relativamente à do artigo feminino.

Quanto aos SN's plural, totalizam 558 ocorrências, assim divididas por género: 291 SN's masculinos e 267 femininos. Os masculinos apresentam 143 inibições (49,2% SN's nus) e 148 inserções (50,8%) conforme seguinte:

- (i) 122 vezes consoante norma do português (41,9%);
- (ii) 2 variações de número (0,7%);
- (iii) 12 variações de número e género (2,4%);
- (iv) 10 variações de género (3,4%);
- (v) 2 vezes com o pronome pessoal "eles" em vez do artigo definido (0,7%).

Os SN's femininos registam 173 inibições (64,8% SN's nus) e 94 inserções (35,2%) com o padrão:

- (i) 75 consoante norma padrão do português (28,1%);
- (ii) 3 variações de número (1,1%);
- (iii) 13 variações de género e número (4,9%);
- (iv) 3 variações de género (1,1%).

O favorecimento e a inibição seguem o mesmo perfil verificado no uso dos artigos definidos plurais masculino e feminino, confirmando-se a tendência de os SN's plurais favorecerem mais a inserção do artigo do que os SN's singulares. Paralelamente, confirma-se que, também no uso da forma plural, o artigo masculino favorece mais a inserção, com percentuais de aplicação próximos do valor neutro (41,9%), enquanto o artigo feminino a desfavorece notoriamente (28,1%).

Geracionalmente, observa-se que a aquisição do SN singular é mais problemática do que a do SN plural, visto que os falantes começam por inserir o artigo definido singular em fases precoces, ancorando a marcação no sistema das L1's (cf. Baxter 2009: 293). Assim, a inserção da marca plural no artigo é feita logo na FE-3, que revela variação mínima em virtude de os falantes já terem adquirido as noções básicas do sistema de pluralização da LA. A variação regista-se acentuadamente apenas nos SN's nus, com maior incidência no género feminino:

- (i) SN's nus plural na FE-3: 38,3% masculinos; 71,1% femininos;
- (ii) SN's nus plural na FE-2: 50,4% masculinos; 65,8% femininos;
- (iii) SN's nus plural na FE-1: 53,6% masculinos; 62,4% femininos.

Os valores permitem concluir que os almoxarifanos só recorrem ao artigo plural quando não tem dúvidas acerca do seu uso, mas hesitando mais no uso do definido feminino.

Os resultados mostram ainda que os SN's singular com semântica de plural totalizam 927 ocorrências, distribuídas da seguinte forma quanto ao gênero: 426 SN's masculinos e 501 femininos. Os masculinos apresentam-se 413 vezes como SN's nus (96,9%) e 13 com inserção de artigo definido masculino singular:

- (i) 4 com padrão de gênero do SN (0,9%);
- (ii) 9 com variação do padrão de gênero do SN (2,2%).

Quanto aos SN's femininos, 469 são SN's nus (93,6%), enquanto 32 registam artigo definido feminino singular:

- (i) 3 com padrão de gênero do SN (0,6%);
- (ii) 29 com variação do padrão de gênero do SN (5,8%).

A variação é maior nos SN's singular com semântica de plural quando se trata da inserção do artigo definido feminino, conduzindo à desidentificação do padrão de gênero (2,2% inserções desviadas de artigo definido masculino; 5,8% inserções desviadas de artigo definido feminino). A variação inicia-se logo na FE-3 (2,8% inserções desviadas de artigo definido masculino; 5,7% de inserções desviadas de artigo definido feminino), mantendo-se na FE-2 (2,2% inserções desviadas de artigo definido masculino; 8,9% inserções desviadas de artigo definido feminino). Contudo, na FE-1, também mais escolarizada, a consciência sobre a correta aplicação do artigo é mais notória, não se registrando qualquer variação resultante da inapropriada inserção do artigo definido em SN's singular com semântica de plural. Portanto, prevalece apenas a variação originada pela realização de nomes nus utilizados inadequadamente na forma singular quando se pretende expressar a pluralização formal de número.

O efeito exercido sobre a variável dependente pela concomitância entre as variáveis *Número do SN*, *Tipo de artigo definido* e *Idade* permitiu confirmar a tendência para os SN's plurais exercerem maior influência na inserção do artigo definido, enquanto os SN's singulares potenciam a realização de SN's nus. Tal deve-se ao fato de a inserção do artigo singular ocorrer na fase precoce da emergência da interlíngua, revelando apego aos sistemas estruturais do santome L1 e que se transmitem geracionalmente. Como a inserção do artigo definido plural ocorre em fases mais tardias, os falantes já detêm um domínio

mais consistente do sistema gramatical da LA, daí a variação ser menos acentuada.

Paralelamente, confirma-se também que os almojarifanos inserem mais o artigo definido masculino do que o feminino, em virtude de o primeiro encerrar o valor de “*default*” que é aplicado na fase em que o uso das categorias funcionais ainda não revela aquisição plena quer de traços quer de formas morfofonológicas. O valor de “*default*” do género masculino é também uma característica do pronome pessoal realizado em vez do artigo definido, que não regista ocorrências com a forma feminina. Os resultados evidenciam ainda que este tipo de inserção apresenta resquícios de sobrevivência até à FE-2, extinguido-se na FE-1 (mais nova e mais escolarizada).

4. Conclusões

Procurámos determinar o quadro da variação do artigo definido no SN do PA, uma variedade que tem o santome como substrato. Comparando perfis de variação similar em variedades de português da África atlântica e do Brasil, buscámos entender as tendências da aquisição e uso do artigo em contextos de *transmissão linguística irregular*, bem como contribuir com mais dados para o conjunto de variáveis que testemunham a reestruturação morfossintática quer de variedades africanas de português quer de variedades do português brasileiro. Paralelamente, objetivámos confirmar se o português adquirido em situação de contato por falantes de um crioulo português do Atlântico manifesta padrões de variação semelhantes aos do português resultante de *transmissão linguística irregular*, na qual a língua ancestral não é um crioulo português do Atlântico.

Assim, analisámos dados de fala de 18 informantes, partindo-se do pressuposto que a variação no PA é motivada por fatores linguísticos e extralinguísticos, que contribuem para o desenho estrutural do SN. Recorremos ao pacote estatístico GOLDVARB X, seguimos o modelo teórico proposto em trabalhos que observam a variação no SN, seja de número seja de inserção do artigo definido, e avaliamos o peso probabilístico de 3 variáveis independentes extralinguísticas e 10 variáveis independentes linguísticas na inserção do artigo definido no PA.

A análise das variáveis estatisticamente relevantes para aplicação da regra de inserção revelou que:

1. A variação no uso do artigo definido do PA é mais influenciada por condicionalismos sócio-estruturais do que semânticos. Ainda assim, o PA segue

os padrões do santome e do CCV quanto à interpretação semântica do nome nu ser motivada pelo traço [+humano], levando os nomes com traço semântico [-humano] a favorecerem a inserção do artigo definido.

2. O sexo masculino realiza menos variação do que o feminino, devido aos condicionalismos sociolinguísticos específicos da comunidade.

3. Os SN's singulares favorecerem a realização de SN's nus, sendo problemática a aplicação do género feminino, pois o género masculino encerra o valor de "default" que é aplicado na fase em que o uso das categorias funcionais ainda não revela aquisição plena quer de traços quer de formas morfofonológicas. O valor de "default" do género masculino era também uma característica do pronome pessoal realizado em vez do artigo definido, nas fases em que ainda sobrevivia.

4. A variável dependente tem-se mantido estabilizada geracionalmente, facilitando a ocorrência de SN's nus, especialmente em SN's singular. Assim, o perfil do SN estará em processo de variação estabilizada que ancora na tipologia do SN sem artigo definido do santome. A possibilidade de uma mudança consumada é de considerar, pois apenas a escolarização pós-primária potencia a inserção do artigo, uma realidade que poderá não se confirmar nas atuais gerações menos idosas de Almocharife, dado os docentes do atual quadro de ensino de São Tomé deterem um padrão de registo mais próximo do dos seus alunos do que do padrão normativo. Contudo, só a análise de dados recolhidos junto das atuais crianças e jovens adolescentes permitirá confirmar se já ocorreu mudança.

5. O padrão de variação do PA é similar ao do PT, do PHEL e da NURC, apesar de o primeiro, pela condição de semi-isolamento mais prolongada, não apresentar um desenvolvimento da regra de inserção ao nível da dos outros três. Portanto, infere-se que o português adquirido por contato por falantes de um crioulo português do Atlântico manifesta idênticos padrões de variação do português adquirido por falantes cuja língua ancestral não é um crioulo português do Atlântico (cf. Figueiredo 2010), conquanto compartilhem semelhanças nos seus antecedentes sócio e etnolinguísticos. Nesse sentido, e dadas as similaridades entre as quatro variedades observadas, é também possível argumentar a favor de um *continuum* afro-brasileiro de português, tendo como ponto em comum o contato, ancestral ou atual, com línguas do grupo banto.

Abreviaturas

CCV=Crioulo de Cabo Verde; **DET**=Determinante; **FE-1**=Faixa Etária 20/40 Anos; **FE-2**=Faixa Etária 41/60 anos; **N**=Nome; **NURC**=Norma Culta de Salvador, Bahia, Brasil; **PA**=Português da Comunidade de Almojarife, São Tomé; **PB**=Português do Brasil; **PE**=Português Europeu (norma padrão); **PHL**=Português Rural de Helvécia, Bahia, Brasil; **PT**=Português dos Tongas, Roça Monte Café, São Tomé; **SDET**=Sintagma Determinante; **SN**=Sintagma Nominal.

Referências

- Alexandre, Nélia & Tjerk Hagemeyer. 2007. Bare-nouns and the nominal domain in Saotome. In Marlyse Baptista & Jacqueline Guéron (eds.), *Noun Phrases in Creole Languages: A Multi-faceted Approach*, 37-59. Amsterdam/Philadelphia: John Benjamins.
- Baptista, Marlyse. 2002. *The Syntax of Cape Verdean Creole: The Sotavento Varieties* (Linguistik Aktuell/Linguistics Today 54). Amsterdam/Philadelphia: John Benjamins.
- Baptista, Marlyse. Baptista, Marlyse. 2007. On the syntax and semantics of DP in cape verdean creole. In Marlyse Baptista & Jacqueline Guéron (eds.), *Noun Phrases in Creole Languages: A Multi-faceted Approach*, 61-105. Amsterdam/Philadelphia: John Benjamins.
- Baxter, Alan Norman. 2004. The development of variable NP plural agreement in a restructured African variety of Portuguese. In Geneviève Escure & Armin Schwegler (eds.), *Creoles, Contact, and Language Change: Linguistic and Social Implications*, 97-126. Amsterdam: John Benjamins.
- Baxter, Alan Norman 2009. A concordância de número. In Dante Lucchesi, Alan Norman Baxter & Ilza Ribeiro (eds.), *O Português Afro-brasileiro*, 269-293. Salvador: EDUFBA.
- Baxter, Alan Norman, Dante Lucchesi & Maximiliano Guimarães. 1997. Gender agreement as a “decreolizing” feature of afro-brazilian dialect of Helvecia. *Journal of Pidgin and Creole Languages* 12(1). 1-57.
- Baxter, Alan Norman & Norma Lopes. 2004a. Variação no uso de artigos: Uma comparação. Comunicação apresentada na *XX Jornada Nacional de Estudos Lingüísticos do GELNE*. Universidade Federal da Paraíba, João Pessoa.
- Baxter, Alan Norman & Norma Lopes. 2004b. Marcação referencial no dialeto afro-brasileiro de Helvécia, Bahia. Comunicação apresentada no *XIX ENANPOLL – Encontro Nacional de Pós-Graduação em Letras e Lingüística*, Maceió, 28 Junho a 2 Julho.
- Baxter, Alan Norman & Norma Lopes. 2005. O artigo definido em variação com zero no SN de referência específica: Re-estruturação em três variedades do português. Comunicação apresentada no *XIV Congreso Internacional ALFAL – Asociación de Lingüística e Filología de América Latina*, Monterrey, México, 17-21 Outubro.
- Baxter, Alan Norman & Norma Lopes 2006. Bare definite reference NPs in an afro-brazilian portuguese dialect. *Revista Internacional de Lingüística Iberoamericana* 4(1). 55-70.
- Baxter, Alan Norman & Norma Lopes. 2009. O artigo definido. In Dante Lucchesi, Alan Baxter & Ilza Ribeiro, *O Português Afro-brasileiro*, 319-330. Salvador: EDUFBA.
- Castro, Ana & Fernanda Pratas. 2006. Capeverdean DP-internal number agreement: Additional

- arguments for a distributed morphology approach. In João Costa & Maria Cristina Figueiredo Silva (eds.), *Studies on Agreement*, 11-24. Amsterdam/Philadelphia: John Benjamins.
- Chomsky, Noam. 1996. *The Minimalist Program*, 2nd edn. Cambridge, MA: The MIT Press.
- Ferrari-Neto, José, Leticia Maria Sicuro Corrêa & Marina R. A. Augusto. 2005. O processamento da informação de interface na aquisição do sistema de número gramatical no DP em português brasileiro. In Lúcia Maria Pinheiro Lobato *et al* (eds.), *Anais do IV Congresso Internacional da ABRALIN*, 1075-1082. Associação Brasileira de Linguística: Universidade de Brasília, <http://www.abralin.org/publicacao/abralin2005.pdf>. (31 de Maio, 2018.)
- Ferraz, Luiz Ivens. 1976. The origin and development of four creoles in the Gulf of Guinea. Reimpressão do *African Studies Quarterly Journal* 35. 33-38. Johannesburg: Witwatersrand University Press / Pietermaritzburg, Natal: The Natal Witness (Pty).
- Ferraz, Luiz Ivens. 1979. *The creole of São Tomé*. Reimpressão do *African Studies Quarterly Journal* 37(1/2). Johannesburg: Witwatersrand University Press / Pietermaritzburg, Natal: The Natal Witness (Pty).
- Ferreira, Carlota. 1984. Remanescentes de um falar crioulo brasileiro. In Carlota Ferreira *et al.*, *Diversidade do Português do Brasil*, 21-32. Salvador: EDUFA.
- Figueiredo, Carlos Filipe Guimarães. 2008. A concordância variável no sintagma nominal plural do português reestruturado de Almojarife (São Tomé). *PAPIA* 18. 23-43.
- Figueiredo, Carlos Filipe Guimarães. 2009. A configuração do SN do português reestruturado da comunidade de Almojarife, São Tomé. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola* 1(1). 28-55.
- Figueiredo, Carlos Filipe Guimarães. 2010. *A concordância plural variável no sintagma nominal do português reestruturado da comunidade de Almojarife, São Tomé: Desenvolvimento das regras de concordância variáveis no processo de transmissão-aquisição geracional*, 2 vols. Macau: Universidade de Macau, dissertação de doutoramento.
- Figueiredo, Carlos Filipe Guimarães. 2012. Variável extralinguística escolaridade: Influência na marcação plural do sintagma nominal do português reestruturado de Almojarife, São Tomé. *PAPIA* 22(1). 41-76.
- Figueiredo, Carlos Filipe Guimarães. 2014. Português de Almojarife, São Tomé: Sentenças relativas com relativizadores "ku" e "com". *Filologia e Linguística Portuguesa*, 16(2). 277-309.
- Figueiredo, Carlos Filipe Guimarães & Márcia Santos Duarte de Oliveira. 2013. Português do Libolo, Angola, e português afro-indígena de Jurussaca, Brasil: Cotejando os sistemas de pronominalização. *PAPIA* 23(2). 105-185.
- Fiorin, José Luiz & Margarida Maria Taddoni Petter (orgs.). 2008. *África no Brasil: A Formação da Língua Portuguesa*. São Paulo: Ed. Contexto.
- Godinho, Ana Paula Cleto de Oliveira. 2005. *A aquisição da concordância de plural no sintagma nominal por aprendentes chineses de português língua estrangeira*. Lisboa: Universidade de Lisboa, dissertação de doutoramento.
- Gonçalves, Perpétua. 2004. Towards a unified vision of classes of language acquisition and change: Arguments from the genesis of Mozambican African languages. *Journal of Pidgin and Creole Languages* 19(2). 225-259.

- Guy, Gregory Riordan. 1981. *Linguistic variation in Brazilian Portuguese: Aspects of the phonology, syntax, and language history*. Pennsylvania: Pennsylvania University, Dissertação de doutoramento.
- Hagemeijer, Tjerk. 1999. As ilhas de Babel: A criouliização no Golfo da Guiné. *Revista Camões* 6. 74-88.
- Hagemeijer, Tjerk. 2000. Verbos e gramaticalização em são-tomense. In Ernesto d'Andrade, Maria Antónia Coelho da Mota & Dulce Pereira (orgs.), *Crioulos de Base Portuguesa: Actas do Workshop sobre Crioulos de Base Lexical Portuguesa. FLUL, 29 e 30 de Abril de 1999*, 111-126. Braga: Associação Portuguesa de Linguística.
- Hagemeijer, Tjerk. 2007. *Clause structure in Santome*. Lisboa: Universidade de Lisboa, dissertação de doutoramento.
- Hagemeijer, Tjerk 2009. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola* 1(1). 2-27, <http://www.acblpe.com/revista/volume-1-2009/as-linguas-de-s-tome-e-principe>. (7 de Maio, 2018.)
- Instituto Nacional de Estatística de São Tomé e Príncipe. 2003/2005. *III Recenseamento Geral da População e da Habitação de 2001*. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe.
- Jon-And, Anna. 2011. *Variação, contato e mudança linguística em Moçambique e Cabo Verde: A concordância variável de número em sintagmas nominais do português*. Estocolmo: Stockholm University, dissertação de doutoramento.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Long, Michael H. 2003. Stabilization and fossilization in interlanguage development. In Catherine J. Doughty & Michael H. Long (eds.), *The Handbook of Second Language Acquisition*, 487-536. Oxford: Blackwell.
- Lopes, Norma da Silva. 2001. *Tópicos de concordância*. Salvador: Universidade Federal da Bahia, dissertação de doutoramento.
- Lorenzino, Gerardo. 1998. *The Angolar Creole Portuguese of São Tomé: Its grammar and sociolinguistic history*. Nova Iorque: City University of New York, dissertação de doutoramento.
- Lucchesi, Dante. 2000. *A variação na concordância de género em uma comunidade de fala afro-brasileira: Novos elementos sobre a formação do português popular do Brasil*. Rio de Janeiro: Universidade Federal do Rio de Janeiro, Faculdade de Letras, dissertação de doutoramento.
- Lucchesi, Dante. 2019. O conceito de transmissão linguística irregular. In Dante Lucchesi (coord.), *Projeto Vertentes do Português Popular do Estado da Bahia*, <http://www.vertentes.ufba.br/a-transmissao-linguistica-irregular>. (14 de Março, 2019.)
- Lyons, John. 1979. *Introdução à Linguística Teórica*. São Paulo: Nacional.
- Marrero, Victoria & Carmen Aguirre, 2003. Plural acquisition and development in Spanish. In Silvina Montrul & Francisco Ordóñez (eds.), *Linguistic Theory and Language Development in Hispanic Languages: Papers from the 5th Hispanic Linguistics Symposium and the 4th Conference on the Acquisition of Spanish and Portuguese*, 275-296. Somerville, MA: Cascadilla Press.
- Meisel, Jürgen M. 1994. Getting FAT: Finiteness, agreement and tense in early grammars. In Jürgen M. Meisel (ed.), *Bilingual First Language Acquisition: French and German Grammatical Development*, 89-129. Amsterdam: John Benjamins Publishing.
- Mollica, Maria Cecília. 1977. *Estudo da cópia nas construções relativas em português*. Rio de Janeiro: Pontifícia Universidade Católica, dissertação de mestrado.

- Montrul, Silvina A. 2004. *The Acquisition of Spanish: Morphosyntactic Development in Monolingual and Bilingual L1 Acquisition and Adult L2 Acquisition*. Amsterdam: John Benjamins.
- Name, Maria Cristina Lobo & Leticia Maria Sicuro Corrêa. 2003. Delimitação perceptual de uma classe correspondente à categoria funcional D: Evidências da aquisição do português. *Fórum Lingüístico* 3(1). 55-88.
- Pereira, Dulce. 2006. *O Essencial sobre Língua Portuguesa: Crioulos de Base Portuguesa*. Lisboa: Caminho.
- Pereira, Dulce, Eva Arim & Nuno Carvalho. 2006. Crioulo de Cabo Verde. In ILTEC, *Diversidade linguística na escola portuguesa: Projecto Diversidade Linguística na Escola Portuguesa (ILTEC)*, edição em CD-rom. Lisboa: ILTEC, Ministério da Educação & Fundação Calouste Gulbenkian, http://www.iltec.pt/divling/_pdfs/linguas_crioulo_cv.pdf. (6 de Junho, 2018.)
- Petter, Margarida Maria Taddoni. 2008. *Variedades lingüísticas em contato: Português angolano, português brasileiro e português moçambicano*. São Paulo: Universidade de São Paulo, dissertação de doutoramento.
- Ritter, Elizabeth. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. In Susan Deborah Rothstein (ed.), *Perspectives in Phrase Structure: Heads and Licensing* (Syntax and Semantics 25), 37-62. San Diego/London: Academic Press.
- Sankoff, David, Sali A. Tagliamonte & Eric Smith. 2005. *Goldvarb X: A Variable Rule Application for Macintosh and Windows*. Toronto: Dept. of Linguistics, University of Toronto, http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm. (9 de junho, 2018.)
- Schang, Emmanuelle. 2000. *L'émergence des créoles portugais du Golfe de Guinée*. Nancy: Université Nancy 2, dissertação de doutoramento.
- Scherre, Maria Marta Pereira. 1988. *Reanálise da concordância nominal em português*. Rio de Janeiro: Universidade Federal do Rio de Janeiro, dissertação de doutoramento.
- Silva, Giselle Machline de Oliveira e. 1996. Realização facultativa do artigo definido diante de possessivo e de patronímico. In Giselle Machline de Oliveira e Silva & Maria Marta Pereira Scherre (orgs.), *Padrões Sociolinguísticos: Análise de Fenômenos Variáveis do Português Falado na Cidade do Rio de Janeiro*, 119-145. Rio de Janeiro: Tempo Brasileiro/ Departamento de Língua e Filologia, UFRJ.
- Tomás, Gil *et al.* 2002. The peopling of São Tomé (Gulf of Guinea): Origins of slave settlers and admixture with the Portuguese. *Human Biology* 74. 397-411.
- White, Lydia *et al.* 2004. Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25. 105-133.
- Young, Richard. 1991. *Variation in Interlanguage Morphology*. New York, NY: Peter Lang.

APÊNDICE

Anexo 1: Informantes de Almoxarife: Aspetos sociais

Código do informante	Escalão etário	Sexo	Ensino	L1	Duração da entrevista
[CELESH3]	FE3 (+60)	Masc.	3ª Classe	SANTOME	01:04:44
[MANOH3]	FE3 (+60)	Masc.	Analfabeto	SANTOME	01:02:55
[SALVH3]	FE3 (+60)	Masc.	Analfabeto	SANTOME	01:10:53
[PRINCEM3]	FE3 (+60)	Fem.	Analfabeta	SANTOME	01:15:05
[MMDEUSM3]	FE3 (+60)	Fem.	Analfabeta	SANTOME	01:11:51
[MINISM3]	FE3 (+60)	Fem.	1ª Classe	SANTOME	01:08:16
[ABILH2]	FE2 (41-60)	Masc.	4ª Classe	SANTOME	01:08:59
[LUIH2]	FE2 (41-60)	Masc.	Analfabeto	SANTOME	01:25:57
[MAURIH2]	FE2 (41-60)	Masc.	3ª Classe	SANTOME	01:03:28
[CLOTIM2]	FE2 (41-60)	Fem.	4ª Classe	SANTOME	01:02:36
[FRANCM2]	FE2 (41-60)	Fem.	Analfabeta	SANTOME	00:56:28
[MANEVEM2]	FE2 (41-60)	Fem.	Analfabeta	SANTOME /PTG	01:01:10
[OSVALH1]	FE1 (20-40)	Masc.	4. Classe	SANTOME /PTG	01:00:28
[ZECAH1]	FE1 (20-40)	Masc.	4ª Classe	SANTOME /PTG	01:10:00
[CASTEH1]	FE1 (20-40)	Masc.	8º Ano	SANTOME /PTG	00:59:14
[ANTOM1]	FE1 (20-40)	Fem.	8º Ano	SANTOME /PTG	01:04:45
[DULCEM1]	FE1 (20-40)	Fem.	3ª Classe	SANTOME /PTG	00:56:20
[CARMOM1]	FE1 (20-40)	Fem.	8º Ano	SANTOME /PTG	01:06:19

Anexo 2: PA: Variáveis dependentes e independentes (análise atomística)

Variáveis dependentes	
	Presença de artigo definido
	Ausência de artigo definido
Variáveis independentes	
Variáveis independentes extralinguísticas	
1.	Sexo
2.	Faixa etária
3.	Escolaridade
Variáveis independentes estruturais e funcionais	
4.	Número do SN
5.	Presença de outros modificadores no SN
6.	Determinante no SN anafórico prévio
7.	Função sintáctica do SN
8.	Presença de preposição nos adjuntos do sintagma verbal e do SN
9.	Tipo de artigo definido
10.	Presença de advérbio locativo no SN
Variáveis independentes semânticas	
11.	Animacidade do substantivo
12.	Contabilidade do substantivo
13.	Referência do SN